Márcio Araújo da Gama

A high dimensional data approach to evaluate the effect of banking diversification on performance

Brasília

2020

Márcio Araújo da Gama

A high dimensional data approach to evaluate the effect of banking diversification on performance

Dissertação apresentada ao Curso de Mestrado Acadêmico em Economia, Universidade de Brasília, como requisito parcial para a obtenção do título de Mestre em Economia

Universidade de Brasília - UnB

Faculdade de Administração Contabilidade e Economia - FACE

Departamento de Economia - ECO

Programa de Pós-Graduação

Orientador: Prof. Dr. Daniel Oliveira Cajueiro

Brasília

2020

Márcio Araújo da Gama
A high dimensional data approach to evaluate the effect of banking diversification on performance/ Márcio Araújo da Gama. – Brasília, 2020-
52p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Daniel Oliveira Cajueiro

Dissertação (Mestrado) – Universidade de Brasília - UnB
Faculdade de Administração Contabilidade e Economia - FACE
Departamento de Economia - ECO
Programa de Pós-Graduação, 2020.

1. Bank Performance. 2. Portfolio Concentration. 2. High-Dimensional Econometrics. I. Orientador Prof. Dr. Daniel Oliveira Cajueiro II. Universidade de Brasília. III. Faculdade de Administração, Contabilidade e Economia - FACE. IV. Departamente de Economia IV. A high dimensional data approach to evaluate the effect of banking diversification on performance

Márcio Araújo da Gama

A high dimensional data approach to evaluate the effect of banking diversification on performance

Dissertação apresentada ao Curso de Mestrado Acadêmico em Economia, Universidade de Brasília, como requisito parcial para a obtenção do título de Mestre em Economia

---

Prof. Dr. Daniel Oliveira Cajueiro
Orientador

Prof. Dr. Benjamin Miranda Tabak
Convidado

Prof. Dr. José Guilherme de Lara Resende
Convidado

Brasília
2020

Abstract

This papers tests the effects of banks' loan portfolio concentration on performance and risk. We measure concentration by geographic region and economic activity. We find that both measures of concentration improve banks returns while geographic concentration reduces risk only for foreign banks and economic activity reduces the risk for all types of banks and foreign banks enjoy even less risk. Moreover, we test if changes in monetary policy affects differently focused and diversified banks. Our results do not show any significant difference between the two groups. We improve and make our models more robust by introducing control variables selected by Lasso. We collected a set of control variables used through banking literature to avoid possible omitted variables bias. All results are robust to the inclusion of the new control variables and one model is improved by this procedure.

Keywords: Bank performance, Portfolio concetration, High-Dimensional Econometrics.

Resumo

Nesse trabalho avaliamos como a escolha de uma banco entre alocar sua carteira de crédito de forma diversificada ou concentrada impacta sua performace e seu nível de risco de crédito. Foram usadas como medida de concentração a alocação da carteira de crédito por região geográfica e por atividade econômica. Os resultados mostram que bancos com a carteira concentrada possuem maiores ganhos e concentração geográfica reduz os riscos apenas para bancos estrangeiros e concentração por atividade econômica reduz os riscos para todos os tipos de bancos, porém a redução é maior para bancos estrangeiros. Além do mais, testamos se os efeitos da política monetária sobre a lucratividade dos bancos é diferente entre os bancos com carteira concentrada e diversificada. Os resultados não reportaram diferença entre os dois grupos. Todos os modelos passaram por um segunda rodade de estimação usando variáveis selecionadas pelo Lasso dentro de um grande conjunto de variáveis usadas na literatura de banking. Os resultados são robustos à inclusão de novas variáveis e um dos modelos apresenta uma melhora de resultado.

Palavras-chave: Performance Bancária, Concentração de Portfolio, Econometria em Alta Dimensão

Lista de tabelas

Sumário

1 Introduction

Banks in their role of financial intermediaries face several types of risk. They have to screen and monitor borrowers to minimize credit risk, analyze the economy and hedge themselves to avoid market risk and make proper management of their assets and liabilities to not incur in liquidity risk. Banks have to balance all these possible sources of distress to maintain their credibility as a financial institution while still begin a profitable business. Traditional portfolio theory argues that to reduce risk, one should diversify their investments. So, is it better for a bank to follow portfolio theory and be as diversified as possible? Or should they be focused, so they can gain more knowledge to screen and monitor better? We can find arguments in favor of both strategies. In this paper, we seek to answer this question by empirically testing the effects of loan portfolio concentration on the performance and credit risk of banks. We also test if changes in monetary policy affects differently the performance of diversified and focused banks and which strategy is more resilient to an economic recession.

We apply our empirical strategy to a dataset composed of Brazilian banks from 2016 to 2020. We measure loan portfolio concentration by economic activities and geographic regions to test its effects in four different ways. First, we test if focused banks are more efficient than diversified ones. To do this, we start with a base fixed effects regression. We find that banks that focus their loan portfolio have better returns. After that, we add to the model control variables selected by the least absolute shrinkage and selecting operator (Lasso) (TIBSHIRANI, 1996) among a rich set of variables used in banking literature. The results are robust to the addition of the new control variables. Second, we test if being more focused increases banks' credit risk. Our results show that focused banks enjoy less credit risk than diversified banks, though, for geographic diversification, only foreign banks sees reduction in their levels of risk. As for economic activity diversification, the reduction in risk is for all types of banks but the reduction is greater for foreign banks. Next, we test if focused banks are more sensitive to changes in monetary policy. We repeat the same procedure but we add to the regressions Brazil's basic interest rate - named Selic rate -, and a interaction effect between the Selic and a concentration measure. We find that interest rates affect negatively the performance of banks but this result is not robust to all especifications. Moreover, we do not find any significant difference in the effect that monetary policy has in the return of diversified and focused banks.

Lastly, we test which diversification strategy is more resilient to a downturn in the business cycle. To do this, we use the Covid-19 pandemic as an exogenous shock in economic activity and apply a difference in difference regression. We find that the return of banks that are focused in fewer sectors decreased less in the first two quarters of 2020.

What is the optimal strategy for banks? Should they focus or diversify their loan portfolio ? Knowing this, banks can improve their allocation of credit to optimize their risk-return. As mentioned before, we can find contradicting theories in the banking literature that argue in favor of both strategies. On one hand, banks should be as diversified as possible as it lowers the costs of financial intermediation and gives them credibility in their role of delegated monitors of credit while reducing their default risk (DIAMOND, 1984). Elsas, Hackethal, and Holzhäuser (2010) give two arguments in favor of diversification. The first is through economies of scope. Banks have a great amount of information about their customers which they could use for market research. Also, their operational structure allows them to expand their business with relative low costs. Second is that constant technological advances and deregulation keep changing the demand for financial products. Having a diversified portfolio of products would give a diversified bank more leverage to adapt to the new market.

However, when diversifying to new regions or economic sectors, banks may face high competition where they do not have expertise and end up facing losses. Moreover, a single sector in downturn is enough to bring a bank to default. Therefore, a diversified bank is facing more downside risk (WINTON, 1999). Ibragimov, Jaffee, and Walden (2011) argue that, if banks follow the diversification strategy, their assets will get similar over time causing an increase in systemic risk. Even if this strategy is beneficial to banks, the social cost would surpass its benefits. These critiques to the idea that banks should be diversified comes from a theory line that argues in favor of a focused strategy.

The theory that advocates for the focused strategy has its roots in corporate finance theory (DENIS; DENIS; SARIN, 1997; BERGER; OFEK, 1995). With a focus strategy, banks gain more expertise in the focused sector improving their screening and monitoring ability and also reduce agency costs. Futhermore, Winton (1999) argues that if technology reduces screening and monitoring costs, there is even less incentive to diversify. Considering the amount of data available now and new credit score methods and default probability models that benefits from larger data sets, the costs of screening and monitoring are as low as ever. Diversification would only be beneficial if loans had moderate levels of risk. When risks are low, there is

little incentive to diversify and if risks are high diversified banks are at greater risk due to their exposure to more sectors (WINTON, 1999).

Results in empirical research are also ambiguous. Several papers approach this problem measuring loan concentration by different criterias such as by industry sectors, geographic regions or products. An important paper that tries to answer this question is Acharya, Hasan, and Saunders (2006). Using data from Italian banks, they test the effect of portfolio concentration by industry and economy sector on bank profitability. They find that diversification, for banks with moderate levels of risk, reduces bank profitability while also increasing its risk. For banks with low levels of risk, diversification only lead to an inefficient tradeoff between risk and returns. In a similar way, Behr et al. (2007) tests this relation for German banks. Their results show that focused banks are more profitable but they are also exposed to more risk. Using an updated data set for German banks, Jahn, Memmel, and Pfingsten (2013) find that focused banks are less risky and expected loss is lower in the focused sector. In a more recent research, Beck, De Jonghe, and Mulier (2018) using a cross country panel of banks also find that a focused portfolio increases the bank stock return and reduces systemic risk and this relation is even stronger in the long run. As far as we know, Tabak, Fazio, and Cajueiro (2011) is the only other paper that investigates the effects of portfolio concentration on profitability for Brazilian banks. They measure loan concentration by industry for Brazilian banks find that a focused portfolio increases return and reduces risk.

Goetz, Laeven, and Levine (2016) investigate banks asset diversification by geographic region. They find that more diversified banks are exposed to less risks and risk reduction is greater when they venture to regions with different industrial structures and business cycles. Although banks enjoy less risk, the paper also finds that geographic expansions do not improve loan quality. Complementing this work, Chu, Deng, and Xia (2019) check the effect of geographic diversification on systemic risk. They find that diversification increases systemic risk. Furthermore, geographic diversification increases asset similarities between banks, which also increases systemic risk.

Considering that there is not a definitive answer for the concentration versus diversification problem, our paper differs from the others since we select among a rich set of variables using Lasso. With this, we can improve our inference by avoiding omitted variable bias. It's worth mentioning that in recent years, we are seeing more and more applied economics work where the traditional econometrics estimation techniques are combined with machine learning algorithms to improve forecasting performance (GARCIA; MEDEIROS; VASCONCELOS, 2017; MEDEIROS et al., 2019),

variable selection (BELLONI; CHERNOZHUKOV; HANSEN, 2014b; BELLONI; CHERNOZHUKOV; HANSEN; KOZBUR, 2016) and causal inference (ATHEY; IMBENS, 2016; WAGER; ATHEY, 2018). The ML toolkit can be useful especially when we are dealing with very large data sets or when the number of possible covariates exceed the number of observations. They also allow new possibilities like using text or image data to create variables never used before (VARIAN, 2014; ATHEY, 2018). Although these type of research is getting more popular, the banking literature has yet to see much work done using this approach (kimura). With this, we hope to shed new light on this problem.

Another innovation that our paper brings is a test to check if monetary policy affects focused and diversified banks differently and a test of resilience to adverse shocks in the business cycle. Empirical research in portfolio concentration for banks is mainly focused in risk and return problems. Therefore ,we hope to incentivise the proposal of new research questions regarding bank concentration.

The rest of this paper is organized as follow. In Section 2 we present our data set and describe in details our empirical strategy. In Section 3 we present our results and Section 4 concludes.

2 Data and methodology

2.1 Sample selection

We draw bank-specific balance sheet data and loan portfolio composition from Brazil's Central Bank's IF.data database. The composition of credit portfolio is given by sectors of industry, geographic region, rating, credit products and index rates. To avoid problems with missing values and outliers, we apply the following filters to the data:

- We exclude observations periods with missing data, negative or zero values for relevant balance sheet data;
- We drop banks that have less than 20% of their assets in lending operations;
- We drop development banks;
- We use banks that are classified as S1, S2 or S3 by Brazil's Central Bank.

Furthermore, if a bank has a gap in the sample, we drop the observations in the smaller group. The final sample is an unbalanced panel including 32 banks during the period 2016-2020 quarterly data totaling 515 observations. We start our sample in the second quarter of 2016 due to a change in how loan data is reported. In table 1 we present summary statistics for relevant balance sheet data.

Variable	Mean	SD	Min.	Max.
Assets	232046.6	456531.2	6071.1	1939420.0
Equity	17586.5	35371.8	747.6	144218.9
Deposits	74793.2	155631.1	88.6	764750.4
Total Loans	97506.1	196831.1	2199.6	720080.1
Net Revenue	689.7	1500.9	-1449.223	8168.235

All values are in millions of Brazilian Reais (R$)
Source: Brazil's Central Bank

Table 1 – Summary statistics for balance sheet data.

2.2 Concentration measure

To measure the concentration of the bank's credit portfolio we use the Hirshmann-Herfindahl index. This index is the sum of the square of relative exposure of bank b to sector i divided by its total exposure.

$$HHI_{bt} = \sum_{i=1}^{n} \left(\frac{NominalExposure_{bti}}{TotalExposure_{bt}}\right)^2 \tag{2.1}$$

We calculate the HHI of loans by economic activities as grouped by Brazil's Central Bank (listed in table 2) , and by Brazil's five geographic regions, henceforth HHI^{EA} and HHI^{Geo} ,respectively. The HHI is equal to 1 when there is complete concentration in 1 sector, and is equal to $\frac{1}{n}$ when there is perfect diversification among n sectors.

Economic activities
Agriculture, live stock, forestry fishing and aquaculture
Transformation industries
Construction
Public utility industry services
Extractive industries
Trade; repair of motor vehicles and motorcycles
Public administration, defense and social security
Transport, storage and mail
Others

Source: Brazil's Central Bank

Table 2 – Economic activities.

In table 3 we present summary statistics of HHI^{EA} and HHI^{Geo} for the whole sample and by type of ownership. We see that diversification by economic activity, compared to the average of 0.316 reported by Tabak, Fazio, and Cajueiro (2011) for the period of 2006-2009, has increased a little especially for state-owned banks, that are more diversified now (0.215 now compared to 0.31 then).

Variable	Total				Private banks				Foreign banks				State-owned banks			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
HHI^{Geo}	0.453	0.172	0.249	0.947	0.417	0.101	0.249	0.756	0.426	0.156	0.256	0.940	0.587	0.253	0.249	0.946
HHI^{EA}	0.303	0.130	0.149	0.870	0.309	0.150	0.183	0.870	0.347	0.108	0.188	0.631	0.215	0.042	0.149	0.340
Obs.	525				247				177				101			

Source: Brazil's Central Bank

Table 3 – Concentration measures by bank ownership.

2.3 Dependent variables

We use Return over Assets, ROA_{bt}, and Return over Equity ROE_{bt} of bank b on time t as dependent variables to account for banks performance, using revenue before taxes as return. We also use the ratio of Non-performing loans to total loans, Npl_{bt}, as our dependent variable in our second test, to account for bank's credit risk.

Variable	Mean	SD	Min.	Max.
ROA	0.0035	0.0054	-0.0437	0.0337
ROE	0.0345	0.0490	-0.4301	0.2118
Npl	0.0361	0.0278	0	0.1479
Size	17.65	1.62	15.62	21.38
Equity ratio	0.1051	0.0499	0.0171	0.2976
Loan ratio	0.4779	0.2161	0.0792	0.9762
Llp	0.0238	0.0151	0	0.0725

Source: Brazil's Central Bank

Table 4 – Summary statistics of dependant variables and Top 4 control variables used in the literature.

Variables	ROA	ROE	Npl	Size	Equity r.	Loan r.	Llp	HHI^{Geo}	HHI^{EA}
ROA	1.000								
ROE	0.8590	1.000							
Npl	-0.284	-0.344	1.000						
Size	-0.013	0.185	-0.166	1.000					
Equity ratio	0.224	-0.060	0.020	-0.486	1.000				
Loan ratio	0.259	0.062	-0.022	-0.331	0.594	1.000			
Llp	0.024	-0.086	0.475	-0.130	0.307	0.634	1.000		
HHI^{Geo}	-0.041	0.088	0.108	-0.105	-0.351	-0.479	-0.195	1.000	
HHI^{EA}	0.112	0.031	-0.338	-0.334	0.220	0.214	-0.187	-0.074	1.000

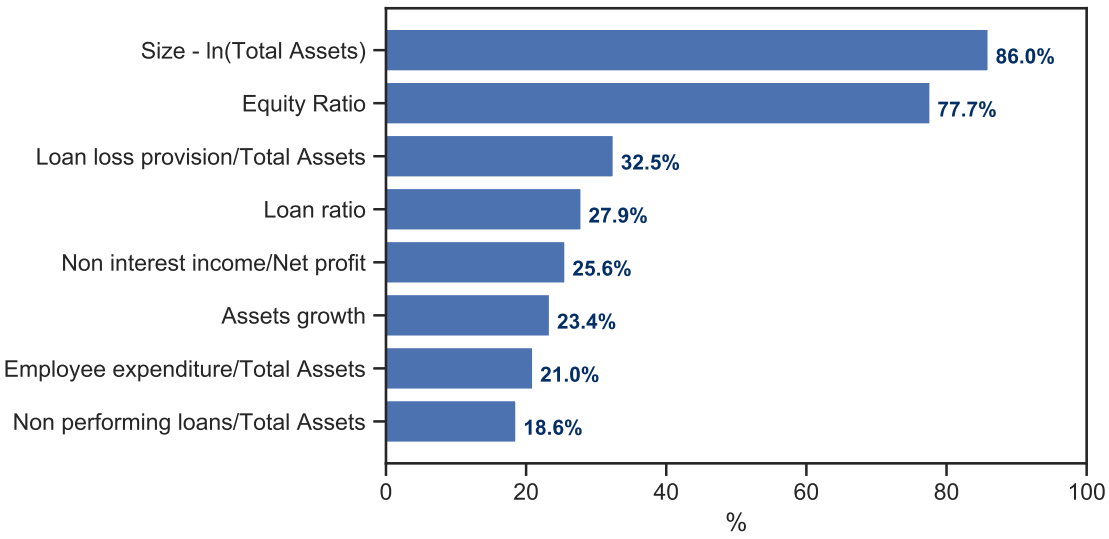
Source: Brazil's Central Bank

Table 5 – Cross-correlation table.

2.4 Control variables

Empirical research shows evidence favoring both diversification strategy and concentration strategy. These differences in results could be because of the country whose banks is sampled for the study. Different regulation laws changes how banks make their decisions. But this difference could also be due to specification problems and some works could be endued with omitted variable bias. We try to improve on past works by using a large set of possible control variables in our regression analysis.

For our possible control variables, we gather them from relevant empiric research on banking. We collect 66 control variables used in 43 banking papers. Some of them had to be dropped due to the impossibility to calculate them with our data set. All variables with a description and frequency of use are presented in the appendix. We show the top 8 variables used in fig. 1 and the percentage of papers they were used in. Even though LASSO could select variables from a number of covariates far exceeding the number of observations, the researcher is not exempt from preparing a set of variables with strong economic intuition, otherwise variable selection will perform poorly if the set of variables is improperly chosen (BELLONI; CHERNOZHUKOV; HANSEN, 2014a). We also create interaction variables between them to capture any non-linearities that may occur. In total, we end up with 1431 possible control variables. Since the number of covariates exceeds the number of observations, typical regression analysis is not feasible.



Source: Data compiled by authors

Figure 1 – This figure shows the top 8 control variables used in banking empirical research in the papers we sampled

We use the top 4 variables used as control variables in our base models. They are the Size of bank measured by the natural logarithm of total assets, the equity ratio, the ratio of total loans to total assets and the ratio of loan loss provision to total assets (Llp). **??** contains summary statistics of them plus our dependent variables and **??** shows cross-correlations between our base control variables, dependent variables and concentration measures.

2.5 Effects of concentration on performance

Starting with the question of the effect of loan portfolio concentration on bank's performance, we regress the performance variable on a concentration measure with the top 4 control variables listed in control. We use lagged explanatory variables to avoid endogeneity problems. The regression goes as follow:

$$Return_{bt} = \beta_{b0} + \beta_{b1} \times HHI_{bt-1} + \gamma \times CV_{bt-1} + \epsilon_{bt} \tag{2.2}$$

We'll use this regression as a base for further tests. A $\beta_{b1} >= 0$ indicates that focused banks perform better than diversified banks. After this, we add to this regression ownership variable interacting with the concentration variable to check if the effects of concentration on performance are different for private, state-owned and foreign banks. Then, we rerun both regressions including time fixed effects to account for macroeconomic conditions in each period.

$$\begin{aligned} Return_{bt} = \beta_{b0} + \beta_{b1} \times HHI_{bt-1} + \sum \alpha_j \times HHI_{bt-1} \times OwnershipDummy \\ + \gamma \times CV_{bt-1} + \epsilon_{bt} \end{aligned} \tag{2.3}$$

Considering that our set of control variables is reasonably large, we wish to select among them a subset of these variables that have strong explanatory power while maintaining a parsimonious model. To do this, we use ℓ_1 penalization, i.e lasso, since it shrinks the coefficient of some variables to 0, this technique has the properties of selecting variables (TIBSHIRANI, 1996). Since lasso alters the values of coefficients, it makes them biased. To correct for that we use post estimation of our regression, that is, we first run the Lasso regression then we run the regression using the variables with coefficients that weren't zeroed out in the first stage (BELLONI; CHERNOZHUKOV, 2013).

The lasso traditional specification is based on OLS, so its application to panel data models is not straightforward. Belloni, Chernozhukov, Hansen, and Kozbur

(2016) argue that using Lasso on a panel data will not properly account for individual specific heterogeneity and its direct application may result in a poor estimation and inference properties. Another problem is the assumption of independent observations does not apply to the majority of panel data used in economics. The authors propose a variation of the Lasso which allows a clustered covariance structure and partialing out the fixed effects (Cluster-Lasso). The procedure goes as following:

1. Run Cluster-Lasso on return against the concentration measure and the control variables used in the basic model. We allow these variables to be unpenalized during Lasso.

$$Return_{bt} = \beta_{b0} + \beta_{b1} \times HHI_{bt-1} + \gamma \times CV_{bt-1} + \eta \times PV_{bt-1} + \epsilon_{bt} \tag{2.4}$$

   Where PV is a vector of possible variables to be selected by Lasso.

2. The we run the regression adding the vector of variables LV that is composed by the variables of PV whose coefficient from the first stage is different than 0.

$$Return_{bt} = \beta_{b0} + \beta_{b1} \times HHI_{bt-1} + \gamma \times CV_{bt-1} + \eta \times LV_{bt-1} + \epsilon_{bt} \tag{2.5}$$

We repeat this process using ownership dummy variables and with time fixed effects.

2.6 Concentration and risk

Next we check the effects of concentration on risk. Traditional portfolio theory says that one should diversify to minimize risk. But is it true for banks? Banks can reduce their risk by selecting better who to lend to and monitoring them. A focused bank has a better understanding of the sectors they are focused on, therefore they can select and monitor better which may decrease the credit risk the bank faces.

To test this, we use the ratio of non-performing loans to total loans as our measure of credit risk and we regress it onto the concentration measure, the bank's size, equity ratio and a return measure as control variables. Again we use lagged variables to avoid any problems of endogeneity. The regression goes as follow:

$$Risk_{bt} = \beta_{b0} + \beta_{b1} \times HHI_{bt-1} + \gamma \times CV_{bt-1} + \epsilon_{bt} \tag{2.6}$$

For this regression, a $\beta_{b1} >= 0$ indicates that focused banks are riskier than diversified banks. To complement this regression, we test the relation when we add ownership dummy variables and time fixed effects:

$$Risk_{bt} = \beta_{b0} + \beta_{b1} \times HHI_{bt-1} + \sum \alpha_j \times HHI_{bt-1} \times OwnershipDummy + \gamma \times CV_{bt-1} + \epsilon_{bt} \tag{2.7}$$

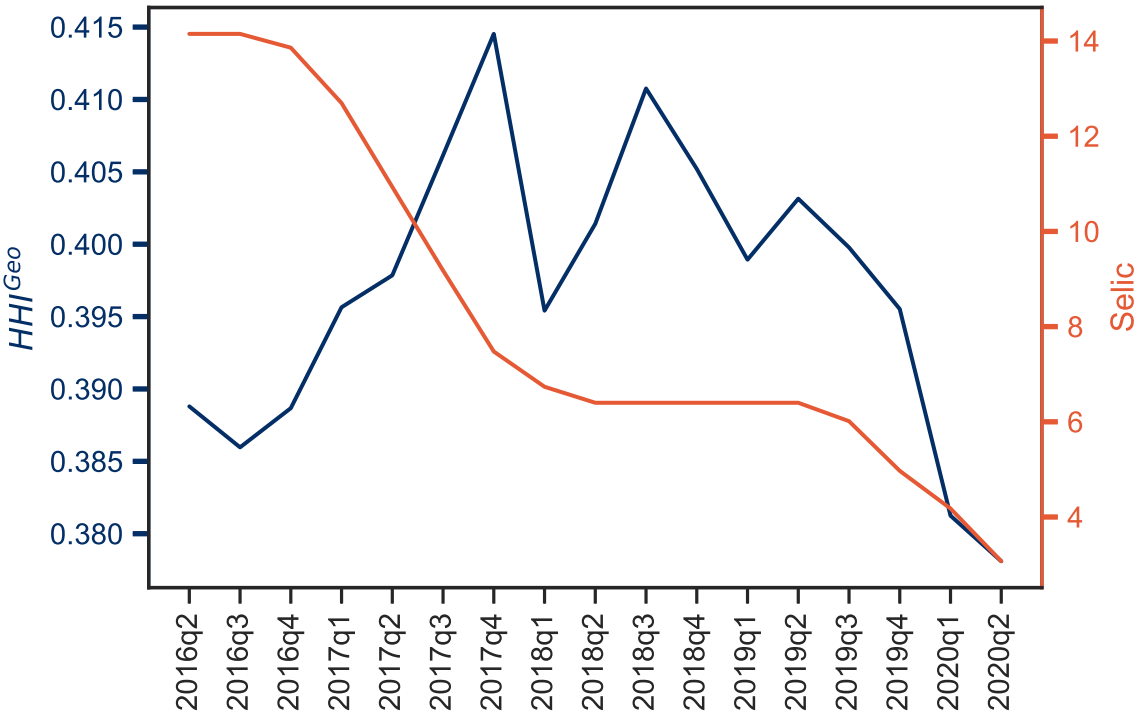
Following the same procedure as the previous test, we again use Post estimation with variables selected by Cluster-Lasso.

2.7 Monetary policy

Banking activity is greatly affected by changes in the monetary policy. Be it a change in the short term interest rate or compulsory deposits, banks have to adjust their strategy accordingly. An interesting question concerning changes in monetary policy is how does it affect bank's performance, especially changes in the basic interest rate. Samuelson (1945) argues that an increase in interest rates benefits the banking sector overall. The only case the author can see where the banking sector would be left worse with an increase in interest rates is if the average time period of their earning assets exceeded the average time period of their liabilities.

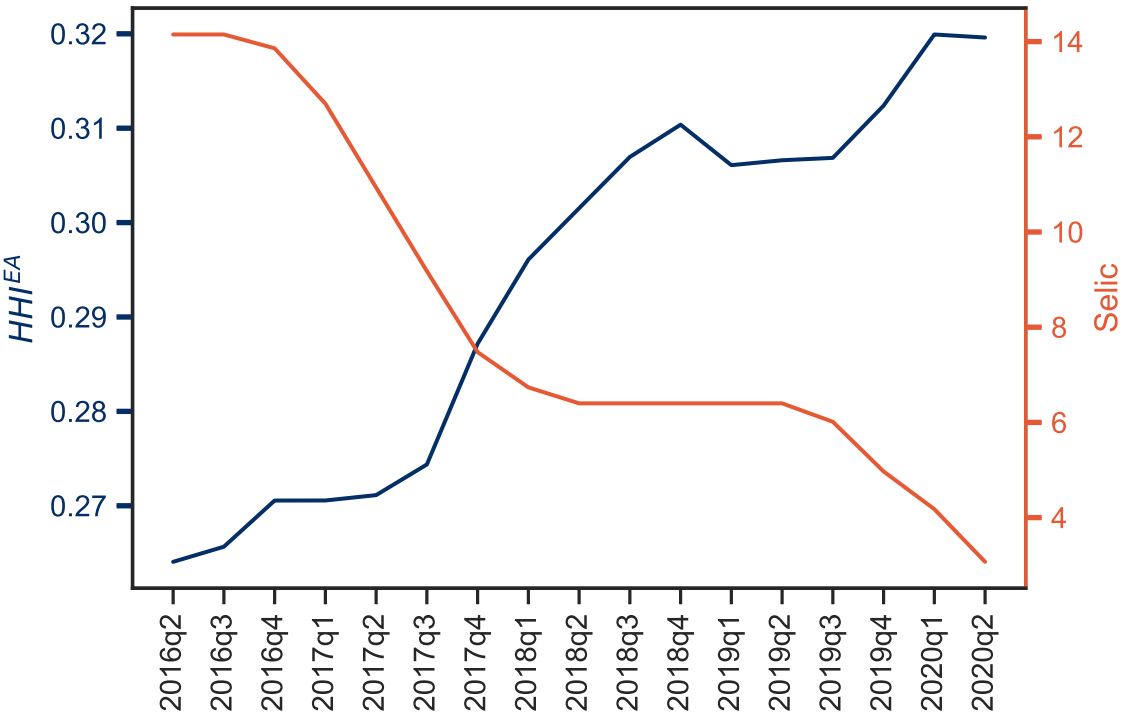
A paper that tries to answer this question empirically is Borio, Gambacorta, and Hofmann (2017). Using a panel of banks from OCDE countries, they investigate empirically the effects of monetary policy on bank profitability. They find that the short term interest rate has a positive impact on bank profitability, but they also impact positively loan loss provisions. Their results shows that the relation is non-linear, indicating that the effects of a increase in interest rates are larger at lower levels and that long periods of low interest rates are detrimental do bank profitability.

In the last years, Brazil has seen a decrease in it's basic interest rate, the Selic rate, to all time lows. In fig. 2 and fig. 3 we show the evolution of the quarterly average of concentration measures over the sample period. We plot the Selic, averaged by quarter, together with the concentration measures to see how banks adapted their portfolios in the presence of lower interest rates. We see that banks, for geographic loan concentration, while there was a increase in concentration at the start, it was eventually decreased again when interest rates were lowered below 6% a year. For concentration by economic activity, banks got more focused when interest rates were getting lower.



Source: Brazil's Central Bank

Figure 2 – This figure shows the average of portfolio concentration by geographic region with the Selic rate



Source: Brazil's Central Bank

Figure 3 – This figure shows the average of portfolio concentration by economic activity region with the Selic rate

Although this seen like a natural question to be investigated, Borio, Gambacorta, and Hofmann (2017) says there is very few empirical work investigating the effects of monetary policy in bank's performance. We take a step further and evaluate if monetary policy effects are different between focused and diversified banks. To test this we run the following regression:

$$Return_{bt} = \beta_{b0} + \beta_{b1} \times HHI_{bt-1} + \beta_{b2} \times Selic_{t-1} + \beta_{b3} \times Selic_{t-1} \times HHI_{bt-1} + \gamma \times CV_{bt-1} + \epsilon_{bt} \tag{2.8}$$

Where CV is a vector of control variables. We again use the top 4 control variables as listed in Control variables.

From the results we can draw 4 different conclusions:

- If $\beta_{b2} > 0$ and $\beta_{b3} > 0$ then an increase in interest rates benefits banks and focused banks have more gains than diversified ones.
- If $\beta_{b2} < 0$ and $\beta_{b3} > 0$ then an increase in interest rates lowers bank's profits but focused banks are not as impacted as diversified banks.
- If $\beta_{b2} > 0$ and $\beta_{b3} < 0$ then an increase in interest rates benefits banks and focused banks have less gains than diversified ones.
- If $\beta_{b2} < 0$ and $\beta_{b3} < 0$ then an increase in interest rates lowers bank's profits and focused banks are more impacted than diversified banks.

2.8 Resilience to an economic downturn

Another interesting question to be answered is how banks respond to downturns in the economy depending on how they allocate their loan portfolio. To test this, we use the Covid-19 pandemic as an exogenous shock in economic activity to

check if diversified or focused banks are more resilient to a recession. Diversified banks, on one hand, are exposed to more sectors meaning it could be less impacted if one particular sector is not performing well but other sectors could compensate for it. But, if the recession is affecting most economic sectors, a diversified bank could see a sharp decrease in their revenue. On the other hand, focused banks could be well worse off specially if the few sectors they lend to are the ones most impacted, and vice versa. Even though a recession reduces economic activity, a lot business depend on banks for credit so they can maintain their basic operations through a rougher time. In this case, focused banks could take advantage of this situation and use their expertise to provide better services and products as to, at least, not suffer greater losses.

We use a difference in difference (DiD) regression as our empirical tool in this test. This is unconventional for did since both of our groups - diversified and focused banks- receive the treatment, in this case, the recession caused by the Covid-19 pandemic. Our test aims to verify if the impact of the recession was different for focused and diversified banks. To use as our focused banks group, we selected the upper quartile of mean HHI through the year 2019. The rest is the diversified banks groups. We use the period from the first quarter of 2019 until the second quarter of 2020, where both 2020 quarters are our post treatment period. The basic DiD regression we run is as follow:

$$\begin{aligned} Return_{bt} = \beta_{b0} + \beta_{b1} \times Post_t + \beta_{b2} \times Post_t \times Focused_b \\ +\gamma \times CV_{bt-1} + \epsilon_{bt} \end{aligned} \tag{2.9}$$

In this regression, our main interest is β_{b2} that could show that diversified banks were less affected during recession if its value is negative, or that focused banks were less affected in the opposite case.

3 Results

3.1 Effects of concentration on performance

The results of first empirical test are reported in table 6 and table 7, full regression tables are presented in the appendix. Our results shows strong evidence that banks with focused portfolio perform better than diversified ones, for both geographic and economic activity concentration measures. In all specifications, the coefficient for the concentration measure was positive indicating that being focused is better for returns than diversification. This result is consistent with the theoretical model of Winton (1999) and the findings of Acharya, Hasan, and Saunders (2006) and Tabak, Fazio, and Cajueiro (2011). The result is robust to the addition of ownership variables and time fixed effects. In regression 4 of table 6, there is some evidence of the effect of geographic diversification on performance is not as strong for foreign banks, though this result is not consistent through all the tests. As for economic activity concentration, there is some evidence that state-owned banks have more benefits from concentrating their loan portfolio. This could be because most state owned banks are regional so they will perform better if they focus in the most prominent sectors in that region. This result differs from the one Tabak, Fazio, and Cajueiro (2011) finds in which state owned banks would actually benefit less from a focused strategy.

The model with post estimation with control variables selected by Cluster Lasso are presented in table 8 and table 9. Results from the base model are sustained with the additional controls selected by Lasso. For the models using ROA as return and with concentration measured by economic activity no variables are selected by Lasso. This could indicate that our set of control variables, even though it was assembled with common control variables used in the literature, still lacks in explanatory power. In regression 4 from table 8, the aforementioned possible weaker effect of geographic concentration on the performance of foreign banks vanishes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	ROA	ROA	ROA	ROA	ROE	ROE	ROE	ROE
HHI_{t-1}^{Geo}	0.0143***	0.0166***	0.0143***	0.0221***	0.185***	0.191***	0.182***	0.243***
	(0.00319)	(0.00568)	(0.00390)	(0.00533)	(0.0285)	(0.0552)	(0.0344)	(0.0602)
$HHI_{t-1}^{Geo} \times State - Owned$		0.0104		0.0130		0.175		0.0785
		(0.0155)		(0.0232)		(0.236)		(0.220)
$HHI_{t-1}^{Geo} \times Foreign$		-0.00471		-0.0151*		-0.0202		-0.116
		(0.00864)		(0.00782)		(0.0815)		(0.0873)
Constant	0.0173	0.0116	0.0349	0.0213	0.380	0.334	0.258	0.154
	(0.0261)	(0.0294)	(0.0297)	(0.0338)	(0.281)	(0.315)	(0.257)	(0.293)
Observations	492	492	492	492	492	492	492	492
R-squared	0.054	0.055	0.144	0.147	0.039	0.040	0.128	0.130
Number of Banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6 – Relationship between return and concentration - Base model - Geographic concentration - Fixed Effects Estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	ROA	ROA	ROA	ROA	ROE	ROE	ROE	ROE
HHI_{t-1}^{EA}	0.00750***	0.00585***	0.00692***	0.00564***	0.0753***	0.0621***	0.0713***	0.0603***
	(0.00163)	(0.00150)	(0.00139)	(0.00127)	(0.0147)	(0.0130)	(0.0147)	(0.0139)
$HHI_{t-1}^{EA} \times State - Owned$		0.0578**		0.0510**		0.455		0.453
		(0.0232)		(0.0244)		(0.279)		(0.267)
$HHI_{t-1}^{EA} \times Foreign$		0.00555		0.00343		0.0454		0.0272
		(0.00571)		(0.00580)		(0.0374)		(0.0333)
Constant	0.0234	0.0233	0.0350	0.0267	0.437	0.437	0.242	0.169
	(0.0275)	(0.0269)	(0.0314)	(0.0330)	(0.298)	(0.294)	(0.293)	(0.311)
Observations	492	492	492	492	492	492	492	492
R-squared	0.060	0.069	0.148	0.154	0.038	0.045	0.125	0.132
Number of Banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 7 – Relationship between return and concentration - Base model - Economic activity concentration - Fixed Effects Estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	ROA	ROA	ROA	ROA	ROE	ROE	ROE	ROE
HHI_{t-1}^{Geo}	0.0151***	0.0162***	0.0143***	0.0209***	0.179***	0.192***	0.178***	0.243***
	(0.00273)	(0.00448)	(0.00336)	(0.00458)	(0.0282)	(0.0562)	(0.0336)	(0.0606)
$HHI_{t-1}^{Geo} \times State - Owned$		0.0174		0.0162		0.179		0.0952
		(0.0157)		(0.0227)		(0.241)		(0.225)
$HHI_{t-1}^{Geo} \times Foreign$		-0.00298		-0.0129		-0.0325		-0.125
		(0.00853)		(0.00775)		(0.0803)		(0.0872)
Constant	0.0241	0.0187	0.0217	0.00984	0.415	0.358	0.306	0.194
	(0.0253)	(0.0294)	(0.0259)	(0.0300)	(0.284)	(0.318)	(0.270)	(0.302)
Observations	492	492	492	492	492	492	492	492
R-squared	0.088	0.088	0.171	0.174	0.061	0.061	0.146	0.148
Number of Banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Selected Variables	1	2	1	1	2	2	2	2

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8 – Relationship between return and concentration - With control variables selected by Cluster Lasso - Geographic concentration - Fixed Effects Estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	ROA	ROA	ROA	ROA	ROE	ROE	ROE	ROE
HHI_{t-1}^{AE}	0.00750***	0.00640***	0.00692***	0.00608***	0.0737***	0.0615***	0.0704***	0.0600***
	(0.00163)	(0.00143)	(0.00139)	(0.00120)	(0.0147)	(0.0130)	(0.0145)	(0.0137)
$HHI_{t-1}^{AE} \times State - owned$		0.0583**		0.0506**		0.457		0.454*
		(0.0238)		(0.0242)		(0.278)		(0.268)
$HHI_{t-1}^{AE} \times Foreign$		0.00326		-3.36e-05		0.0375		0.0224
		(0.00487)		(0.00470)		(0.0364)		(0.0335)
Constant	0.0234	0.0233	0.0350	0.0175	0.473	0.473	0.293	0.219
	(0.0275)	(0.0241)	(0.0314)	(0.0294)	(0.306)	(0.305)	(0.313)	(0.332)
Observations	492	492	492	492	492	492	492	492
R-squared	0.060	0.085	0.148	0.165	0.060	0.067	0.144	0.151
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lasso	Single	Single	Single	Single	Single	Single	Single	Single
Selected Variables	0	1	0	3	2	2	2	2

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9 – Relationship between return and concentration - With control variables selected by Cluster Lasso - Economic activity concentration - Fixed Effects Estimation

3.2 Effects of concentration on risk

With the results from the previous section, we have strong evidence that performance is improved by having a focused portfolio we have to check how does it effect risk. Classic portfolio theory says that we have to diversify to decrease risk, and banks want to avoid unnecessary risk as a bank's downturn or bank runs can be very costly to society as a whole. So we proceed with our test to check the effects of concentration on the banks credit risk.

Our base model results presented in table 10 and table 11 shows that having a focused credit portfolio decreases banks' credit risk measured by the proportion of non-performing loans to total loans for both of ours concentration measures . Again, this result is consistent with the theory of Winton (1999) and other empirical results for economic activity concentration and geographic concentration, including the previous results for Brazilian banks of Tabak, Fazio, and Cajueiro (2011). Considering the results for Geographic diversification of Chu, Deng, and Xia (2019), their results show that geographic concentration, although decreases the banks' risk, do not improve loan quality. We consider that a reduction on NPL is in fact a improvement in loan quality. With the addition of ownership variables, the effect of the concentration measure on risk is lost and only the interaction between the concentration measure and Foreign ownership has statistical significance. This could indicate that the first result was mainly driven by foreign banks. Foreign banks are at greater disadvantage when it comes to information about specificities about regional differences and the operation of each sector. So, naturally, foreign banks can gain expertise and knowledge faster by specializing in fewer sector and regions.

When we add variables selected by Lasso, the models for geographic concentration (table 12) do not see much change. But the the models for economic activity concentration (table 13), the inclusion of the new control variables makes the concentration measure statistically significant for explaining bank credit risk even with the inclusion of ownership variables. This is a improvement over the models presented in table 11. This confirms that being focused in fewer sector does indeed reduce banks' risk, not only to foreign banks, thought the decrease in risk for foreign banks is higher. This result confirms that gains in expertise improve screening and monitoring for all types of banks, and it helps even further for foreign banks for the reasons aforementioned.

VARIABLES	(1) NPL	(2) NPL	(3) NPL	(4) NPL	(5) NPL	(6) NPL	(7) NPL	(8) NPL
HHI_{t-1}^{Geo}	-0.141***	-0.0486	-0.145***	-0.0493	-0.140***	-0.0484	-0.144***	-0.0491
	(0.0459)	(0.0622)	(0.0489)	(0.0543)	(0.0465)	(0.0629)	(0.0493)	(0.0548)
$HHI_{t-1}^{Geo} \times State - Owned$		0.0245		-0.0319		0.0263		-0.0325
		(0.153)		(0.142)		(0.151)		(0.140)
$HHI_{t-1}^{Geo} \times Foreign$		-0.174*		-0.178**		-0.174*		-0.177**
		(0.0909)		(0.0840)		(0.0907)		(0.0838)
ROA_{t-1}	-0.225**	-0.233**	-0.205*	-0.233**				
	(0.0889)	(0.0873)	(0.103)	(0.0944)				
ROE_{t-1}					-0.0195	-0.0203	-0.0172	-0.0205
					(0.0139)	(0.0141)	(0.0155)	(0.0152)
Constant	0.587**	0.443**	0.353	0.200	0.584**	0.441**	0.347	0.193
	(0.216)	(0.201)	(0.215)	(0.180)	(0.216)	(0.201)	(0.216)	(0.182)
Observations	492	492	492	492	492	492	492	492
R-squared	0.175	0.204	0.214	0.242	0.173	0.202	0.213	0.240
Number of Banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 10 – Relationship between risk and concentration - Base model - Geographic concentration - Fixed Effects Estimation

VARIABLES	(1) NPL	(2) NPL	(3) NPL	(4) NPL	(5) NPL	(6) NPL	(7) NPL	(8) NPL
HHI_{t-1}^{EA}	-0.0446**	-0.00869	-0.0449**	-0.0107	-0.0446**	-0.00939	-0.0448**	-0.0113
	(0.0200)	(0.0131)	(0.0201)	(0.0132)	(0.0200)	(0.0130)	(0.0203)	(0.0132)
$HHI_{t-1}^{EA} \times State - Owned$		-0.184*		-0.136		-0.182*		-0.132
		(0.102)		(0.103)		(0.102)		(0.104)
$HHI_{t-1}^{EA} \times Foreign$		-0.243***		-0.238**		-0.241***		-0.236**
		(0.0874)		(0.0869)		(0.0878)		(0.0871)
ROA_{t-1}	-0.217**	-0.235***	-0.199*	-0.238***				
	(0.0811)	(0.0687)	(0.105)	(0.0794)				
ROE_{t-1}					-0.0229*	-0.0177	-0.0204	-0.0176
					(0.0121)	(0.0105)	(0.0149)	(0.0121)
Constant	0.554**	0.549***	0.377*	0.414***	0.551**	0.547***	0.369*	0.407***
	(0.234)	(0.168)	(0.188)	(0.129)	(0.232)	(0.169)	(0.189)	(0.133)
Observations	492	492	492	492	492	492	492	492
R-squared	0.139	0.293	0.176	0.318	0.139	0.290	0.175	0.316
Number of Banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 11 – Relationship between risk and concentration - Base model - Economic Activity concentration - Fixed Effects Estimation

VARIABLES	(1) NPL	(2) NPL	(3) NPL	(4) NPL	(5) NPL	(6) NPL	(7) NPL	(8) NPL
HHI_{t-1}^{Geo}	-0.116**	-0.0504	-0.124***	-0.0493	-0.0993**	0.00443	-0.118**	-0.0491
	(0.0428)	(0.0589)	(0.0415)	(0.0543)	(0.0449)	(0.0450)	(0.0442)	(0.0548)
$HHI_{t-1}^{Geo} \times State - Owned$		0.102		-0.0319		0.0194		-0.0325
		(0.135)		(0.142)		(0.120)		(0.140)
$HHI_{t-1}^{Geo} \times Foreign$		-0.171*		-0.178**		-0.195***		-0.177**
		(0.0850)		(0.0840)		(0.0642)		(0.0838)
ROA_{t-1}	-0.330***	-0.193*	-0.317**	-0.233**				
	(0.103)	(0.0980)	(0.122)	(0.0944)				
ROE_{t-1}					-0.0312	-0.0301	-0.0374	-0.0205
					(0.0230)	(0.0235)	(0.0241)	(0.0152)
Constant	0.545***	0.352**	0.422*	0.200	0.501***	0.334**	0.381*	0.193
	(0.166)	(0.161)	(0.211)	(0.180)	(0.147)	(0.145)	(0.207)	(0.182)
Observations	492	492	492	492	492	492	492	492
R-squared	0.261	0.251	0.292	0.242	0.295	0.330	0.285	0.240
Number of Banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Selected Variables	3	2	3	0	3	2	1	0

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 12 – Relationship between risk and concentration - With control variables selected by Cluster Lasso - Geographic concentration - Fixed Effects Estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	NPL	NPL	NPL	NPL	NPL	NPL	NPL	NPL
HHI_{t-1}^{EA}	-0.0428**	-0.0165**	-0.0469**	-0.0164*	-0.0425**	-0.0170**	-0.0460**	-0.0169*
	(0.0183)	(0.00703)	(0.0179)	(0.00861)	(0.0184)	(0.00706)	(0.0181)	(0.00848)
$HHI_{t-1}^{EA} \times State - Owned$		-0.0557		-0.102		-0.0496		-0.0910
		(0.0751)		(0.0962)		(0.0737)		(0.0966)
$HHI_{t-1}^{EA} \times Foreign$		-0.198**		-0.211**		-0.195**		-0.208**
		(0.0758)		(0.0768)		(0.0764)		(0.0768)
ROA_{t-1}	-0.0393	-0.262**	-0.298**	-0.330***				
	(0.202)	(0.112)	(0.146)	(0.111)				
ROE_{t-1}					-0.0102	-0.0269	-0.0407	
					(0.0327)	(0.0200)	(0.0259)	
Constant	0.463***	0.426***	0.418**	0.435***	0.462***	0.422***	0.396**	0.420***
	(0.157)	(0.0926)	(0.189)	(0.131)	(0.157)	(0.0916)	(0.179)	(0.132)
Observations	492	492	492	492	492	492	492	492
R-squared	0.282	0.411	0.285	0.387	0.283	0.410	0.281	0.386
Number of Banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Selected Variables	3	3	2	1	3	3	1	1

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 13 – Relationship between risk and concentration - With control variables selected by Cluster Lasso - Economic Activity concentration - Fixed Effects Estimation

3.3 Monetary policy and loan concentration

The results for our test of effects of monetary policy in focused and diversified banks are presented in table 14 and table 15. For all regressions where the Selic rate coefficient is significant, the relation with return seens to be negative. This result is counterintuitive to the operation of banks. Although higher interest rates reduces the demand for credit, banks compensate that by allocating more of its assets in the treasury. This results also goes against the argument made by Samuelson (1945) and the findings of Borio, Gambacorta, and Hofmann (2017). This could be due to the time period of our sample where the Selic rate received continuous cuts throughout the period and banks' profits still increased every year. Although this could explain why the Selic rate has a negative coefficient, the models which we include time fixed effects should have controlled for other causes of the increase in profits.

The models with the interaction effect, the Selic rate times the concentration measure is not significant in any of the specifications. Also, specially for economic activity concentration, the inclusion of the interaction effect removes significance from the Selic rate. Plus, the models with control variables selected by Lasso sees more lost of significance for the Selic rate. This gives more indication that the first result that banks have more returns with lower interest rates is biased.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	ROA	ROA	ROA	ROA	ROE	ROE	ROE	ROE
HHI_{t-1}^{Geo}	0.0164***	0.0156***	0.0143***	0.0139***	0.200***	0.195***	0.182***	0.180***
	(0.00365)	(0.00370)	(0.00390)	(0.00398)	(0.0353)	(0.0365)	(0.0344)	(0.0359)
$Selic_{t-1}$	-0.000230***	-0.000308**	-0.000235**	-0.000277*	-0.00167***	-0.00216*	-0.000973*	-0.00117
	(6.17e-05)	(0.000132)	(9.11e-05)	(0.000149)	(0.000447)	(0.00118)	(0.000561)	(0.00125)
$Selic_{t-1} \times HHI_{t-1}^{Geo}$		0.000176		9.70e-05		0.00110		0.000456
		(0.000226)		(0.000243)		(0.00220)		(0.00223)
Constant	0.0725***	0.0735***	0.0383	0.0388	0.779***	0.786***	0.272	0.274
	(0.0252)	(0.0257)	(0.0304)	(0.0303)	(0.264)	(0.268)	(0.259)	(0.259)
Observations	492	492	492	492	492	492	492	492
R-squared	0.074	0.074	0.144	0.144	0.052	0.052	0.128	0.128
Number of Banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 14 – Difference of the effects of monetary policy in the performance of focused and diversified banks - Base model - Geographic concentration - Fixed Effects Estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	ROA	ROA	ROA	ROA	ROE	ROE	ROE	ROE
HHI_{t-1}^{EA}	0.00738***	0.0150***	0.00692***	0.0146***	0.0744***	0.0998**	0.0713***	0.0965**
	(0.00138)	(0.00505)	(0.00139)	(0.00499)	(0.0144)	(0.0411)	(0.0147)	(0.0410)
$Selic_{t-1}$	-0.000206***	-4.13e-06	-0.000196**	3.96e-06	-0.00138**	-0.000712	-0.000502	0.000160
	(6.05e-05)	(0.000139)	(9.00e-05)	(0.000165)	(0.000510)	(0.00137)	(0.000623)	(0.00139)
$Selic_{t-1} \times HHI_{t-1}^{AE}$		-0.000703		-0.000701		-0.00233		-0.00232
		(0.000463)		(0.000456)		(0.00400)		(0.00394)
Constant	0.0726**	0.0811**	0.0378	0.0467	0.767**	0.795**	0.249	0.279
	(0.0281)	(0.0307)	(0.0320)	(0.0340)	(0.303)	(0.310)	(0.296)	(0.301)
Observations	492	492	492	492	492	492	492	492
R-squared	0.076	0.080	0.148	0.151	0.047	0.048	0.125	0.126
Number of Banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 15 – Difference of the effects of monetary policy in the performance of focused and diversified banks - Base model - Economic Activity concentration - Fixed Effects Estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	ROA	ROA	ROA	ROA	ROE	ROE	ROE	ROE
HHI_{t-1}^{Geo}	0.0164***	0.0156***	0.0143***	0.0147***	0.179***	0.193***	0.178***	0.179***
	(0.00365)	(0.00370)	(0.00336)	(0.00348)	(0.0348)	(0.0351)	(0.0336)	(0.0345)
$Selic_{t-1}$	-0.000230***	-0.000308**	-0.000110	-7.88e-05	-0.00161***	-0.00165*	-0.000807	-0.000705
	(6.17e-05)	(0.000132)	(8.41e-05)	(0.000162)	(0.000480)	(0.000969)	(0.000594)	(0.00118)
$Selic_{t-1} \times HHI_{t-1}^{Geo}$		0.000176		-6.99e-05		3.35e-05		-0.000234
		(0.000226)		(0.000271)		(0.00171)		(0.00197)
Constant	0.0725***	0.0735***	0.0232	0.0228	0.779***	0.776***	0.318	0.317
	(0.0252)	(0.0257)	(0.0263)	(0.0258)	(0.270)	(0.277)	(0.271)	(0.271)
Observations	492	492	492	492	492	492	492	492
R-squared	0.074	0.074	0.171	0.171	0.072	0.073	0.146	0.146
Number of Banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Selected Variables	0	0	1	1	1	2	2	2

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 16 – Difference of the effects of monetary policy in the performance of focused and diversified banks - With control variables selected by Cluster Lasso - Geographic concentration - Fixed Effects Estimation

VARIABLES	(1) ROA	(2) ROA	(3) ROA	(4) ROA	(5) ROE	(6) ROE	(7) ROE	(8) ROE
HHI^{EA}_{t-1}	0.00738***	0.0150***	0.00692***	0.0112**	0.0724***	0.109***	0.0704***	0.102**
	(0.00138)	(0.00505)	(0.00139)	(0.00416)	(0.0144)	(0.0393)	(0.0145)	(0.0396)
$Selic_{t-1}$	-0.000206***	-4.13e-06	-0.000196**	5.29e-05	-0.00138**	-0.000397	-0.000339	0.000501
	(6.05e-05)	(0.000139)	(9.00e-05)	(0.000145)	(0.000530)	(0.00127)	(0.000639)	(0.00134)
$Selic_{t-1} \times HHI^{AE}_{t-1}$		-0.000703		-0.000482		-0.00341		-0.00294
		(0.000463)		(0.000384)		(0.00382)		(0.00379)
Constant	0.0726**	0.0811**	0.0378	0.0294	0.818**	0.860**	0.298	0.336
	(0.0281)	(0.0307)	(0.0320)	(0.0319)	(0.346)	(0.359)	(0.316)	(0.322)
Observations	492	492	492	492	492	492	492	492
R-squared	0.076	0.080	0.148	0.173	0.074	0.075	0.144	0.145
Number of Banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Selected Variables	0	0	0	1	2	2	2	2

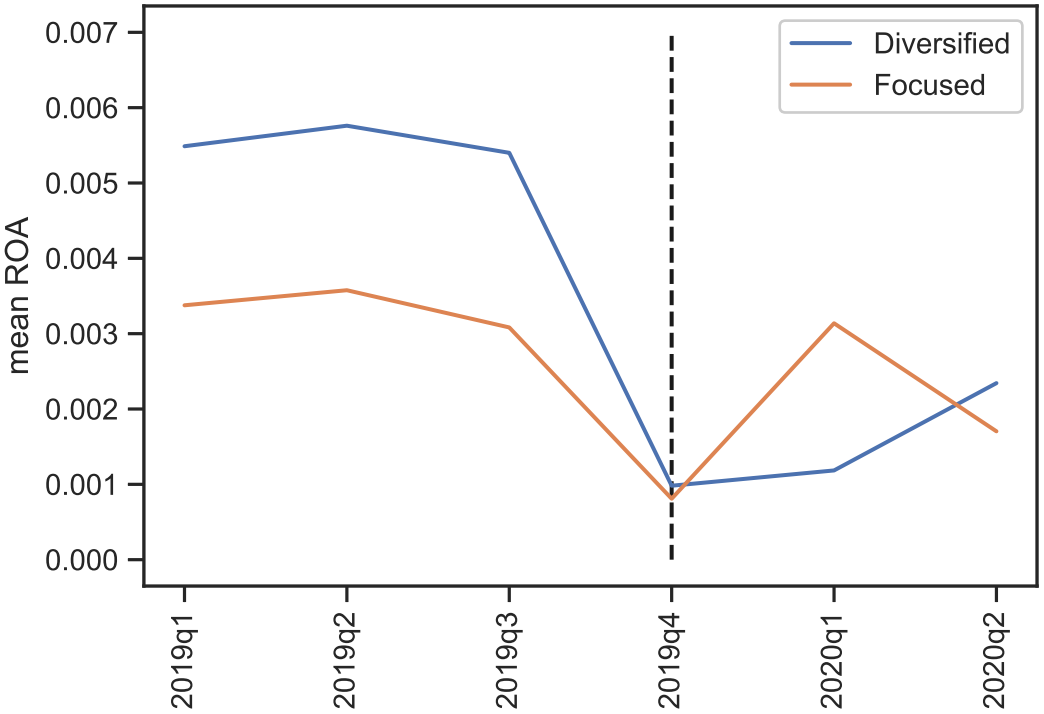
Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 17 – Difference of the effects of monetary policy in the performance of focused and diversified banks - With control variables selected by Cluster Lasso - Economic Activity concentration - Fixed Effects Estimation

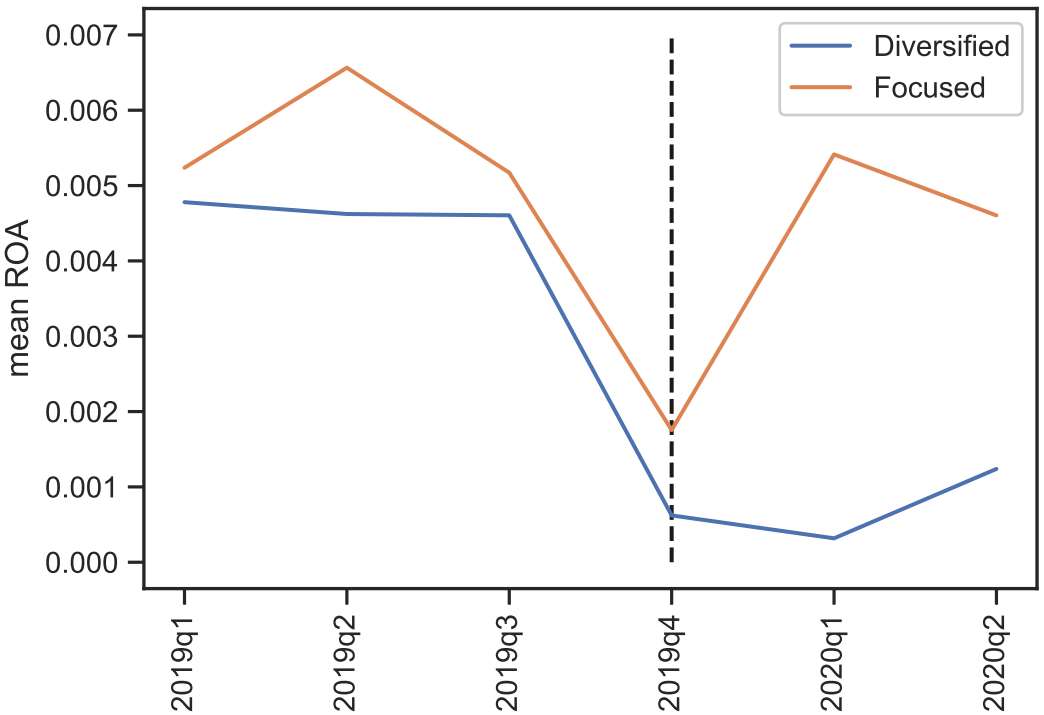
3.4 Resilience to economic downturns

We begin this test by checking if the parallel trends hypothesis is met for our sample. The test is presented in fig. 4 and fig. 5. Is worth mentioning that ROA had an atypical low value in the last quarter of 2019. This might give the wrong impression that there was an increase in revenue during the pandemic. Although it was a recovery from the drip before, profits were still smaller than the period before the pandemic. For geographic region concentration measure, we clearly see that the parallel trends hypothesis is met in the period before the intervention. But in the post period, it do not seen that the impact was much different between groups. Focused banks had a steep increase while diversified banks had a more modest increase in the first quarter, nonetheless in the second quarter diversified banks kept recovering while focused banks returns took a turn for the worse. When the groups are formed using economic activity concentration, the parallel trends hypothesis is not as clear, although it could be argued that there is a slight confirmation. However in the post treatment period, there is a clear difference in response between the two groups. Focused banks had a great increase in return, almost matching before pandemic levels, whereas diversified banks had a modest increase but still far from before. We consider that his big difference in the post period is worth testing despite the fact that parallel trends are not that evident.



Source: Brazil's Central Bank

Figure 4 – This figure shows the test for parallel trends hypothesis for geographic region concentration measure



Source: Brazil's Central Bank

Figure 5 – This figure shows the test for parallel trends hypothesis for economic activity concentration measure

Results from the DiD estimation are presented in table 18. The post variable is statistically significant and negative confirming that there was a decrease in returns in the first two quarters of 2020. For loan concentration measured by geographic regions, there is no evidence that focused banks and diversified banks were affected differently, as we inferred from our visual inspection in fig. 4. What could explain this is if the recession affected all regions proportionally to their share of banking activity. Looking at concentration by economic activity, results show that focused banks actually had a better performance than diversified ones. This could mean that focused banks, with their expertise, would know how to renegotiate and adjust prices in a way as to make their costumers still be able to pay their loans and get new ones.

We acknowledge that the full effects of the pandemic are yet to be seen, therefore it may still be early to draw conclusions using only two quarters of data.

VARIABLES	(1) ROA	(2) ROA
DiD Variables		
Post	-0.00155**	-0.00157*
	(0.000726)	(0.000877)
Post $\times Focused^{Geo}$	0.00157	
	(0.00156)	
Post $\times Focused^{EA}$		0.00249**
		(0.00115)
Control Variables		
Size_{t-1}	-0.00104	-0.00391
	(0.00327)	(0.00378)
Eq. Ratio_{t-1}	0.0914*	0.0751
	(0.0479)	(0.0447)
Llp_{t-1}	0.0695	0.0290
	(0.152)	(0.133)
Loan Ratio_{t-1}	-0.0355*	-0.0362*
	(0.0180)	(0.0186)
Constant	0.0281	0.0821
	(0.0619)	(0.0720)
Observations	145	145
R-squared	0.133	0.144
Number of Banks	29	29

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 18 – Difference in difference estimation of how the economic recession during the Covid-19 pandemic affected differently focused and diversified banks

4 Conclusion

In this paper, we investigated how the choice made by banks between diversifying or focusing their loan portfolio by geographic regions and by economic activity. This question is of great importance for banks, who have to adjust their business plans in order to profit as much as they can without facing excessive risks, and also for regulators, so they can improve legislation and allow banks to achieve the best possible allocation of their assets.

We run fixed effect models to estimate the relation between returns and risk against the HHI measuring concentration by geographic region and economic activity sector. We find that having a focused loan portfolio increases the bank's return and reduces their level of credit risk. This follows the line of banking theory more associated with corporate finance. Focusing in fewer sector or regions, allows banks to acquire more expertise, improving their screening of borrowers, monitoring risky loans and providing better services and products. We complement this models by adding control variables selected by Lasso among a rich set of variables used in banking empirical research. This is a novel approach for banking empirical approach. Although the inclusion of new variables only improved results for one test, their inclusion gives robustness to the tests since they help avoid omitted variable bias.

Lastly, we make two more tests to check how diversified and focused banks react to external factors. First, we test if monetary policy affects differently focused and diversified banks. Results shows some evidence that higher interest rates reduces bank returns although this result is not consistent and we discuss why it could have bias. Moreover, there was no evidence that monetary policy affects differently diversified and focused banks. In the second test, we check which strategy is more robust to an economic recession. Our findings show that banks that are focused in fewer sectors suffered less during the Covid-19 pandemic. Although some results were not much enlightening, we hope that this tests promotes more questions to be researched in the banking loan concentration literature, which is mainly focused in risk and return questions.

References

ACHARYA, Viral V; HASAN, Iftekhar; SAUNDERS, Anthony. Should banks be diversified? Evidence from individual bank loan portfolios. **The Journal of Business**, JSTOR, v. 79, n. 3, p. 1355–1412, 2006.

ATHEY, Susan. The impact of machine learning on economics. In: THE economics of artificial intelligence: An agenda. [S.l.]: University of Chicago Press, 2018. P. 507–547.

ATHEY, Susan; IMBENS, Guido. Recursive partitioning for heterogeneous causal effects. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 113, n. 27, p. 7353–7360, 2016.

BECK, Thorsten; DE JONGHE, Olivier; MULIER, Klaas. Bank sectoral concentration and (systemic) risk: Evidence from a worldwide sample of banks. **Available at SSRN 2959273**, 2018.

BEHR, Andreas et al. Diversification and the banks' risk-return-characteristics: evidence from loan portfolios of German banks. Bundesbank Series 2 Discussion Paper, 2007.

BELLONI, Alexandre; CHERNOZHUKOV, Victor. Least squares after model selection in high-dimensional sparse models. **Bernoulli**, Bernoulli Society for Mathematical Statistics and Probability, v. 19, n. 2, p. 521–547, 2013.

BELLONI, Alexandre; CHERNOZHUKOV, Victor; HANSEN, Christian. High-dimensional methods and inference on structural and treatment effects. **Journal of Economic Perspectives**, v. 28, n. 2, p. 29–50, 2014.

_____. Inference on treatment effects after selection among high-dimensional controls. **The Review of Economic Studies**, Oxford University Press, v. 81, n. 2, p. 608–650, 2014.

BELLONI, Alexandre; CHERNOZHUKOV, Victor; HANSEN, Christian; KOZBUR, Damian. Inference in high-dimensional panel models with an application to gun control. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 34, n. 4, p. 590–605, 2016.

BERGER, Philip G; OFEK, Eli. Diversification's effect on firm value. **Journal of financial economics**, North-Holland, v. 37, n. 1, p. 39–65, 1995.

BORIO, Claudio; GAMBACORTA, Leonardo; HOFMANN, Boris. The influence of monetary policy on bank profitability. **International Finance**, Wiley Online Library, v. 20, n. 1, p. 48–63, 2017.

CHU, Yongqiang; DENG, Saiying; XIA, Cong. Bank Geographic Diversification and Systemic Risk. **The Review of Financial Studies**, 2019.

DENIS, David J; DENIS, Diane K; SARIN, Atulya. Agency problems, equity ownership, and corporate diversification. **The Journal of Finance**, Wiley Online Library, v. 52, n. 1, p. 135–160, 1997.

DIAMOND, Douglas W. Financial intermediation and delegated monitoring. **The review of economic studies**, Wiley-Blackwell, v. 51, n. 3, p. 393–414, 1984.

ELSAS, Ralf; HACKETHAL, Andreas; HOLZHÄUSER, Markus. The anatomy of bank diversification. **Journal of Banking & Finance**, Elsevier, v. 34, n. 6, p. 1274–1287, 2010.

GARCIA, Márcio GP; MEDEIROS, Marcelo C; VASCONCELOS, Gabriel FR. Real-time inflation forecasting with high-dimensional models: The case of Brazil. **International Journal of Forecasting**, Elsevier, v. 33, n. 3, p. 679–693, 2017.

GOETZ, Martin R; LAEVEN, Luc; LEVINE, Ross. Does the geographic expansion of banks reduce risk? **Journal of Financial Economics**, Elsevier, v. 120, n. 2, p. 346–362, 2016.

IBRAGIMOV, Rustam; JAFFEE, Dwight; WALDEN, Johan. Diversification disasters. **Journal of financial economics**, Elsevier, v. 99, n. 2, p. 333–348, 2011.

JAHN, Nadya; MEMMEL, Christoph; PFINGSTEN, Andreas. Banks' concentration versus diversification in the loan portfolio: new evidence from Germany. Bundesbank Discussion Paper, 2013.

MEDEIROS, Marcelo C et al. Forecasting Inflation in a data-rich environment: the benefits of machine learning methods. **Journal of Business & Economic Statistics**, Taylor & Francis, p. 1–22, 2019.

SAMUELSON, Paul A. The effect of interest rate increases on the banking system. **The American economic review**, JSTOR, v. 35, n. 1, p. 16–27, 1945.

TABAK, Benjamin M; FAZIO, Dimas M; CAJUEIRO, Daniel O. The effects of loan portfolio concentration on Brazilian banks' return and risk. **Journal of Banking & Finance**, Elsevier, v. 35, n. 11, p. 3065–3076, 2011.

TIBSHIRANI, Robert. Regression shrinkage and selection via the lasso. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 58, n. 1, p. 267–288, 1996.

VARIAN, Hal R. Big data: New tricks for econometrics. **Journal of Economic Perspectives**, v. 28, n. 2, p. 3–28, 2014.

WAGER, Stefan; ATHEY, Susan. Estimation and inference of heterogeneous treatment effects using random forests. **Journal of the American Statistical Association**, Taylor & Francis, v. 113, n. 523, p. 1228–1242, 2018.

WINTON, Andrew. Don't put all your eggs in one basket? Diversification and specialization in lending. **Diversification and Specialization in Lending (September 27, 1999)**, 1999.

Appendix

APPENDIX A – Full regression tables

VARIABLES	(1) ROA	(2) ROA	(3) ROA	(4) ROA	(5) ROE	(6) ROE	(7) ROE	(8) ROE
HHI_{t-1}^{Geo}	0.0143***	0.0166***	0.0143***	0.0221***	0.185***	0.191***	0.182***	0.243***
	(0.00319)	(0.00568)	(0.00390)	(0.00533)	(0.0285)	(0.0552)	(0.0344)	(0.0602)
$HHI_{t-1}^{Geo} \times State - owned$		0.0104		0.0130		0.175		0.0785
		(0.0155)		(0.0232)		(0.236)		(0.220)
$HHI_{t-1}^{Geo} \times Foreign$		-0.00471		-0.0151*		-0.0202		-0.116
		(0.00864)		(0.00782)		(0.0815)		(0.0873)
$Size_{t-1}$	-0.00121	-0.000968	-0.00242	-0.00181	-0.0226	-0.0211	-0.0179	-0.0131
	(0.00138)	(0.00155)	(0.00165)	(0.00188)	(0.0147)	(0.0164)	(0.0137)	(0.0155)
$Eq.Ratio_{t-1}$	0.0284*	0.0289*	0.0199	0.0200	-0.0360	-0.0289	-0.0736	-0.0732
	(0.0140)	(0.0143)	(0.0139)	(0.0140)	(0.112)	(0.116)	(0.127)	(0.129)
$LlpRatio_{t-1}$	0.150**	0.152**	0.174**	0.182**	0.761	0.775	0.833*	0.891**
	(0.0665)	(0.0655)	(0.0698)	(0.0686)	(0.489)	(0.479)	(0.437)	(0.413)
$LoanRatio_{t-1}$	-0.0117**	-0.0117**	-0.00830	-0.00843*	-0.0931*	-0.0934*	-0.0425	-0.0433
	(0.00474)	(0.00462)	(0.00502)	(0.00465)	(0.0472)	(0.0462)	(0.0444)	(0.0416)
Constant	0.0173	0.0116	0.0349	0.0213	0.380	0.334	0.258	0.154
	(0.0261)	(0.0294)	(0.0297)	(0.0338)	(0.281)	(0.315)	(0.257)	(0.293)
Observations	492	492	492	492	492	492	492	492
R-squared	0.054	0.055	0.144	0.147	0.039	0.040	0.128	0.130
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 19 – Full Regression table of table 6

VARIABLES	(1) ROA	(2) ROA	(3) ROA	(4) ROA	(5) ROE	(6) ROE	(7) ROE	(8) ROE
HHI_{t-1}^{AE}	0.00750***	0.00585***	0.00692***	0.00564***	0.0753***	0.0621***	0.0713***	0.0603***
	(0.00163)	(0.00150)	(0.00139)	(0.00127)	(0.0147)	(0.0130)	(0.0147)	(0.0139)
$HHI_{t-1}^{AE} \times State-owned$		0.0578**		0.0510**		0.455		0.453
		(0.0232)		(0.0244)		(0.279)		(0.267)
$HHI_{t-1}^{AE} \times Foreign$		0.00555		0.00343		0.0454		0.0272
		(0.00571)		(0.00580)		(0.0374)		(0.0333)
$Size_{t-1}$	-0.00116	-0.00137	-0.00203	-0.00174	-0.0209	-0.0225	-0.0121	-0.00949
	(0.00144)	(0.00141)	(0.00170)	(0.00179)	(0.0156)	(0.0153)	(0.0155)	(0.0163)
$Eq.Ratio_{t-1}$	0.0354**	0.0284**	0.0282*	0.0238	0.0489	-0.00601	0.0281	-0.0117
	(0.0129)	(0.0138)	(0.0141)	(0.0145)	(0.111)	(0.125)	(0.138)	(0.146)
$LlpRatio_{t-1}$	0.158**	0.177**	0.179**	0.189**	0.810*	0.960**	0.842**	0.933**
	(0.0676)	(0.0674)	(0.0703)	(0.0706)	(0.443)	(0.429)	(0.402)	(0.399)
$LoanRatio_{t-1}$	-0.0192***	-0.0160***	-0.0154**	-0.0124**	-0.171***	-0.145**	-0.118**	-0.0906*
	(0.00585)	(0.00549)	(0.00590)	(0.00587)	(0.0562)	(0.0534)	(0.0530)	(0.0533)
Constant	0.0234	0.0233	0.0350	0.0267	0.437	0.437	0.242	0.169
	(0.0275)	(0.0269)	(0.0314)	(0.0330)	(0.298)	(0.294)	(0.293)	(0.311)
Observations	492	492	492	492	492	492	492	492
R-squared	0.060	0.069	0.148	0.154	0.038	0.045	0.125	0.132
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 20 – Full Regression table of table 7

VARIABLES	(1) ROA	(2) ROA	(3) ROA	(4) ROA	(5) ROE	(6) ROE	(7) ROE	(8) ROE
HHI^{Geo}_{t-1}	0.0151***	0.0162***	0.0143***	0.0209***	0.179***	0.192***	0.178***	0.243***
	(0.00273)	(0.00448)	(0.00336)	(0.00458)	(0.0282)	(0.0562)	(0.0336)	(0.0606)
$HHI^{Geo}_{t-1} \times State-owned$		0.0174		0.0162		0.179		0.0952
		(0.0157)		(0.0227)		(0.241)		(0.225)
$HHI^{Geo}_{t-1} \times Foreign$		-0.00298		-0.0129		-0.0325		-0.125
		(0.00853)		(0.00775)		(0.0803)		(0.0872)
$Size_{t-1}$	-0.00152	-0.00134			-0.0246	-0.0226	-0.0206	-0.0155
	(0.00134)	(0.00156)			(0.0149)	(0.0166)	(0.0144)	(0.0160)
$Eq.Ratio_{t-1}$	0.0211	0.0217			-0.0130	-0.00523	-0.0527	-0.0516
	(0.0131)	(0.0134)			(0.111)	(0.114)	(0.125)	(0.126)
$LlpRatio_{t-1}$	0.160**	0.162**			0.782	0.801*	0.882*	0.946**
	(0.0657)	(0.0647)			(0.476)	(0.467)	(0.441)	(0.417)
$LoanRatio_{t-1}$	-0.00946**	-0.00958**			-0.0943*	-0.0946**	-0.0468	-0.0478
	(0.00413)	(0.00399)			(0.0476)	(0.0464)	(0.0460)	(0.0429)
Lasso Selected Variables								
Npl_TotalAssets $\times 1 - HHI\_Domestic_{t-1}$		0.00212						
		(0.00892)						
NonDepStFund_TotalFunding $\times LoanRatio_{t-1}$	-0.0138***	-0.0136***						
	(0.00279)	(0.00336)						
NonIntIncToIntIncome $\times 1 - HHI\_Domestic_{t-1}$					0.0210***	0.0209***	0.0235***	0.0234***
					(0.00700)	(0.00698)	(0.00518)	(0.00487)
ROE_Volatility $\times G\_Equity_{t-1}$					-0.465***	-0.469***	-0.367***	-0.375***
					(0.0957)	(0.0941)	(0.0964)	(0.0917)
Constant	0.0241	0.0187	0.0217	0.00984	0.415	0.358	0.306	0.194
	(0.0253)	(0.0294)	(0.0259)	(0.0300)	(0.284)	(0.318)	(0.270)	(0.302)
Observations	492	492	492	492	492	492	492	492
R-squared	0.088	0.088	0.171	0.174	0.061	0.061	0.146	0.148
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Selected Variables	1	2	1	1	2	2	2	2

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 21 – Full Regression table of table 8

VARIABLES	(1) ROA	(2) ROA	(3) ROA	(4) ROA	(5) ROE	(6) ROE	(7) ROE	(8) ROE
HHI^{AE}_{t-1}	0.00750***	0.00640***	0.00692***	0.00608***	0.0737***	0.0615***	0.0704***	0.0600***
	(0.00163)	(0.00143)	(0.00139)	(0.00120)	(0.0147)	(0.0130)	(0.0145)	(0.0137)
$HHI^{AE}_{t-1} \times State-owned$		0.0583**		0.0506**		0.457		0.454*
		(0.0238)		(0.0242)		(0.278)		(0.268)
$HHI^{AE}_{t-1} \times Foreign$		0.00326		-3.36e-05		0.0375		0.0224
		(0.00487)		(0.00470)		(0.0364)		(0.0335)
$Size_{t-1}$	-0.00116	-0.00134		-0.00119	-0.0230	-0.0247	-0.0149	-0.0123
	(0.00144)	(0.00126)		(0.00159)	(0.0161)	(0.0159)	(0.0167)	(0.0175)
$Eq.Ratio_{t-1}$	0.0354**	0.0303**		0.0283**	0.0724	0.0158	0.0504	0.00915
	(0.0129)	(0.0131)		(0.0135)	(0.107)	(0.122)	(0.134)	(0.143)
$LlpRatio_{t-1}$	0.158**	0.201***		0.208***	0.831*	0.977**	0.893**	0.981**
	(0.0676)	(0.0649)		(0.0704)	(0.430)	(0.418)	(0.409)	(0.403)
$LoanRatio_{t-1}$	-0.0192***	-0.0179***		-0.0138**	-0.171***	-0.145**	-0.121**	-0.0947*
	(0.00585)	(0.00512)		(0.00550)	(0.0568)	(0.0537)	(0.0541)	(0.0548)
Lasso Selected Variables								
NPL_Total Assets $\times LoanDepositsRatio_{t-1}$				-0.000152				
				(0.000720)				
$LlpRatio_{t-1} \times LoanDepositsRatio_{t-1}$				-0.00128				
				(0.00134)				
LoanDepositsRatio $\times ROA\_Volatiliry_{t-1}$				-0.00718				
				(0.00655)				
FIIncomeLiquidAssetsRatio $\times LoanDepositsRatio_{t-1}$		-0.000148***						
		(2.75e-05)						
NonIntIncToIntIncome $\times 1 - HHI\_Domestic_{t-1}$					0.0177**	0.0189**	0.0200***	0.0210***
					(0.00790)	(0.00802)	(0.00550)	(0.00542)
ROE_Volatility $\times G\_Equity$					-0.523***	-0.505***	-0.434***	-0.418***
					(0.106)	(0.102)	(0.100)	(0.0974)
Constant	0.0234	0.0233	0.0350	0.0175	0.473	0.473	0.293	0.219
	(0.0275)	(0.0241)	(0.0314)	(0.0294)	(0.306)	(0.305)	(0.313)	(0.332)
Observations	492	492	492	492	492	492	492	492
R-squared	0.060	0.085	0.148	0.165	0.060	0.067	0.144	0.151
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Selected Variables	0	1	0	3	2	2	2	2

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 22 – Full Regression table of table 9

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	NPL	NPL	NPL	NPL	NPL	NPL	NPL	NPL
HHI^{Geo}_{t-1}	-0.141***	-0.0486	-0.145***	-0.0493	-0.140***	-0.0484	-0.144***	-0.0491
	(0.0459)	(0.0622)	(0.0489)	(0.0543)	(0.0465)	(0.0629)	(0.0493)	(0.0548)
$HHI^{Geo}_{t-1} \times State - owned$		0.0245		-0.0319		0.0263		-0.0325
		(0.153)		(0.142)		(0.151)		(0.140)
$HHI^{Geo}_{t-1} \times Foreign$		-0.174*		-0.178**		-0.174*		-0.177**
		(0.0909)		(0.0840)		(0.0907)		(0.0838)
$Size_{t-1}$	-0.0266**	-0.0195*	-0.0138	-0.00589	-0.0265**	-0.0194*	-0.0135	-0.00553
	(0.0104)	(0.00958)	(0.0112)	(0.00944)	(0.0103)	(0.00958)	(0.0113)	(0.00955)
$Eq.Ratio_{t-1}$	-0.0321	-0.0287	0.0174	0.0165	-0.0389	-0.0356	0.0118	0.0105
	(0.107)	(0.0913)	(0.108)	(0.0841)	(0.107)	(0.0910)	(0.108)	(0.0840)
$LoanRatio_{t-1}$	-0.0280	-0.0293	-0.0188	-0.0179	-0.0264	-0.0276	-0.0173	-0.0163
	(0.0430)	(0.0401)	(0.0401)	(0.0362)	(0.0429)	(0.0400)	(0.0400)	(0.0362)
ROA_{t-1}	-0.225**	-0.233**	-0.205*	-0.233**				
	(0.0889)	(0.0873)	(0.103)	(0.0944)				
ROE_{t-1}					-0.0195	-0.0203	-0.0172	-0.0205
					(0.0139)	(0.0141)	(0.0155)	(0.0152)
Constant	0.587**	0.443**	0.353	0.200	0.584**	0.441**	0.347	0.193
	(0.216)	(0.201)	(0.215)	(0.180)	(0.216)	(0.201)	(0.216)	(0.182)
Observations	492	492	492	492	492	492	492	492
R-squared	0.175	0.204	0.214	0.242	0.173	0.202	0.213	0.240
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	No	No	No	No	No	No

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 23 – Full Regression table of table 10

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	NPL	NPL	NPL	NPL	NPL	NPL	NPL	NPL
HHI^{AE}_{t-1}	-0.0446**	-0.00869	-0.0449**	-0.0107	-0.0446**	-0.00939	-0.0448**	-0.0113
	(0.0200)	(0.0131)	(0.0201)	(0.0132)	(0.0200)	(0.0130)	(0.0203)	(0.0132)
$HHI^{AE}_{t-1} \times State-owned$		-0.184*		-0.136		-0.182*		-0.132
		(0.102)		(0.103)		(0.102)		(0.104)
$HHI^{AE}_{t-1} \times Foreign$		-0.243***		-0.238**		-0.241***		-0.236**
		(0.0874)		(0.0869)		(0.0878)		(0.0871)
$Size_{t-1}$	-0.0285**	-0.0256***	-0.0190*	-0.0186***	-0.0284**	-0.0255***	-0.0186*	-0.0183***
	(0.0116)	(0.00785)	(0.00973)	(0.00646)	(0.0115)	(0.00787)	(0.00979)	(0.00664)
$Eq.Ratio_{t-1}$	-0.0943	-0.0789	-0.0610	-0.0649	-0.0985	-0.0873	-0.0644	-0.0727
	(0.0993)	(0.0770)	(0.0914)	(0.0776)	(0.0987)	(0.0789)	(0.0913)	(0.0799)
$LoanRatio_{t-1}$	0.0197	-0.0249	0.0303	-0.0114	0.0210	-0.0225	0.0316	-0.00900
	(0.0546)	(0.0506)	(0.0525)	(0.0475)	(0.0542)	(0.0504)	(0.0522)	(0.0475)
ROA_{t-1}	-0.217**	-0.235***	-0.199*	-0.238***				
	(0.0811)	(0.0687)	(0.105)	(0.0794)				
ROE_{t-1}					-0.0229*	-0.0177	-0.0204	-0.0176
					(0.0121)	(0.0105)	(0.0149)	(0.0121)
Constant	0.554**	0.549***	0.377*	0.414***	0.551**	0.547***	0.369*	0.407***
	(0.234)	(0.168)	(0.188)	(0.129)	(0.232)	(0.169)	(0.189)	(0.133)
Observations	492	492	492	492	492	492	492	492
R-squared	0.139	0.293	0.176	0.318	0.139	0.290	0.175	0.316
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	No	No	No	No	No	No

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 24 – Full Regression table of table 11

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	NPL	NPL	NPL	NPL	NPL	NPL	NPL	NPL
HHI^{Geo}_{t-1}	-0.116**	-0.0504	-0.124***	-0.0493	-0.0993**	0.00443	-0.118**	-0.0491
	(0.0428)	(0.0589)	(0.0415)	(0.0543)	(0.0449)	(0.0450)	(0.0442)	(0.0548)
$HHI^{Geo}_{t-1} \times State - owned$		0.102		-0.0319		0.0194		-0.0325
		(0.135)		(0.142)		(0.120)		(0.140)
$HHI^{Geo}_{t-1} \times Foreign$		-0.171*		-0.178**		-0.195***		-0.177**
		(0.0850)		(0.0840)		(0.0642)		(0.0838)
$Size_{t-1}$	-0.0252***	-0.0151*	-0.0186	-0.00589	-0.0233***	-0.0149**	-0.0163	-0.00553
	(0.00831)	(0.00769)	(0.0113)	(0.00944)	(0.00741)	(0.00732)	(0.0110)	(0.00955)
$Eq.Ratio_{t-1}$	-0.0601	0.00169	-0.0154	0.0165	-0.0457	-0.0389	-0.0244	0.0105
	(0.106)	(0.0762)	(0.107)	(0.0841)	(0.0866)	(0.0695)	(0.106)	(0.0840)
$LoanRatio_{t-1}$	-0.00701	-0.0292	0.000762	-0.0179	-0.00593	-0.00784	0.00227	-0.0163
	(0.0260)	(0.0385)	(0.0252)	(0.0362)	(0.0259)	(0.0244)	(0.0243)	(0.0362)
ROA_{t-1}	-0.330***	-0.193*	-0.317**	-0.233**				
	(0.103)	(0.0980)	(0.122)	(0.0944)				
ROE_{t-1}					-0.0312	-0.0301	-0.0374	-0.0205
					(0.0230)	(0.0235)	(0.0241)	(0.0152)
Lasso Selected Variables								
BranchRatio $\times ROE\_Volat_{t-1}$		64,952***						
		(18,335)						
1 - HHI_Interest $\times 1 - HHI\_Domestic_{t-1}$		-0.0215**						
		(0.00996)						
NonIntIncTotalIncomeRatio $\times G\_Equity_{t-1}$	0.00180***							
	(0.000209)							
MarketShare $\times ROE\_Vol_{t-1}$	1.234***							
	(0.382)							
RecLoansProbLoans $\times 1 - HHI\_Domestic_{t-1}$	0.0535***		0.0500***		0.0503***	0.0447***	0.0522***	
	(0.00437)		(0.00478)		(0.00869)	(0.00396)	(0.00551)	
StDepositsTotalDeposits $\times G\_Assets$			0.164***					
			(0.0418)					
ROE_Volatility $\times G\_Equity$			0.0232					
			(0.0206)					
FIExpFundingRatio $\times 1 - HHI\_Domestic_{t-1}$					-0.0755			
					(0.118)			
$Roe\_Vol^2_{t-1}$					1.636***	1.855***		
					(0.402)	(0.326)		
Constant	0.545***	0.352**	0.422*	0.200	0.501***	0.334**	0.381*	0.193
	(0.166)	(0.161)	(0.211)	(0.180)	(0.147)	(0.145)	(0.207)	(0.182)
Observations	492	492	492	492	492	492	492	492
R-squared	0.261	0.251	0.292	0.242	0.295	0.330	0.285	0.240
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	No	No	No	No	No	No
Selected Variables	3	2	3	0	3	2	1	0

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 25 – Full Regression table of table 12

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	NPL	NPL	NPL	NPL	NPL	NPL	NPL	NPL
HHI^{AE}_{t-1}	-0.0428**	-0.0165**	-0.0469**	-0.0164*	-0.0425**	-0.0170**	-0.0460**	-0.0169*
	(0.0183)	(0.00703)	(0.0179)	(0.00861)	(0.0184)	(0.00706)	(0.0181)	(0.00848)
$HHI^{AE}_{t-1} \times State-owned$		-0.0557		-0.102		-0.0496		-0.0910
		(0.0751)		(0.0962)		(0.0737)		(0.0966)
$HHI^{AE}_{t-1} \times Foreign$		-0.198**		-0.211**		-0.195**		-0.208**
		(0.0758)		(0.0768)		(0.0764)		(0.0768)
$Size_{t-1}$	-0.0239***	-0.0204***	-0.0217**	-0.0204***	-0.0239***	-0.0203***	-0.0204**	-0.0196***
	(0.00791)	(0.00461)	(0.00990)	(0.00684)	(0.00787)	(0.00459)	(0.00941)	(0.00689)
$Eq.Ratio_{t-1}$	-0.0985	-0.0773	-0.0906	-0.0928	-0.0975	-0.0839	-0.0950	-0.0992
	(0.0912)	(0.0629)	(0.0989)	(0.0811)	(0.0875)	(0.0628)	(0.0958)	(0.0827)
$LoanRatio_{t-1}$	0.0383	0.0198	0.0535	0.0124	0.0384	0.0224	0.0540*	0.0159
	(0.0358)	(0.0243)	(0.0317)	(0.0272)	(0.0356)	(0.0242)	(0.0309)	(0.0261)
ROA_{t-1}	-0.0393	-0.262**	-0.298**	-0.330***				
	(0.202)	(0.112)	(0.146)	(0.111)				
ROE_{t-1}					-0.0102	-0.0269	-0.0407	-0.0353*
					(0.0327)	(0.0200)	(0.0259)	(0.0201)
Lasso Selected Variables								
$RecLoansProbLoans \times 1-HHI\_Domestic_{t-1}$	0.0568***	0.0455***	0.0596***	0.0504***	0.0574***	0.0462***	0.0619***	0.0513***
	(0.00272)	(0.00596)	(0.00325)	(0.00404)	(0.00343)	(0.00610)	(0.00456)	(0.00404)
$TradeIncomeTotalAssets \times G\_Equity$		-0.491***				-0.490***		
		(0.115)				(0.116)		
$TreasuryCreditRatio \times ROE\_Vol_{t-1}$		0.243**				0.243***		
		(0.0888)				(0.0883)		
$NonIntIncTotalIncomeRatio \times G\_Equity$	0.00117***				0.00116***			
	(0.000389)				(0.000409)			
$CostRevenueRatio \times ROE\_Vol_{t-1}$	-0.131***				-0.127***			
	(0.0346)				(0.0316)			
$1 - HHI\_Domestic \times G\_Assets$			0.0909***					
			(0.0209)					
Constant	0.463***	0.426***	0.418**	0.435***	0.462***	0.422***	0.396**	0.420***
	(0.157)	(0.0926)	(0.189)	(0.131)	(0.157)	(0.0916)	(0.179)	(0.132)
Observations	492	492	492	492	492	492	492	492
R-squared	0.282	0.411	0.285	0.387	0.283	0.410	0.281	0.386
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	No	No	No	No	No	No
Selected Variables	3	3	2	1	3	3	1	1

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 26 – Full Regression table of table 13

VARIABLES	(1) ROA	(2) ROA	(3) ROA	(4) ROA	(5) ROE	(6) ROE	(7) ROE	(8) ROE
HHI^{Geo}_{t-1}	0.0164***	0.0156***	0.0143***	0.0139***	0.200***	0.195***	0.182***	0.180***
	(0.00365)	(0.00370)	(0.00390)	(0.00398)	(0.0353)	(0.0365)	(0.0344)	(0.0359)
$Selic_{t-1}$	-0.000230***	-0.000308**	-0.000235**	-0.000277*	-0.00167***	-0.00216*	-0.000973*	-0.00117
	(6.17e-05)	(0.000132)	(9.11e-05)	(0.000149)	(0.000447)	(0.00118)	(0.000561)	(0.00125)
$Selic_{t-1} \times HHI^{Geo}_{t-1}$		0.000176		9.70e-05		0.00110		0.000456
		(0.000226)		(0.000243)		(0.00220)		(0.00223)
$Size_{t-1}$	-0.00422***	-0.00424***	-0.00242	-0.00243	-0.0444***	-0.0445***	-0.0179	-0.0179
	(0.00139)	(0.00142)	(0.00165)	(0.00166)	(0.0139)	(0.0140)	(0.0137)	(0.0137)
$Eq.Ratio_{t-1}$	0.0126	0.0124	0.0199	0.0198	-0.151	-0.152	-0.0736	-0.0740
	(0.0126)	(0.0127)	(0.0139)	(0.0140)	(0.125)	(0.127)	(0.127)	(0.128)
$LlpRatio_{t-1}$	0.182***	0.183***	0.174**	0.175**	0.995**	0.998**	0.833*	0.834*
	(0.0628)	(0.0629)	(0.0698)	(0.0698)	(0.392)	(0.393)	(0.437)	(0.439)
$LoanRatio_{t-1}$	-0.0120**	-0.0126**	-0.00830	-0.00868*	-0.0950*	-0.0993*	-0.0425	-0.0443
	(0.00501)	(0.00481)	(0.00502)	(0.00475)	(0.0510)	(0.0512)	(0.0444)	(0.0435)
Constant	0.0725***	0.0735***	0.0383	0.0388	0.779***	0.786***	0.272	0.274
	(0.0252)	(0.0257)	(0.0304)	(0.0303)	(0.264)	(0.268)	(0.259)	(0.259)
Observations	492	492	492	492	492	492	492	492
R-squared	0.074	0.074	0.144	0.144	0.052	0.052	0.128	0.128
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 27 – Full Regression table of table 14

VARIABLES	(1) ROA	(2) ROA	(3) ROA	(4) ROA	(5) ROE	(6) ROE	(7) ROE	(8) ROE
HHI^{AE}_{t-1}	0.00738***	0.0150***	0.00692***	0.0146***	0.0744***	0.0998**	0.0713***	0.0965**
	(0.00138)	(0.00505)	(0.00139)	(0.00499)	(0.0144)	(0.0411)	(0.0147)	(0.0410)
$Selic_{t-1}$	-0.000206***	-4.13e-06	-0.000196**	3.96e-06	-0.00138**	-0.000712	-0.000502	0.000160
	(6.05e-05)	(0.000139)	(9.00e-05)	(0.000165)	(0.000510)	(0.00137)	(0.000623)	(0.00139)
$Selic_{t-1} \times HHI^{AE}_{t-1}$		-0.000703		-0.000701		-0.00233		-0.00232
		(0.000463)		(0.000456)		(0.00400)		(0.00394)
$Size_{t-1}$	-0.00379**	-0.00438**	-0.00203	-0.00264	-0.0385**	-0.0405**	-0.0121	-0.0141
	(0.00150)	(0.00168)	(0.00170)	(0.00183)	(0.0158)	(0.0164)	(0.0155)	(0.0159)
$Eq.Ratio_{t-1}$	0.0219	0.0177	0.0282*	0.0240	-0.0419	-0.0556	0.0281	0.0143
	(0.0132)	(0.0138)	(0.0141)	(0.0148)	(0.139)	(0.141)	(0.138)	(0.142)
$LlpRatio_{t-1}$	0.185***	0.187***	0.179**	0.181**	0.990***	0.995***	0.842**	0.848**
	(0.0638)	(0.0635)	(0.0703)	(0.0700)	(0.347)	(0.348)	(0.402)	(0.403)
$LoanRatio_{t-1}$	-0.0195***	-0.0193***	-0.0154**	-0.0153**	-0.172***	-0.172***	-0.118**	-0.117**
	(0.00563)	(0.00572)	(0.00590)	(0.00606)	(0.0569)	(0.0572)	(0.0530)	(0.0535)
Constant	0.0726**	0.0811**	0.0378	0.0467	0.767**	0.795**	0.249	0.279
	(0.0281)	(0.0307)	(0.0320)	(0.0340)	(0.303)	(0.310)	(0.296)	(0.301)
Observations	492	492	492	492	492	492	492	492
R-squared	0.076	0.080	0.148	0.151	0.047	0.048	0.125	0.126
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 28 – Full Regression table of table 15

VARIABLES	(1) ROA	(2) ROA	(3) ROA	(4) ROA	(5) ROE	(6) ROE	(7) ROE	(8) ROE
HHI^{Geo}_{t-1}	0.0164***	0.0156***	0.0143***	0.0147***	0.179***	0.193***	0.178***	0.179***
	(0.00365)	(0.00370)	(0.00336)	(0.00348)	(0.0348)	(0.0351)	(0.0336)	(0.0345)
$Selic_{t-1}$	-0.000230***	-0.000308**	-0.000110	-7.88e-05	-0.00161***	-0.00165*	-0.000807	-0.000705
	(6.17e-05)	(0.000132)	(8.41e-05)	(0.000162)	(0.000480)	(0.000969)	(0.000594)	(0.00118)
$Selic_{t-1} \times HHI^{Geo}_{t-1}$		0.000176		-6.99e-05		3.35e-05		-0.000234
		(0.000226)		(0.000271)		(0.00171)		(0.00197)
$Size_{t-1}$	-0.00422***	-0.00424***	-0.00157	-0.00156	-0.0445***	-0.0444***	-0.0206	-0.0205
	(0.00139)	(0.00142)	(0.00141)	(0.00139)	(0.0140)	(0.0147)	(0.0144)	(0.0144)
$Eq.Ratio_{t-1}$	0.0126	0.0124	0.0179	0.0180	-0.0912	-0.111	-0.0527	-0.0525
	(0.0126)	(0.0127)	(0.0158)	(0.0158)	(0.143)	(0.117)	(0.125)	(0.125)
$LlpRatio_{t-1}$	0.182***	0.183***	0.177**	0.176**	0.725**	0.989**	0.882*	0.882*
	(0.0628)	(0.0629)	(0.0689)	(0.0690)	(0.323)	(0.384)	(0.441)	(0.443)
$LoanRatio_{t-1}$	-0.0120**	-0.0126**	-0.00595	-0.00566	-0.0870	-0.0915*	-0.0468	-0.0459
	(0.00501)	(0.00481)	(0.00431)	(0.00381)	(0.0548)	(0.0484)	(0.0460)	(0.0439)
Lasso Selected Variables								
NonDepStFund_TotalFunding $\times LoanRatio_{t-1}$			-0.0131***	-0.0132***				
			(0.00232)	(0.00247)				
NplLlpRatio $\times TreasuryAssetsRatio_{t-1}$					-1.959***			
					(0.675)			
NonIntIncIntIncome $\times ROE\_Vol_{t-1}$						0.0654**		
						(0.0262)		
ROE_Volatility $\times G\_Equity$						-0.415***	-0.367***	-0.368***
						(0.125)	(0.0964)	(0.0963)
NonIntIncToIntIncome $\times 1 - HHI\_Domestic_{t-1}$							0.0235***	0.0235***
							(0.00518)	(0.00515)
Constant	0.0725***	0.0735***	0.0232	0.0228	0.779***	0.776***	0.318	0.317
	(0.0252)	(0.0257)	(0.0263)	(0.0258)	(0.270)	(0.277)	(0.271)	(0.271)
Observations	492	492	492	492	492	492	492	492
R-squared	0.074	0.074	0.171	0.171	0.072	0.073	0.146	0.146
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Selected Variables	0	0	1	1	1	2	2	2

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 29 – Full Regression table of table 16

VARIABLES	(1) ROA	(2) ROA	(3) ROA	(4) ROA	(5) ROE	(6) ROE	(7) ROE	(8) ROE
HHI^{AE}_{t-1}	0.00738***	0.0150***	0.00692***	0.0112**	0.0724***	0.109***	0.0704***	0.102**
	(0.00138)	(0.00505)	(0.00139)	(0.00416)	(0.0144)	(0.0393)	(0.0145)	(0.0396)
$Selic_{t-1}$	-0.000206***	-4.13e-06	-0.000196**	5.29e-05	-0.00138**	-0.000397	-0.000339	0.000501
	(6.05e-05)	(0.000139)	(9.00e-05)	(0.000145)	(0.000530)	(0.00127)	(0.000639)	(0.00134)
$Selic_{t-1} \times HHI^{AE}_{t-1}$		-0.000703		-0.000482		-0.00341		-0.00294
		(0.000463)		(0.000384)		(0.00382)		(0.00379)
$Size_{t-1}$	-0.00379**	-0.00438**	-0.00203	-0.00164	-0.0416**	-0.0445**	-0.0149	-0.0175
	(0.00150)	(0.00168)	(0.00170)	(0.00171)	(0.0184)	(0.0193)	(0.0167)	(0.0171)
$Eq.Ratio_{t-1}$	0.0219	0.0177	0.0282*	0.0233	-0.0338	-0.0543	0.0504	0.0329
	(0.0132)	(0.0138)	(0.0141)	(0.0169)	(0.139)	(0.143)	(0.134)	(0.138)
$LlpRatio_{t-1}$	0.185***	0.187***	0.179**	0.180**	1.077***	1.087***	0.893**	0.901**
	(0.0638)	(0.0635)	(0.0703)	(0.0688)	(0.364)	(0.364)	(0.409)	(0.408)
$LoanRatio_{t-1}$	-0.0195***	-0.0193***	-0.0154**	-0.0124**	-0.169***	-0.168***	-0.121**	-0.121**
	(0.00563)	(0.00572)	(0.00590)	(0.00564)	(0.0572)	(0.0576)	(0.0541)	(0.0547)
Lasso Selected Variables								
NonDepStFund_TotalFunding $\times LoanRatio_{t-1}$				-0.0118***				
				(0.00220)				
NonIntIncTotalIncomeRatio $\times G\_Equity$					-0.00783***	-0.00805***		
					(0.00146)	(0.00149)		
ROE_Volatility $\times G\_Equity$					-0.151	-0.141	-0.434***	-0.435***
					(0.145)	(0.143)	(0.100)	(0.100)
NonIntIncToIntIncome $\times 1 - HHI\_Domestic_{t-1}$							0.0200***	0.0204***
							(0.00550)	(0.00539)
Constant	0.0726**	0.0811**	0.0378	0.0294	0.818**	0.860**	0.298	0.336
	(0.0281)	(0.0307)	(0.0320)	(0.0319)	(0.346)	(0.359)	(0.316)	(0.322)
Observations	492	492	492	492	492	492	492	492
R-squared	0.076	0.080	0.148	0.173	0.074	0.075	0.144	0.145
Number of banks	32	32	32	32	32	32	32	32
Bank Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Selected Variables	0	0	0	1	2	2	2	2

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 30 – Full Regression table of table 17

APPENDIX B – Control variables

Variable	Description	Frequency
Size	Logarithm of total assets	86.0%
Eq.Ratio	Total equity divided by total assets	77.7%
LlpRatio	Loan loss provision divided by total assets	32.5%
LoanRatio	Total loans divided by total assets	27.9%
NonIntIncTotalIncome	Non interest income divided by net profits	25.6%
G_Assets	Growth of total assets	23.4%
EmpExpRatio	Employee expenditure divided by total Assets	21.0%
NplRatio	Non performing loans divided by total assets	18.6%
Npl	Non performing loans divided by total loans	16.3%
Zscore	Zscore of bank using 4 periods to aggregate	16.3%
OverheadRatio	Overhead costs divided by total assets	16.3%
DepositRatio	Total Deposits divided by total assets	14.0%
LiquidityRatio	Liquid assets divided by total assets	14.0%
LlpLoans	Loan loss provision divided by total loans	11.6%
MarketPower	Total loans/ total loans of all banks	9.3%
FIreveLiqAssets	Net financial intermediation revenue divided by liquid assets	7.0%
DepositsFunding	Total deposits divided by total funding	7.0%
LoanDeposit	Total loans divided by total deposits	7.0%
NonIntExpRatio	Non interest expenses divided by total assets	4.6%
NonIntExpLiab	Non interest expenses divided by total liabilities	4.6%
AssetDiversity	1 - HHI between earning assets and non earning assets	4.6%
NonDepStFund_TotalFunding	Non deposit short term funding divided total short term funding	4.6%
TradeIncomeTotalAssets	Trading income divided by total assets	4.6%
FIrevenueRatio	F. intermediation revenue divided by total assets	4.6%
CostRevenueRatio	Total costs divided by total revenue	4.6%
G_revenue	Growth of total revenue	2.3%
LlpIntRev	Loan loss provision divided by interest revenue	2.3%
SD_NplTA	Standard deviation of non performing loans divided by total assets of 4 periods	2.3%
BranchRatio	Number of branchs divided by total Assets	2.3%
Log_NPL	Logarithm of non performing loans	2.3%
IBR	Dues from other banks divideds by dues to other banks	2.3%
LiqAs_StFundDep	Liquid Assets divided by short term funding and deposits	2.3%
TranscDep	Transaction depositos divided by total deposits	2.3%
RecLoansProbLoans	Recovered loans divided by problematic loans	2.3%
mRoa_sd	Mean ROA divided standar deviation of roa from the past 4 periods	2.3%
mRoe_sd	Mean ROE divided standar deviation of roe from the past 4 periods	2.3%
1 - HHI_Domestic	1 minuis the HHI of domestic loans and foreign loans	2.3%
1 - HHI_Interest	1 minus the HHI of interest and non interest revenue	2.3%
LoansEarAssets	1 - absolute value of loans minus other earning assets divided by liquid assets	2.3%
G_loans	Growth of loans	2.3%
ROE_Volatility	Standard deviation of ROE	2.3%
ROA_Volatility	Standard deviation of ROA	2.3%
G_deposits	Growth of deposits	2.3%
OthRev	Other operational revenue divided by total assets	2.3%
DepLiab	Total deposits divided by total liabilities	2.3%
NplLlpRatio	Non performing loans plus loan loss provision divided by total loans	2.3%
Npl2Prov	Non performing loans divided by two minus loan loss provision divided by total equity	2.3%
RevBranch	Operational revenue divided by branches	2.3%
NonIntIncTotalIncomeRatio	Non interest income divided by total income	2.3%
TreasuryAssetsRatio	Total assets in the treasury divided by total assets	0%*
TreasuryCreditRatio	Total assets in the treasury divided by total loans	0%*

*: Included by authors

Table 31 – Control variables used in banking literature sampled through 43 papers