The impact of technology investment on commercial bank's financial performance

Brenda Olivia Rangel Barrón[1] | Alexia Valeria Barrera Torres[1] | Luis Miguel Pulido[1]

[1] Faculty of Business and Finance, EGADE Business School, Mexico

Prof. Dr. Tauhidul Tanin
Comprehensive Bank and Non-bank Management

Abstract

The research objective of this report is to examine the impact of information technology on the financial performance of commercial banks. Data from six different commercial banks in Germany and Spain were obtained. As proxy for our focus variable, which is information technology investment, the variable of "Technology and Communications expense" was used. The dependent variable used to represent the financial performance of banks was "Net Interest". Additionally, two macroeconomic were chosen as control variables: growth rate of the GDP and inflation. Trough the development of a fixed effect panel regression modeling and interpretation process we find that investment in information technology has a week positive impact on the financial performance of a bank.. Our results are robust to a variety of econometric tests and specifications. As future work, we can consider different proxys to be used for the chosen focus variable and analyze the difference in results.

KEY WORDS

Financial performance, technological investment, net income, tech and communication expense, panel data, FE model, RE model

1 | Introduction

It has become clear that the digital revolution and social computing has caused a great impact in business. Information technology has become essential for becoming and staying competitive in today's market through the creation and support of various digital systems and applications. As technology continues to advance, the banking industry has seen the need to accelerate the acquisition of new tools and techniques to improve the customer experience, increase efficiency and improve security and avoid risks to customers. From mobile banking to blockchain, technology is changing the way we interact with financial institutions and the way we manage money.

One of the most important impacts of technology in banking is digital banking. Where the widespread use of smartphones and the Internet has become part of our daily routine, banks have developed cell phone applications and online platforms that allow customers to carry out a wide variety of financial transactions from anywhere. This has made banking more convenient and accessible and creating a convenience for customers, something that has never been seen before, one of the operations that can be carried out is verifying the balances of their accounts, transferring funds and paying invoices with just a few steps.

It has become a necessity for banks to manage, constantly change to improve, and master IT (Information Technology) to maintain the objective of ensuring their competitiveness in the market. The presence of IT eases the management and optimization of information, allowing an increase in efficiency in the bank. Taking all of these into consideration, it can be clearly expected that investment in Information Technology (IT) can have a great impact in the financial performance of a commercial bank, considering that technology has become an essential aspect, if not the core, of many operational processes within this sector. Focusing on the role that IT has had in the commercial banking sector as mentioned above; the following paper aims to examine the real impact that technology investment can has on the financial performance of a bank. A regression model was created with data from six different commercial banks from Germany and Spain, adding other macroeconomic variables to the process, and with this model the relationship between technology investment and financial performance will be examined. The paper will cover the explanation of variable selections, the methodology that was carried out, the tests that were executed, finalizing with the empirical results and conclusions.

2 | Literature Review

Among the existing literature there are studies that indicate either a negative or a positive statistically significant relationship between IT Investment and financial performance, either in the banking sector or in other sectors.

Some authors have found that financial performance is negatively affected by IT investment (Dandago, 2015) or that IT expenditure could lead to "existing network effects and reducing profits" in banks (Ho & Mallick, 2006), placing the role of IT more as a strategic necessity rather than a variable that would make way to a competitive advantage.

On the other hand, there have been also studies that concluded with a positive relationship between these two variables. Authors like Shu and Strassmann (2005) in their study "Does information technology provide banks with profit?", concluded that a bank's performance is indeed positively impacted by IT. Even Gupta, Raychaudhuri and Haldar (2018) in their study "Information technology and profitability: evidence from Indian banking sector" were also able to confirm this positive relationship in Indian banks.

Even though a few studies were found that stated there was no relationship between IT investment and profitability (Strassman, 1990), most of the literature indicates a relation between these two variables, either this relationship being positive or being a negative one.

With this in mind we can carry on with the chosen research objective, having a clear understanding that previous studies have indeed found either a positive or negative relationship between IT investment and financial performance, both in banking and non-banking sectors.

2| Dataset & Variables

2.1 | Dataset & Variables

Taking into consideration that the research objective is to examine the impact of information technology on the financial performance of commercial banking sectors, the dependent variable had to be a metric that was able to express relevant information of a bank's financial performance. A set of possible variables were considered; however, net income was chosen due to its direct relation with a bank's profitability, and thus, its overall financial performance.

Net income is easily calculated with the following formula:

$$Net\ Income = Total\ Revenue - Total\ Expenses$$

For this model, net income was obtained through the income statement. An important aspect to consider is that, unlike regular businesses, a bank's net income is primarily derived from the interest they earn on loans, subtracting the interest they usually pay on deposits.

The focus variable chosen for the development of this model had to be related with technology investment, to evaluate the impact of IT practices in a bank's performance. Due to a limited amount of data from public financial statements, the variable chosen was "Tech and Communications Expense", which can be easily found in financial statements. Even though it is not strictly related solely with technological investment, it is the near best available data that can measure in some way the reliability of a bank on technology.

On the other hand, control variables also had to be included to enhance the internal validity of the research that is being performed. A set of different control variables were considered, but finally only two variables were chosen: GDP Growth Rate and Inflation. Both variables have been very popular macroeconomic factor when examining the determinants of a bank's profitability. These variables are generally positively related with bank profitability.

The GDP growth rate implies an improvement in a country's economic performance, since as this growth rate increases, there is a greater population and business demand for banking activities. This greater demand for banking activities leads to a greater profitability trend for these entities. Profitability is further assured since during periods of economic growth, loan default cases reduce, ensuring profitability.

Inflation is not typically considered a "control variable" in the banking context, as it is an economic phenomenon that is beyond the direct control of individual banking institutions. However, inflation is an important economic variable that banks must consider when making decisions and formulating policies. By definition, inflation is the raise in process of goods and services in an economy over time. When it comes to bank management, inflation can have several impacts that banks must consider and manage, although they cannot control it directly. Although inflation is not a direct control variable for banks, they must consider it when developing strategies and policies. Central banks and financial authorities typically have a more direct role in managing inflation through monetary and fiscal policies.

2.2 | Chosen Bank Entities

BBVA (Spain): Banco Bilbao Vizcaya Argentaria or commonly known as BBVA, is a multinational Spanish banking group that has a very important presence in the global financial landscape.

Founded in 1857, they have grown to now become one of the largest and most important banks in Spain. Digital innovation is one of their main focuses and has played a key role in shaping modern banking, mainly due to its focus on technology and customer solutions.

Commerzbank (Germany): Commerzbank is a leading German Bank with rich history dating all the way back to the year 1970, making them one of the oldest banks currently operating. Their headquarters are in Frankfurt and is a main force not only in Germany but in European financial markets. They are recognized for their commitment to supporting businesses no matter their size and magnitude in a domestical and international way by building strong international relationships.

Deutsche Bank (Germany): They are one of the largest and most influential banks, not only domestically but worldwide. They offer a broad spectrum of financial services, which are investment and retail banking but also asset management. They have a significant global presence and play a crucial role in shaping the international financial landscape. They have achieved this goal by focusing on their international relations and this has been their goal since their beginnings.

DZBank (Germany): DZ Bank AG, or Deutsche Zentral-Genossenschaftsbank, is a central institution of Germany's cooperative financial network, which has more than 800 local cooperative banks in Germany, called Volksbanken Raiffeisenbanken. They operate as the central bank of these cooperative institutions, providing them with various services and offering a wide variety of financial services, including corporate banking, retail banking, capital markets, treasury services and international banking. Given its role as a central institution supporting a large network of cooperative banks, DZ Bank plays a crucial role in contributing to the stability of the German cooperative banking sector.

Hamburg Commercial Bank (Germany): Hamburg Commercial Bank AG, commonly known by the acronym HCOB, is a German private commercial bank headquartered in Hamburg, Germany, which is owned by a group of financial investors, including Cerberus Capital Management. It offers a wide variety of financial services to support the banking needs of businesses. It offers a variety of financial services, including corporate lending, trade finance, treasury services and capital market solutions, serving the financial needs of clients in various industries. The bank has special experience in sectors such as maritime transport, real estate, renewable energy, and aviation.

Santander (Spain): As well as BBVA, Banco Santander is a global banking entity that is headquartered in Spain. It is one of the largest banks in the world in terms of market capitalization and is present in numerous countries. It has presence is in Europe, America, Asia and Africa and offers a wide range of financial services including personal banking, commercial banking, private banking, asset management, insurance and investment services, and has stood out for its focus on technological innovation in the financial sector.

2.3| Data Bases

During the elaboration of this project, we relied on the databases "CapitalIQ" and "The World Bank". They both play critical roles in their respective domains making them leaders in their areas. They provide tools for businesses and to researchers to manage and analyze data in a more effective way in the realms of business management and market analytics.

The variables of "Net Income" and "Tech and Communications Expense" froom 2006 to 2022 were extracted from CapitalIQ. CapitalIQ is a market intelligence platform designed and developed by Standard & Poor's (S&P), which is a leading company in the financial sector, publishing world-recognized indexes like the S&P 500 and providing first-class financial information and analysis on both markets and industries (S&P Global Ratings | Spanish | Homepage | S&P Global Ratings, n.d.). CapitalIQ is specifically designed to provide data and research on companies, economies, markets, and industries. This database is one of the most reliable ones in the market and offers a wide variety of features: financial statements (balance sheet, profit and loss, cash flow, ratios, etc.), estimates, ESG data, latest news, ownership, and even integrates visualization tools. The database can organize financial data in a certain way, that allows us to have a more comprehensive and customizable reporting.

On the other hand, the control variables of "GDP Growth Rate" and "Inflation" were obtained from The World Bank, also from 2006 to 2022. The World Bank maintains a vast and comprehensive database which is used to serve as a valuable resource for different people such as researchers, policymakers and mainly the public in general. It encompasses a wide range of social, environmental, and more importantly economic indicators from countries around the globe. Its main objective is to provide financing and technical assistance to developing countries to reduce poverty and support economic and social development. As well as CapitalIQ, The World Bank database also has key features that help us have an easier experience and allows us to get reliable macroeconomic data and social indicators from countries worldwide. It contains customizable reports and dashboards that allows to have a better visualization and interpretation of data. With this feature the usability of the database gets enhanced.

Even though they both databases are different, they both play critical roles in their respective domains by providing tools for

businesses and researchers to manage and analyze data in a more effective way in topics of business management and global development. We decided to use these databases since they are both worldwide known and contain what is considered truly reliable data, which is of keen importance when developing a regression model.

3 | Methodology

3.1| Overview

As mentioned in this paper, the data used for this specific model consists in the net income and tech and communication expenses of six different commercial banks from Germany and Spain, from the years 2006 to 2022. Additionally, the GDP Growth rate as well as the Inflation variables were introduced as control variables.

We can describe the nature of the obtained data as being panel data. Panel data, by definition, is " [Panel data is] a two-dimensional concept, where the same individuums are observed repeatedly over different periods in time" (Hsiao, 2005). In other words, it is a combination of cross-sectional and time-series data, making if a multi-index data set with both year and observed individuum (which in this case is the bank's entity). This dataset nature plays an important role in the development of the model since a proper method had to be chosen to take into consideration both heterogeneity and endogeneity biases.

For this model, a plain static panel technique was executed due to its simplicity, since it does not include any lagged or future values as dependent variables. Taken this into consideration, our baseline equation is as follows:

$$NetInc_{it} = a + b_1 TechEx_{it} + b_2 GDP.GR_t + b_3 INF_t + e_{it}$$

Where $NetInc_{it}$ stands for net income, and is the dependent variable used as proxy for financial performance. $TechEx_{it}$ represents tech and communication expense, which was used as proxy for the focus variable of technology investment. Additionally, the baseline model also includes macroeconomic variables that act as control variables, $GDP.GR_t$ which stands for GDP growth rate and INF_t which stands for inflation.

3.2| Data Pre-Processing

The first step toward the development of a model was to pre-process the data. Since the variables that were chosen had a very distinct numerical range, a process of normalization was chosen. Normalization is a data preparation technique used to transform the features of a dataset to be on a similar scale, usually between 0 and 1.

The common formula used for data normalization is the following:

$$x_{\text{norm}} = \frac{x - \min(x)}{\max(x) - \min(x)}$$

In this case, the sklearn library in python was used for normalizing the dataset.

3.3| Algorithms and Tests

Different types of panel data regressions were evaluated during the model's development process. A decision had to be made between choosing PooledOLS or FE/RE regression methods.

PooledOLS simple Ordinary Least Squared model to be performed on panel data. Pooled Ordinary Least Squares (Pooled OLS) is a method of estimating the parameters of a linear regression model when data is structured in a panel format structure. Pooled OLS assumes that there is a common set of coefficients for all cross-sectional units and time periods. Its focus is on dependencies between the individuals (in this case banks) and requires having no correlation between independent variables and unobserved variables.

Fixed-Effects (FE) Model is a panel data regression model that accounts for individual-specific effects, also known as fixed effects. Panel data involves observations on multiple entities (cross-sectional units) over multiple time periods, and the Fixed Effects model is designed to control for individual heterogeneity by including fixed effects for each individual unit. It assumes that individual effects of unobserved, independent variables (or variables that are not considered in the model) are fixed (or constant) over time.

Random-Effects (RE) Model is a statistical model commonly used in the analysis of panel data. Panel data involves observations on multiple entities (in this case banks) over multiple time periods (in this case, from 2006 to 2022). The Random Effects model allows for unobserved individual-specific effects by assuming that these effects are random variables with specific distributional properties. Assumes that individual effects of unobserved, independent variables (or variables that are not considered in the model) are random over time.

To choose between these models, three main assumptions will be evaluated:

Homoskedasticity: It's a term used in statistics and econometrics to provide information on the constant variances of the errors in a regression model, which assumes that the variability of the errors or residuals is constant at all levels of the independent variables, in simple terms. The dispersion of the errors is constant at all levels of the independent variables, which implies that the variance of the errors is constant.

Non-autocorrelation: It indicates that the errors or residuals of the model are not correlated with each other over a period. In other words, the errors do not show systematic patterns of temporal correlation.

Endogeneity: Indicates a situation were an independent variable is correlated with the errors of the model. This correlation can be problematic because it makes causal interpretation of the relationship between the independent variable and the dependent variable difficult. In other words, endogeneity raises the possibility that the independent variable is simultaneously determined by other unobserved variables and affects the dependent variable.

If the assumption of either homoskedasticity or non-autocorrelation are violated, then it is more suitable to use a FE/RE model. Else, the PoolsOLS model is recommended. These assumptions will be tested using the following tests:

White – Test: General test for heteroskedasticity, also known as the White general test for heteroskedasticity, is another statistical test used to check for heteroscedasticity in a regression model. Like the Breusch-Pagan test, the White test assesses whether the variances of the errors in a regression model are constant across all levels of the independent variable(s). It is based on the idea that the squared residuals from the regression model will be correlated with the independent variables if there is indeed heteroskedasticity in the model. This test allows for the independent variable to have a nonlinear effect on the error variance. If the p-value is less than <0.05, then heteroskedasticity is indicated.

Breusch – Pagan – Test: Like the White-Test, the Breusch-Pagan test also measures the heteroskedasticity in a model. The Breusch-Pagan test, also known as the Cook-Weisberg test, is a statistical test used to check for heteroscedasticity in a regression model. The relationship between the variables in a dataset is considered to be heteroskedastic when the variation nof the errors (residuals) in a regression model is not constant. However, is only allows for the independent variable to have a linear effect on the error balance. If the p-value is less than <0.05, then heteroskedasticity is indicated.

Durbin – Watson Test: Statistical test used to assess the presence of autocorrelation in the residuals. This test has an outcome that ranges from 0 to 4, indicating either positive or negative autocorrelation. An answer near the mean (= 2) indicates there is little to no autocorrelation. A result from 0 – 2 means there is a positive autocorrelation (the nearer to 0 it is, the more positive correlation is presents). On the other hand, a result from 2 - 4 means there is a negative autocorrelation (the nearer to 4 it is, the more negative correlation is presents).

Additionally, the election between using a fixed or random effect model depends on if there is endogeneity or no endogeneity in the model. In other words, if the individual, unobserved heterogeneity is a constant or a random effect. For this, the Hausman - Test can be used.

Hausman Test: It is a statistic typically used in econometrics to help evaluate whether the random effects in a random effects model are uncorrelated with the regressors. The goal of the Hausman Test is to support researchers in choosing between a fixed effects model and a random effects model in the context of panel data analysis. The null hypothesis of the Hausman Test is that the random effects are not correlated with the regressors, what it recommends is that the random effects model be consistent. The alternative hypothesis is that the random effects are correlated with the regressors, indicating that the fixed effects model is a better fit.

4 | Empirical Results and Discussion

4.1| Normalization

The first step towards the model creation was data pre-processing. We started off with the creation of two simple PooledOLS, one using raw data and another normalizing the data beforehand. As expected, the model was able to explain more of the results with normalized data (Figure 1).

PooledOLS Estimation Summary

Dep. Variable:	Net Income	R-squared:	0.3393
Estimator:	PooledOLS	R-squared (Between):	0.4581
No. Observations:	102	R-squared (Within):	0.1023
Date:	Tue, Nov 28 2023	R-squared (Overall):	0.3393
Time:	03:32:21	Log-likelihood	-999.02
Cov. Estimator:	Clustered		
		F-statistic:	16.950
Entities:	6	P-value	0.0000
Avg Obs:	17.000	Distribution:	F(3,99)
Min Obs:	17.000		
Max Obs:	17.000	F-statistic (robust):	3.4648
		P-value	0.0191
Time periods:	17	Distribution:	F(3,99)
Avg Obs:	6.0000		
Min Obs:	6.0000		
Max Obs:	6.0000		

Parameter Estimates

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
Tech And Communications Expense	0.5838	0.7332	0.7962	0.4278	-0.8710	2.0386
GDP Growth Rate	324.86	162.69	1.9969	0.0486	2.0556	647.67
Inflation	625.27	319.42	1.9575	0.0531	-8.5358	1259.1

PooledOLS Estimation Summary

Dep. Variable:	Net Income	R-squared:	0.8697
Estimator:	PooledOLS	R-squared (Between):	0.9256
No. Observations:	102	R-squared (Within):	0.0035
Date:	Mon, Nov 27 2023	R-squared (Overall):	0.8697
Time:	23:46:39	Log-likelihood	19.108
Cov. Estimator:	Clustered		
		F-statistic:	220.21
Entities:	6	P-value	0.0000
Avg Obs:	17.000	Distribution:	F(3,99)
Min Obs:	17.000		
Max Obs:	17.000	F-statistic (robust):	31.347
		P-value	0.0000
Time periods:	17	Distribution:	F(3,99)
Avg Obs:	6.0000		
Min Obs:	6.0000		
Max Obs:	6.0000		

Parameter Estimates

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
Tech And Communications Expense	0.1286	0.1951	0.6594	0.5112	-0.2584	0.5156
GDP Growth Rate	0.5873	0.0631	9.3007	0.0000	0.4620	0.7126
Inflation	0.1283	0.0496	2.5848	0.0112	0.0298	0.2268

Figure 1: PooledOLS model using raw data (left) and normalized data (right)

4.2| Assumptions Evaluation and Model Selection

Afterwards, the violation of homoscedasticity and non-correlation had to be addressed. If either of these assumptions were to be violated, the PooledOLS model should be discarded, and a Fixed Effect or Random Effect model should be developed.

The residuals and predicted values of the PooledOLS model mentioned earlier were first plotted to see if we can find heteroskedasticity graphically. Since the plotted data spreads out [Figure 2], it is a clear indicator for growing variance and thus, for heteroskedasticity. With this we can state that the assumption of homoscedasticity is being violated. To enforce this conclusion, the White-Test and Breusch-Pagan-Test were also executed.

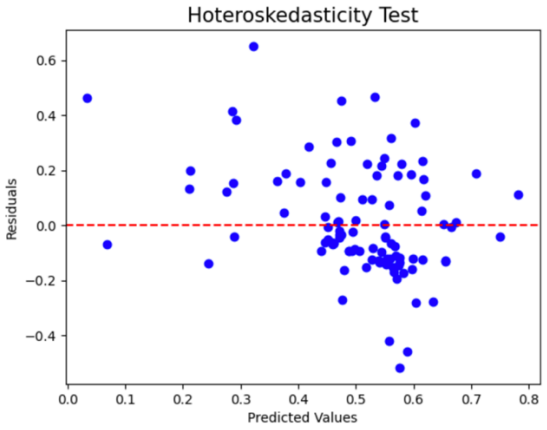

Figure 2: residual and predicted values graphically plotted

The White-Test and Breusch-Pagan-Test were additionally executed to reinforce the conclusion of the model being heteroskedastic. The p-value of both tests is lower than 0.05 (Figure 3). With these results we can clearly state that there is indeed a violation to the homoskedasticity assumption. Additionally, the assumption of non-correlation was also tested through the Durbin-Watson-Test (Figure 3), the result of .9206 clearly stating that there is a positive autocorrelation and thus, violation the second assumption as well. Since both assumptions (homoskedasticity and non-correlation) are violated, then either a fixed or random effect model is recommended.

	White-Test	Breusch-Pagan Test
LM-Stat	43.02702993	6.01562455
LM p-val	4.54E-10	0.014179759
F-Stat	36.11549459	6.267295612
F p-val	1.67E-12	0.013917519

Durbin-Watson Test	0.920465259

Figure 3: White-Test, Breusch-Pagan-Test, Durbin-Watson Test Results

To choose between a fixed or random effect model, the assumption to evaluate is endogeneity. In simple terms, if there is endogeneity present (if the null hypothesis is rejected), then the most suitable model is the fixed effect, else the random effect model should be chosen. However, the essence of choosing between these models relays on the assumption if the individual, unobserved heterogeneity is a constant or a random effect.

In this case, the Hausman-Test result (Figure 4) indicates that a random model should be used over a fixed model.

	Hausman Test
chi-squared	-3.538565596
degrees of freedom	3
p-value	1.00

Figure 4: Hausman-Test result

While the Hausman test indicates that a random effect model might be recommended for this research objective, in this case the random effect model was not giving us meaningful results. We believe that in this case a fixed effect panel regression is rather a more appropriate choice since it is more consistent with the stated research objective, more realistic with the gathered data, and more robust to the potential problems that a random effect model might not address. In general, we decided to move forward with the results of the fixed effect model due to the following reasons:

Correlation between unobserved heterogeneity and explanatory variables

A random effects model assumes that there is independency between unobserved heterogeneity (bank-specific characteristics like size) and explanatory variables (Rodriguez & Barbagallo, 2019)., but this might not hold true. A fixed effects model can control for these omitted variables, allowing for a more robust assessment of the relationship between IT investments and financial performance.

uno

Endogeneity problem between IT investments and financial performance

Even though the Hausman-Test states there is no endogeneity, the reality is that the assumption that IT investment has a casual, undirectional relation with Net income might not reflect the reality. Financial performance can indeed affect IT investments, or both variables might be affected by a third factor. In this case, the usage of a fixed effects model can help mitigate this endogeneity problem.

4.3| Model Evaluation and Interpretation

The model that was kept was the fixed effects panel model regression. In this case, the chosen evaluation metric to be used was the R^2. This metric measures "the proportion of variation in the dependent variable (Y) that can be explained by the independent variables (X)." (Verma, 2021). In simple terms, the goodness of fit of the model.

The R^2 obtained was of .8697 (Figure 5), meaning that 86.97% of the variation in the output variable can be explained by the chosen input variables, which is a good fit. The dependent variable for the model, which was "Tech and Communications Expense", obtained a p-value of 0.0727, meaning that it is statistically significant to the model at the 10% level. In other words, there is only a 7.27% chance that the relationship between the dependent variable and the independent variable is by chance alone. In this case, the coefficient of the dependent variable is positive, meaning that an increase Tech and Communications Expense (Investment in IT) usually leads to an increase in Net Income (Financial Performance). However, the magnitude of the coefficient is not as strong, suggesting that the effect of the focus variable on the dependent variable is not very strong.

As of the control variables, we find that both GDP Growth Rate and Inflation have a positive relation with Net Income, as expected. GPD Growth Rate is statistically significant and its relation to a bank's net income is strong (.5873), while Inflation is under an inferior level of significance and its relation is not as strong (.1283).

Overall, we can argue that there is indeed a positive impact in the financial performance of a bank when there is investment in technology. However, there can be other variables that are able to explain the variance in financial performance better than technology investment alone.

PanelOLS Estimation Summary

Dep. Variable:	Net Income	**R-squared:**	0.8697
Estimator:	PanelOLS	**R-squared (Between):**	0.9256
No. Observations:	102	**R-squared (Within):**	0.0035
Date:	Tue, Nov 28 2023	**R-squared (Overall):**	0.8697
Time:	19:45:54	**Log-likelihood**	19.108
Cov. Estimator:	Unadjusted		
		F-statistic:	220.21
Entities:	6	**P-value**	0.0000
Avg Obs:	17.000	**Distribution:**	F(3,99)
Min Obs:	17.000		
Max Obs:	17.000	**F-statistic (robust):**	220.21
		P-value	0.0000
Time periods:	17	**Distribution:**	F(3,99)
Avg Obs:	6.0000		
Min Obs:	6.0000		
Max Obs:	6.0000		

Parameter Estimates

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
Tech And Communications Expense	0.1286	0.0709	1.8142	0.0727	-0.0121	0.2693
GDP Growth Rate	0.5873	0.0526	11.165	0.0000	0.4830	0.6917
Inflation	0.1283	0.1097	1.1698	0.2449	-0.0893	0.3460

Figure 5: Fixed Effects Panel Model Regression

4.4 | Future Work

To examine the impact of technology investment on the financial performance of commercial banks, different variables can be considered, and various analytical approaches can be used.

In addition to inflation and GDP Growth Rate, which are the variables used in this analysis, other variables relevant to the banking sector can also be considered, such as:

Interest Rate: Since it can influence interest margins and profitability.

Levels of Competition in the Market: Competitive intensity can affect market share and profitability.

Financial Regulations: There are changes in regulation which could have significant effects on the operations and financial performance of banks.

Technology or Innovation Indices: This variable can obtain the general level of technological advancement in the industry.

Working with a low quantity of data, it is important to know how to choose which machine learning algorithm is the right one; one of the qualities is that they are robust and not prone to overfitting. There are other algorithms that can be considered to examine the impact of investment in technology on the financial performance of commercial banks:

***LASSO Regression*:** By its acronym (Least Absolute Shrinkage and Selection Operator), it is a regression technique that involves penalties to reduce the complexity of the model and avoid overfitting. It is especially useful when you have a limited number of observations.

Ridge Regression: Like LASSO, Ridge regression also incorporates penalties, but uses a norm penalty. It can be of great help in mitigating overfitting on small data sets.

***Support Vector Machines (SVM) with Linear Kernel*:** SVMs can be effective on small data sets, especially with a linear kernel. Choosing a linear kernel can help avoid overfitting and provide an easier-to-interpret model.

***XGBoost and LightGBM*:** It is important to emphasize that boosting algorithms like XGBoost and LightGBM can also be efficient on small data sets. These models are robust and have regularization options.

5 | Conclusions

In conclusion, information technology has a positive impact on the financial performance of commercial banks, but its relationship is not as strong.

Technologies have indeed opened huge roads for banks to acquire more clients and expand their market in a more global way. Automation and large datasets allow them to lower operating costs, manage credit risk in a more effective way and overall enhance customer value which will eventually lead to a better bank-customer relationship. In the topic of competitiveness banks now must excel in the areas of precision, efficiency, and speed in their response to customer needs which are becoming the topics that customers want better results. There is a clear reason to think that information technology has a strong impact of financial performance.

However, there might be other variables that have a stronger impact on the net income variable, or there might also be a more reliable proxy aside from “Technology and Communications Expense” to use for the focus variable of Investment in Technology. Additionally, other regression algorithms might be used to further experiment on this research objective.

6 | References

Bbva. (2023, September 11). *Información Corporativa BBVA*. BBVA NOTICIAS. https://www.bbva.com/es/informacion-corporativa/#historia-de-bbva

Brugger, B. (2021, December 26). A Guide to Panel Data Regression: Theoretics and Implementation with Python. Medium. https://towardsdatascience.com/a-guide-to-panel-data-regression-theoretics-and-implementation-with-python-4c84c5055cf8

Commerzbank AG - History. (n.d.). https://www.commerzbank.com/en/hauptnavigation/presse/150_jahre/150_jahre__zeitstrahl/geschichte.html

Dandago, K.I., Farouk, M., & Usman, B.K. (2012). Impact of investment in information technology on the return on assets of selected banks in Nigeria. International Journal of Arts and Commerce, 1(5).

Deutsche Bank, El Banco Alemán con Mucha Historia en España. Vozpópuli. (2023, April 1). https://www.vozpopuli.com/economia_y_finanzas/deutsche-bank-historia-banco-aleman.html

Gupta, S.D., Raychaudhuri, A., & Haldar, S.K. (2018). Information technology and profitability: evidence from Indian banking sector. International Journal of Emerging Markets.

Hamdan, M. N. (s/f). *The impact of information technology investment on the financial performance of the banks*. Abacademies.org. Recuperado el 26 de noviembre de 2023, de https://www.abacademies.org/articles/the-impact-of-information-technology-investment-on-the-financial-performance-of-the-banks.pdf

Hamdan, M. N., Gharaibeh, O. K., Quran, A. Z. A.-, & Nusairat, N. M. (2021). The impact of information technology investment on the financial performance of the banks. *Academy of Strategic Management Journal*, *20*(6S), 1–12. https://www.abacademies.org/articles/the-impact-of-information-technology-investment-on-the-financial-performance-of-the-banks-13768.html

Ho, S. and Mallick, S. (2006) The Impact of Information Technology on the Banking Industry Theory and Empirics. National Chengchi University, 15-23. - References - Scientific Research Publishing. (n.d.). https://www.scirp.org/(S(lz5mqp453edsnp55rrgjct55.))/reference/referencespapers.aspx?referenceid=2351371

Hsiao, C. (2005). *WHY PANEL DATA?* The Singapore Economic Review, 50(02), 143–154. https://doi.org/10.1142/s0217590805001937

Nuestra Historia: Sobre nosotros: Banco Santander. Sobre Nosotros | Banco Santander. (2023, February 3). https://www.santander.com/es/sobre-nosotros/nuestra-historia#2016-2010

Pradhan, K. (2023, febrero 22). *The impact of technology in banking*. Linkedin.com. https://www.linkedin.com/pulse/impact-technology-banking-kaustuva-pradhan/

Rodriguez, L. D., & Barbagallo, M. (2019). Mediterranean diet and longevity. In *Elsevier eBooks*. https://doi.org/10.1016/b978-0-12-801238-3.62178-5

Shu, W., & Strassmann, P. A. (2005). Does information technology provide banks with profit? Information & Management, 42(5), 781–787. https://doi.org/10.1016/j.im.2003.06.007

S&P Global Ratings | Spanish | Homepage | S&P Global Ratings. (n.d.). https://www.spglobal.com/ratings/es/

Strassmann, P. A. (1990). The business value of computers: an executive's guide: Information Economics Press.

Verma, U. (2021, December 13). R-Square(R2) and Adjusted R-Square - Analytics Vidhya - Medium. *Medium*. https://medium.com/analytics-vidhya/r-square-r%C2%B2-and-adjusted-r-square-63b77d4a6bd7

Wood, A., & Mc Conney, S. (s/f). *The impact of risk management on the financial performance of the commercial banking sector in Barbados*. Cert-net.com. Recuperado el 26 de noviembre de 2023, de https://cert-net.com/files/publications/conference/2017/AnthonyWoodandShaniseMcConneySHANWOOD2.pdf