Determinants of profitability in the Portuguese banking sector: panel evidence from the period 2015-2023

Pedro Guilherme Branco Martins Pestana Soares

Dissertation written under the supervision of

Professor Diptes Chandrakante Prabhudas Bhimjee

Dissertation submitted in partial fulfilment of requirements for the MSc in Economics with major in Finance & Banking, at the Universidade Católica Portuguesa, September 2024.

{Page intentionally left blank}

Determinants of profitability in the Portuguese banking sector: panel evidence from the period 2015-2023

by Pedro Guilherme Branco Martins Pestana Soares

September 2024

Abstract

The present Dissertation revisits the main determinants of bank profitability in the context of the Portuguese banking sector. For this purpose, the research engages a balanced panel dataset comprising half-yearly observations of the six largest Portuguese commercial banks covering the period 2015-2023. Bank-specific, industry-specific, and macroeconomic variables were selected after critically reviewing the existing academic literature. All data were retrieved from publicly available sources, namely the banks' consolidated financial reports, Bpstat, and Instituto Nacional de Estatística. The first difference technique is applied across the whole panel to ensure stationary dependent and independent variables. After the Hausman specification test deemed the Random Effects approach as efficient, three econometric regressions - one for each dependent variable: ROAE, ROAA, and PBT - were estimated using the PANEL EGLS estimator with cluster-robust standard errors. The baseline findings for bank-specific predictors are generally in line with the existing empirical research: (i) the Equity-to-Assets ratio (capitalization) and Size are positively correlated with bank performance, whereas (ii) the Loan Loss Provisions (credit risk) and Cost-to-income ratio (operating efficiency) have the opposite effect on profitability. The control variables have weak to no effect on the dependent variables. Once the baseline findings are obtained, further three robustness checks are performed to corroborate the overall stability of the baseline findings, namely for: (i) the Random Effects model; (ii) a different solvency-related independent variable; and (iii) the pre-pandemic subsample.

Keywords: Bank profitability, Bank-specific determinants, Control variables, Panel Data, Random Effects model

J.E.L. Classification: C10, E50, G21

{Page intentionally left blank}

Determinants of profitability in the Portuguese banking sector: panel evidence from the period 2015-2023

por Pedro Guilherme Branco Martins Pestana Soares

Setembro 2024

Resumo

A presente Dissertação revisita os principais determinantes da rentabilidade bancária no âmbito do sector bancário Português. Para este efeito, a pesquisa utiliza dados em painel balanceados e compostos por observações semestrais dos seis maiores bancos comerciais Portugueses para o período compreendido entre 2015 e 2023. Os fatores específicos do banco, os fatores específicos do setor e as variáveis macroeconómicas foram selecionadas após a revisão crítica da literatura académica. Todos os dados foram retirados de fontes de informação públicas, nomeadamente os relatórios e contas consolidados dos bancos, Bpstat e o Instituto Nacional de Estatística. Aplicou-se a técnica de diferenciação em todo o painel para salvaguardar que as variáveis dependentes e independentes são estacionárias. Após o teste Hausman ter concluído que o modelo de Efeitos Aleatórios é eficiente, as três regressões econométricas (uma para cada variável dependente: ROAE, ROAA, and PBT) foram estimadas através do estimador PANEL EGLS e recorrendo a erros padrão robustos. Os resultados base para os fatores específicos do banco estão, de um modo geral, em linha com investigação empírica existente: (i) a capitalização e dimensão estão positivamente correlacionadas com a rentabilidade, (i) já o rácio de imparidade (risco de crédito) e o *Cost-to-income* (eficiência operacional) demonstram um efeito contrário. As variáveis de controlo não demonstram efeitos significativos sobre a rentabilidade. Após a obtenção dos resultados base, procedeu-se a três verificações de robustez, corroborando a estabilidade dos resultados base, nomeadamente: (i) modelo de Efeitos Fixos; (ii) nova varável de solvência; e (iii) sub-amostra para o período pré-pandémico.

Palavras-chave: Rentabilidade bancária, Fatores específicos do banco, Variáveis de controlo, Dados em painel, Modelo de Efeitos Aleatórios

Classificação J.E.L.: C10, E50, G21

{Page intentionally left blank}

Table of Contents

Table of Figures

Table of Tables

Glossary

APB – Associação Portuguesa de Bancos

EBA – European Banking Authority

ECB – European Central Bank

EFAP - Economic and Financial Assistance Program

EGLS - Estimated Generalized Least Squares

FE – Fixed Effects

GDP – Gross Domestic Product

GFC – Great Financial Crisis

HHI - Herfindahl-Hirschman Index

INE – Instituto Nacional de Estatística

LDR – Loan-to-Deposit Ratio

LLP – Loan Loss Provisions

NIM – Net interest margin

NPL – Non-Performing Loans

PBT – Profit before tax

POLS – Pooled Ordinary Least Squares

RE – Random Effects

ROAA – Return on average assets

ROAE – Return on average equity

Acknowledgments

After completing all the M.Sc. in Economics courses and joining Caixa Geral de Depósitos, the major milestone still missing was the master Dissertation. It was not by accident that this step took so long to accomplish, as successfully concluding a thesis while working full-time is no easy task. Despite the many times of frustration, I was able to reach this goal that I had for so many years.

I would like to express my deep gratitude to my Dissertation supervisor, Professor Diptes C. P. Bhimjee for his most valuable guidance throughout the whole seminar process. All meetings, comments, and revisions were of the utmost importance in developing the present thesis.

To my parents, thank you for your wholehearted support and inspiring examples of setting goals and achieving them. During my years as a full-time student, I was granted the precious opportunity to focus on my studies without any other concerns. This work is a recognition of your endless care and motivation.

To my friends, thank you for your companionship which has made the Dissertation process more enjoyable.

To my girlfriend, thank you for your tenacity, patience, and much-needed encouraging words which helped me get through rougher times.

I would also like to extend my appreciation to my former colleagues at Caixa Geral de Depósitos as they witnessed my progress and provided helpful advice. I am also grateful to Novo Banco for providing me the chance to conciliate the thesis work with my current position at the bank.

An investment in knowlegde pays the best interest.

Benjamin Franklin

1. Introduction

The surge in profitability registered by Portuguese commercial banks following the pandemic, and the energy-driven inflation aggravated by the war in Ukraine, have brought the narrative of windfall bank profits into the limelight. Indeed, banks' net interest margins have hitherto soared to record highs amid the European Central Bank's contractionist stance to taper inflation. Against this backdrop, bank profits are expected to continue to grow – although at a slower pace – given the precautious data-dependent and meeting-by-meeting approach carried out by the Governing Council.

Although increasing bank profitability is not verified in Portugal alone, the lending-deposit spread of this industry began to increase faster than its European counterparts and is currently higher than most Euro Area countries. In fact, according to the ECB, the average cost of borrowing for Portuguese households was 4.12% whereas the average rate verified in the Euro Area was 3.97%. Further, the average interest rate on deposits from Portuguese households was 3.1% whereas the average rate verified in the Euro Area was 3.29% as of December 2023.

The increasing cost of living and higher debt burden of households, alongside soaring bank profits, originated a great deal of public and political scrutiny on the said banks. While the Portuguese Government imposed temporary windfall taxes on energy and food distribution firms in 2022 and 2023, the banking sector was not required to comply with an additional tax burden on extraordinary profits. However, banks need to sustain the so-called Portuguese Banking Sector Surcharge in place since 2020, despite profitability remaining below the cost of capital only until very recently.

Notwithstanding, other factors determine bank performance besides market rates. These factors have been worthy of pivotal studies and numerous research articles, although the existing empirical evidence on the Portuguese case is quite scarce.

The present Dissertation revisits the main determinants of bank profitability in the context of the Portuguese banking sector, thus providing an important contribution to the existing literature. To address this ever-relevant topic, the present Dissertation uses an econometric methodology supported by a balanced panel dataset comprising the six largest Portuguese commercial banks. The determinants of bank profitability are split into three different groups: i) bank-specific (internal factors), ii) industry-specific, and iii)

macroeconomic; publicly available information is used to derive both dependent and independent variables.

The main findings advocate for a robust effect that capitalization, bank size, credit risk, and operating efficiency have on the various profitability metrics. These relationships are generally in line with the results described in the academic literature. On the other hand, the external factors do not exhibit a statistically strong effect on the dependent variables which is a common characteristic found in the existing empirical studies.

The Dissertation is structured as follows: section 2 delves into the Portuguese banking sector by providing insights on the liberalization process, portraying the dynamics of relevant aggregated indicators for the industry as a whole since the Great Financial Crisis (most of which will be used in the employed methodology), and summarizing the main recapitalization events underpinned by stricter requirements imposed by the Basel III Accord and the Economic and Financial Assistance Program. Section 3 presents the most relevant literature review pertaining to the academic research question addressed herein, including an overview of the different clusters of studies, several variables typically used by practitioners to assess bank profitability, and the conclusions derived from analogous articles. Section 4 describes the econometric specification, the panel dataset and the corresponding data collection procedure, as well as the definition of both dependent and explanatory variables. Section 5 reports and critically discusses the empirical findings associated with the applied methodology in the context of the Portuguese banking sector, including robustness checks to evaluate the stability of the baseline findings. Lastly, section 6 concludes and suggests a direction for future research.

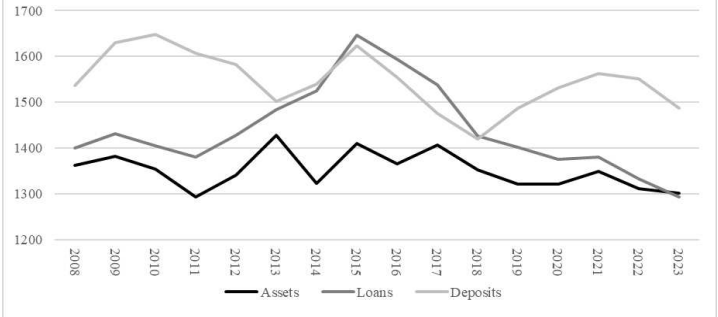
2. Contextual Background

2.1 Developments and key indicators of the Portuguese banking sector

The Portuguese banking sector has accomplished important liberalization milestones throughout the last decades in order to diminish public sector involvement in banking activity and increase competitiveness. This gradual process was chiefly prompted by Portugal's accession to the European Economic Community in 1986 and later on by the complete integration into the European Single Market in 1993, thus contrasting with the then-established background underpinned by an almost fully nationalized banking system since the mid-1970s (Boucinha and Ribeiro, 2009), and characterized by strong regulatory interference, legal constraints, barriers to entry and weak competition (Canhoto, 2004; Ribeiro, 2007). Furthermore, the liberalization of interest rates and the abolishment of credit ceilings materialized in the mid-1990s. This, in turn, readily intensified market competition which used to be heavily based on non-price strategies such as branch expansion, advertising, and product differentiation (Barros and Modesto, 1999). As acknowledged by Mendes and Rebelo (1999), the number of branches and banks operating in Portugal more than doubled between 1985 and 1995.

In parallel, the banking industry has undergone several consolidation events via mergers and acquisitions (more markedly in the 1990s). This consolidation trend stems from the deregulation efforts and spawned increasing market concentration. In fact, M&A activity was the main driver for foreign bank penetration, as opposed to *de novo* growth strategies. On the other hand, privatizations also marked the last decade of the twentieth century: in 1991 there were ten state-owned banking groups, whereas in 1996 there was a single public banking institution – Caixa Geral de Depósitos (CGD). Despite recent fluctuations in market concentration as depicted in Figure 1 (based on the six largest banks), the current metrics are similar to those registered in 2008.

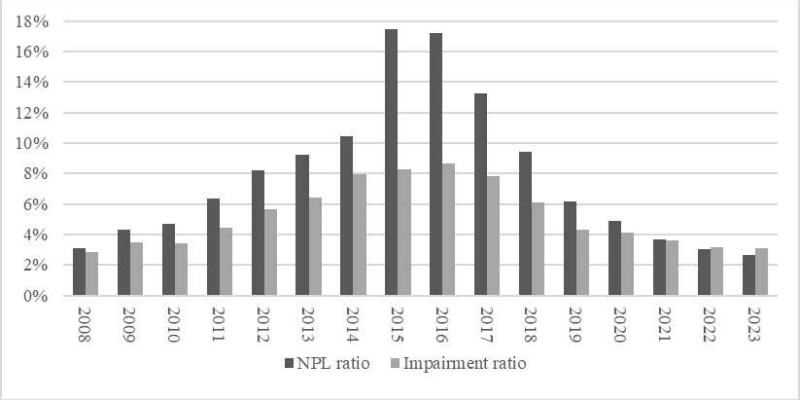
Figure 1 – HHI measured by the six largest banks (2008-2023)



Source: Own calculation based on Bpstat and banks' financial reports

Moreover, the Portuguese financial sector has witnessed a fair number of private and government capital injections. This is particularly true following the Great Financial Crisis (GFC henceforth) and the ensuing European sovereign debt crisis (Augusto and Félix, 2014), the latter being a direct consequence of an international financial contagion process connected to the "Subprime" crisis (Bhimjee, 2013). This global event has led to an increase in the NPL and Impairment ratios up until 2015 and 2016, respectively, as illustrated in Figure 2.

Figure 2 – Credit risk indicators (2008-2023)



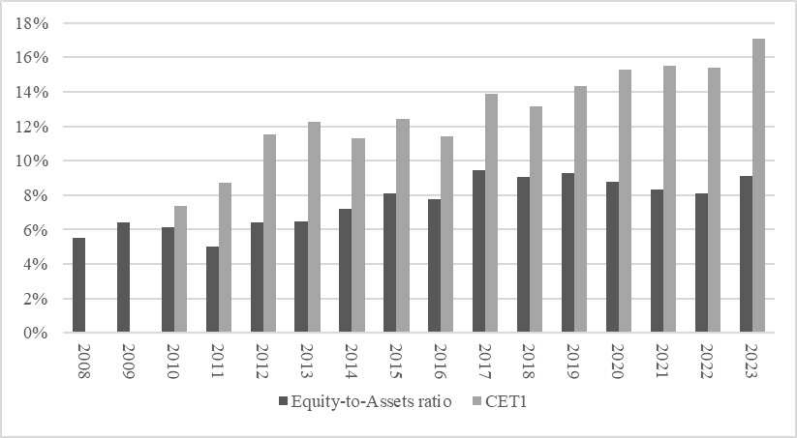
Source: NPL - Bpstat and Statista; Impairment - Own calculation based on Bpstat

Against the background of the GFC's devastating effects on the real economy, the Bank for International Settlements established the Basel III protocol in 2010. This newer Basel Accord framework seeks to implement stricter capital requirements – with the introduction of conservation and countercyclical buffers, the leverage ratio, and higher minimum solvency ratios – as well as novel liquidity standards such as a Liquidity Coverage Ratio and a Net Stable Funding Ratio of at least 100%. In this context, i) the need to comply with a more demanding solvency structure under Basel III, ii) the extraordinary on-site inspections led by the Bank of Portugal, iii) the additional robustness for minimum capital requirements reinforced by the Economic and Financial Assistance Program (EFAP) signed in 2011, iv) the stress test exercises conducted by EBA, alongside v) the worsening macroeconomic conditions, the accumulated bank losses, and the dire funding conditions, ultimately gave rise to numerous bank recapitalizations through the issuance of Contingent Convertible Bonds (CoCos) with high interest rates associated[1]. These bailouts were concurrent with certain private

[1] The pricing is consistent with the discouragement of resorting to State budget funds, ensuring the Portuguese Republic as lender of last resort (Augusto and Félix, 2014).

recapitalizations. Figure 3 outlines the evolution of the Equity-to-Assets ratio and the Common Equity Tier 1, including a noticeable increase in 2012.

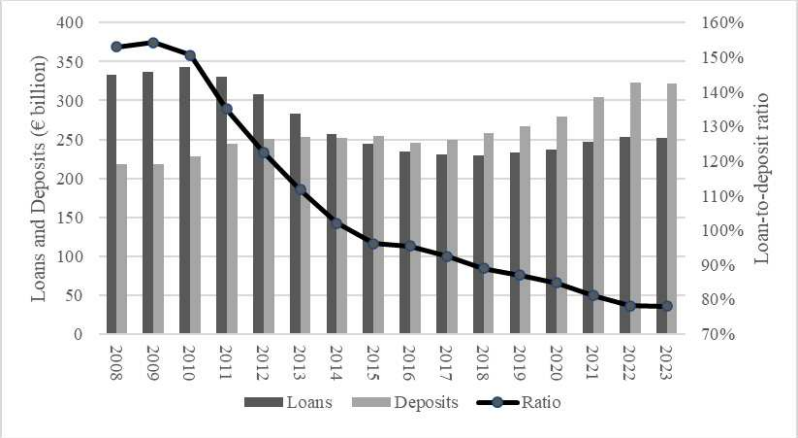
Figure 3 – Solvency ratios (2008-2023)



Source: Bpstat

In short, the EFAP's rescue package totaled €78 billion, of which €12 billion was directly aimed at recapitalizing financial institutions. Furthermore, as part of the Economic Adjustment Programme for Portugal agreed with the "Troika" - and after a recent credit boom - the main banking groups should envisage measures to decrease the loan-to-deposit ratio to below 120%, depicting the onset of a long deleveraging process in the domestic financial system (Crosignani et al., 2015) which was largely targeted to the private sector. Accordingly, households and corporations reduced consumption and postponed further investments, despite negative real interest rates after 2012. Part of the deleveraging process was due to a steady increase in deposits, in great contrast to Greece where there were many bank run episodes. The bank deposits to GDP ratio is currently 122%, translating the economic importance of these institutions in Portugal. Figure 4 portrays the deleveraging process after the GFC.

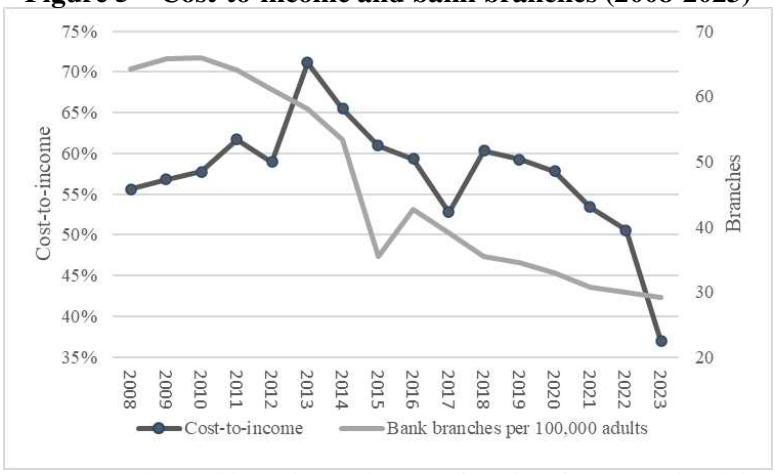
Figure 4 – Loan-to-deposit ratio (2008-2023)



Source: Bpstat

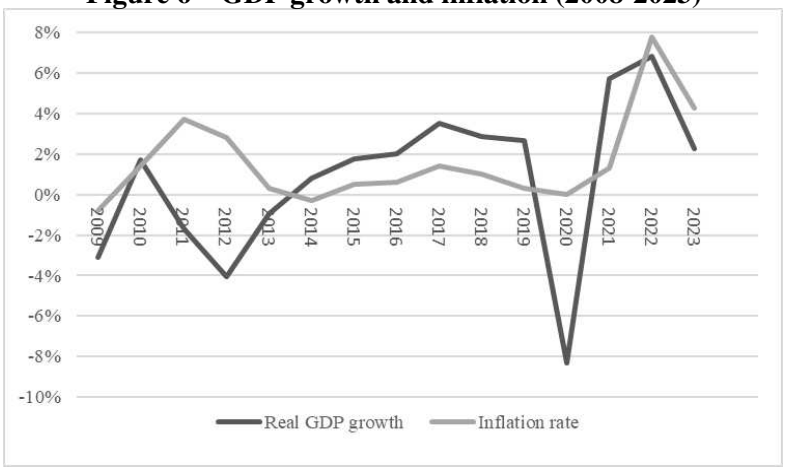
Moreover, in contrast to the diminishing levels of the Loan-to-deposit ratio, the Portuguese banking sector has been witnessing increasing levels of operating efficiency in the last decade, in part explained by the closure of bank branches (Figure 5), a reduction in the workforce, digital transformation, and automatization of repetitive processes as well as increasing key ECB rates amid the recent hawkish narrative to address record high inflation (Figure 6).

Figure 5 – Cost-to-income and bank branches (2008-2023)



Source: Bpstat and World Bank. Author's estimation for 22-23 bank branches

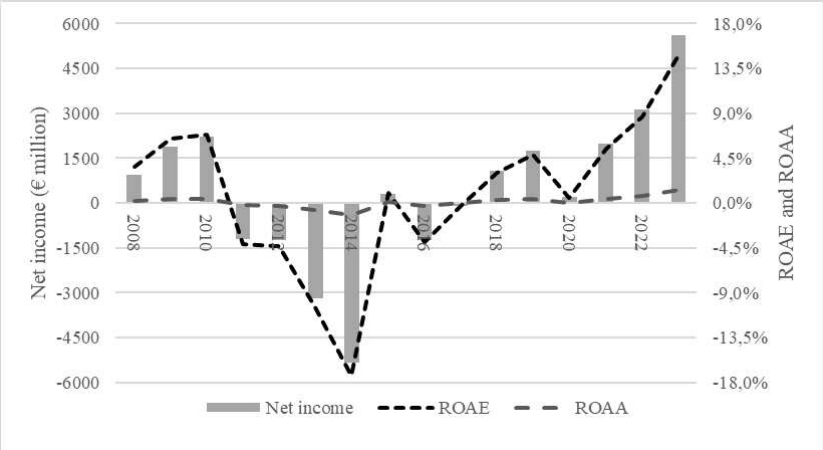
Figure 6 – GDP growth and inflation (2008-2023)



Source: Instituto Nacional de Estatística

As pictured in Figure 7, bank profitability has soared in recent years. While increasing Euribor rates (market benchmark) and operating efficiency are key drivers for such an evolution, the upward revision in commission pricing and the reversal of loan loss provisions played an important role in boosting profitability before the ECB took action against the general surge in prices of goods and services in the post-pandemic phase.

Figure 7 – Net income and profitability (2008-2023)



Source: Bpstat

2.2 Main recapitalizations in the industry

The most notable bailouts occurred in June 2012: in the case of Millennium BCP, the issuance of CoCos totaled €3 billion; regarding BPI, the issuance amounted to €1.5 billion; as for CGD, the bailout corresponded to €900 million. Shortly after, BANIF was rescued in December 2012 through the issuance of the above-mentioned hybrid debt security worth €400 million which was complemented in early 2013 by another €700 million in new shares with a subscription price of 1 cent per share. Notwithstanding, BANIF couldn´t avoid a resolution in December 2015 in the wake of a bank run in the same month. This bailout required circa €1.76 billion in state aid and €489 million from the Resolution Fund. The higher-quality assets were sold to Santander Totta for €150 million.

In 2014, Banco Espírito Santo (BES) entered into resolution following a reported €3.6 billion in losses during the year's first half given the imprudent loans conceded to the nonfinancial arm of the Espírito Santo (ES) group. That said, the “toxic” assets were transferred to a so-called “bad bank” whereas the remaining assets were moved to the newly formed Novo Banco, the latter with a bridge bank status. This operation entailed a €4.9 billion capital injection from the Resolution Fund, of which €1.0 billion was funded by private banks. In 2017, Lone Star acquired a 75% stake in the institution and injected an additional €1.0 billion, thus ceasing the status as a bridge bank. Moreover, Novo Banco raised €3.4 billion from the Portuguese Republic over 2018–2021 under the Contingent Capital Agreement. At present, Senior management intends to pursue an IPO after lifting

the ban on dividend payout, currently in force under the terms of the bailout until the end of 2025.

In 2017, CGD underwent a recapitalization amounting €3.9 billion. Therefore, the state-owned bank was required to implement a strategic plan (2017-2020) which was designed by the Portuguese authorities in conjunction with the European Commission. The plan encompassed scaling down international operations, reducing the number of branches, trimming the cost-to-income ratio, and reducing non-performing assets (Stadheim, 2020). Prior to this recapitalization, the bank's executive committee conducted a Management Assessment of Asset Value (MAAV) in 2016. This translated into €2.4 billion in additional provisions that year and reclassification of assets to the delinquency categories. Nonetheless, DG Comp and Credit Rating Agencies recognized the overall success of the strategic plan's implementation. The strategic plan (2021-2024) is currently underway, wherein the Executive Committee's focus is to increase operating efficiency, accelerate digital transformation, and strengthen asset quality (with special emphasis on reducing the NPL ratio in the non-financial corporations' portfolio).

Millennium BCP raised €1.3 billion in 2017 with the issuance of new shares. Part of this capital injection was aimed at increasing the soundness of the balance sheet, whereas €700 million was allocated to reimburse the Portuguese State concerning the CoCos.

3. Literature Review

3.1 Overview

Existing studies on the determinants of banks' profitability seek to establish either a correlation or causation between a set of explanatory variables and some key profitability ratios. A plethora of research questions are scrutinized under the scope of the above-mentioned topic, hence the broad array of sample sizes, input variables, and statistical techniques employed among the various articles/papers. Such variety in research methodologies ultimately yields differences in explanatory power and empirical results derived from each research. Be that as it may, practitioners have pinpointed some common variables that influence profitability.

Empirical literature covering bank profitability stems from earlier studies (Short, 1979; Bourke, 1989) and can be divided into multiple-country studies (usually encompassing banks from several countries within a given region) and single-country studies (country-specific) where emphasis is placed on the banking sector of a single country, as pointed out by Tan (2016). Most articles belong to the latter cluster of studies (Ercegovac et al., 2020).

Contrary to the case of independent variables discussed next, the choice of dependent variables is typically rather limited. In general, authors consider the return on equity (ROE), return on assets (ROA) and net interest margin (NIM) as profitability ratios. Furthermore, a ratio considering profit before tax (PBT) on the numerator is also considered but less frequently (e.g., Tan, 2016). Notwithstanding, these accounting-based ratios are universally used by banks to report performance and thus deemed appropriate to address research questions alike.[2]

Where the case of the explanatory variables is concerned, there is a fair number of determinants proposed in academic research, all of which fall into two categories: "bank-specific (internal) factors and industry-specific and macroeconomic (external) factors" (Petria, Capraru, and Ihnatov, 2015: p. 520). On the one hand, the former category comprises intrinsic characteristics of each financial institution resulting from their business and risk management activities. Examples of internal factors are "ratios

[2] In the case of the Portuguese banking industry, Banco de Portugal has issued the Instrução n.º 6/2018 which contains three profitability ratios that every bank should disclose to the public, besides any other metric the management team decides to report. PBT is one of those three metrics and thus will be used in this Dissertation aside from ROAE and ROAA.

representing cost efficiency, liquidity, asset quality, and capital adequacy" (Kosmidou et al., 2005: p. 6). On the other hand, the latter category includes control variables not directly linked to bank management activities but rather to the social, economic and legal environment underpinning the operation and performance of a particular banking industry[3] (Athanasoglou, Brissmis and Delis, 2005). Most often, industry-specific factors are described by ownership status, industry size and market concentration. Concerning competition, authors frequently resort to the Herfindahl-Hirschman Index[4] (HHI) or n-firm concentration ratio (Berger et al., 2004). Moreover, the Lerner index is also used as a control variable for market concentration, as in Cipollini and Fiordelisi (2012) and Tan (2016), although with diminished ease of implementation alongside some criticism (e.g., Lindenberg and Ross, 1981; Elzinga and Mills, 2011). In their turn, macroeconomic determinants for profitability are typically perceived to be inflation, cyclical output and national accounts aggregates.

3.2 Relevant empirical studies

As previously mentioned, many statistical techniques have been used to conduct investigations regarding the relationship between a proposed set of determinants and bank profitability. The chosen research methodology for this Dissertation engages a panel data model. Therefore, only articles relying on this type of specification - while still covering the most scrutinized variables - will be addressed herein.

Goddard et al. (2004) assess the relationship between a set of bank-specific variables and profitability in six European banking sectors: Denmark, France, Germany, Italy, Spain and the UK, for the 1992–98 period. A total of 665 banks are selected, with some countries being excluded due to a lack of data (including Portugal). Only ROE is considered as a performance measure since these authors acknowledge that off-balance-sheet (OBS) activities contribute to total profit (numerator) but are excluded from the denominator of ROA[5]. The main explanatory variables are as follows: i) natural logarithm of total assets (size); ii) OBS business; iii) Capital-assets ratio (CAR); and iv) ownership type. Accounting data are retrieved from Bankscope Database. An unbalanced dynamic panel dataset with 3325 observations is built and the two-step GMM dynamic panel

[3] Albeit banks' profitability in the same industry may be affected by external factors to different extents.
[4] It accounts for the share of each bank in the market (for a given metric), endows greater weight to the firms with larger market shares (Petria et al., 2015) and is calculated on a country-specific basis.
[5] Results using ROA were qualitatively similar to those using ROE as the year-on-year variation in the numerator is greater than in the denominators of both measures.

estimator is employed. Their main findings suggest that size lacks statistical significance except for the German and UK banking sectors. In Germany, the lower the size the higher the profitability. The opposite is true for the UK, thus implying the existence of scope economies among larger banks in such a highly concentrated banking sector. Off-balance-sheet items display different effects between countries, with the UK comprising the only sector where OBS activities positively influence profitability. Regarding CAR, the empirical results attained in the article differ from the authors' initial expectations because a high CAR translates into weakened profitability. In contrast, possible explanations that have been put forward for an expected positive relationship are that high capitalization decreases the costs of insurance against bankruptcy, reduces funding costs, and acts as a signaling hypothesis. According to the latter, managers make use of robust solvency to send signals about future profitability. Moreover, there is a paucity of evidence that sustains a systematic relationship between ownership status (savings, cooperative, and commercial banks) and performance. Lastly, the authors also employ the lagged dependent variable as an additional predictor, which is shown to have a persistent positive effect on the following year's profitability.

Kosmidou et al. (2005) research the effect of both internal and external factors on UK-based commercial banks' profitability during the 1995-2002 period. The selected performance measures are the return on average assets (ROAA) and NIM. Regarding the former, the authors confirm its utmost importance in comparing bank profitability, while averaging assets on the denominator to control for fluctuations that occur during the fiscal year. The chosen bank-specific variables are as follows: i) cost-to-income ratio (efficiency); ii) liquidity; iii) Loan Loss Reserves to Gross Loans (credit risk); iv) Equity-to-Asset ratio (leverage); and v) Total assets (size). Industry-specific variables are the five-firm concentration ratio and the market capitalization to total assets of the deposit money banks (stock market development). As for macroeconomic variables, the article resorts to GDP growth and inflation. The first six variables are retrieved from Bankscope Database while the others are available in the Euromonitor International Database. An unbalanced panel dataset[6] of 224 observations encompassing 32 banks is built and all models are estimated using fixed-effects regression (the random effects model was disregarded after the Hausman test). Regarding internal factors, empirical results demonstrate that leverage is the most statistically significant variable positively affecting

[6] Not all banks had the same number of observations.

profitability, suggesting that well-capitalized banks face lower funding costs. Cost-to-income ratio and bank size are also deemed significant, negatively affecting bank profits. The effect of liquidity and credit risk on bank performance is not clear-cut and exhibits opposite impacts on ROAA and NIM. In addition, all external variables are found to be significant and positively related to performance, albeit adding a somewhat small change to the explanatory power of the model.

Athanasoglou et al. (2006) address the Southeastern European (SEE) region by analyzing the effect of a set of determinants on bank profitability (commercial and savings banks) in seven countries, namely Albania, Bosnia-Herzegovina, Bulgaria, Croatia, FYROM, Romania and Serbia-Montenegro, for the 1998–2002 period. The dataset includes 71 banks in 1998 and 132 in 2002, covering roughly 80% of the industry's total assets. The chosen proxies for profitability are ROA and ROE[7]. Internal factors considered are i) Loans-to-Assets ratio (liquidity); ii) average loan loss provisions to total loans ratio (credit risk); iii) Equity-to-Asset ratio (capitalization); iv) ratio of operating expenses to total assets (efficiency); v) logarithm of bank assets (size); and vi) foreign ownership. Industry-specific variables are market concentration (proxied by the HHI) and the EBRD index to capture the degree of liberalization and institutional reform of the banking sector. The macroeconomic variables used are inflation and real GDP per capita (RGC). Annual bank data is retrieved from Bankscope Database, the banking reform index from the European Bank for Reconstruction and Development, and macroeconomic variables from IMF's International Financial Statistics. An unbalanced panel dataset is formed after data cleansing for inconsistencies. The GLS estimator is employed using the random effects (RE) model as the Hausman test strongly recommends this approach. This article's findings suggest that liquidity has a positive but nonsignificant effect on profitability, possibly because the SEE banking system's resources are still insufficient to meet the liquidity standards of the developed banking systems. Another surprising result concerns RGC with no relevant effect on the dependent variables, perhaps due to the hawkish monetary stance that constrained bank lending. The results for the remaining variables are significant and in line with expectations: credit risk presents a negative effect, capitalization exhibits a positive impact, operating expenses diminish profitability, bank size suggests economies of scale, foreign banks outperform their peers, concentration increases profitability, the improvement in the regulatory framework

[7] Despite using this nomenclature, the authors consider running yearly averages and PBT on the numerator as opposed to net income.

increases competition which narrows the interest rate spread and finally inflation causes bank income (e.g., interest income and commissions) to increase more than costs (e.g., interest on deposits and operating costs).

Alexiou and Sofoklis (2009) conduct research to infer the most relevant factors underpinning the profitability of the six major Greek commercial banks for the 2000-2007 period[8]. ROE and ROA are chosen as dependent variables. The latter is assumed to be an indicator of managerial efficiency and the former is a proxy for the net benefit that the shareholders received from investing their capital in the firm. The following internal factors are considered: i) ratio of provisions or allowance for loan losses to total net loans (credit risk); ii) Equity-to-Asset ratio (contemplates a bank's ability to absorb unexpected losses); iii) logarithm of bank assets (seeks to accommodate a nonlinear size-profitability relationship); iv) Loan-to-Deposit ratio (liquidity); v) cost-to-income ratio (efficiency); and vi) net assets over the number of employees (productivity). Macroeconomic factors considered are inflation (proxied by CPI), GDP growth and private consumption. Internal factors are obtained from the published financial statements of the banks, whereas macroeconomic variables are collected from the National Statistics Bureau and the ECB. These authors employ a panel data analysis by utilizing both fixed and random effects models and resorting to pooled OLS and GLS. The fixed effects model is preferred after considering the Hausman test. All bank-specific estimated parameters pass the t-test at the 1% significance level. Credit risk has a negative effect on the dependent variables, suggesting the higher the level of defaulted loans[9] the lower the profitability. Capital displays a positive relationship with profitability. Bank size seems to have a positive effect on ROE and ROA, thus indicating that larger firms with a broader product offering can secure funding at lower costs. Liquidity showcases an inverse relationship with the dependent variables, outlining the trade-off between liquidity and profitability. Regarding efficiency, it comes as no surprise that the higher the ratio the lower the performance measures. Unexpected results are achieved for productivity with the model suggesting a negative effect on performance. On the other hand, external factors are found to have a positive but insignificant effect on profitability.

Petria et al. (2015) present the first-ever study that evaluates bank performance in an overarching framework covering all EU 27 countries over the period 2004-2011. The said research encompasses 1098 commercial banks and there was no need to exclude any

[8] These banks comprise more than 80% of the market share in terms of assets, deposits and loans.
[9] Usually associated with a higher NPL ratio.

financial institution due to data availability issues. The performance proxies are ROAE and ROAA, the latter considering the risks derived from the leverage. The following internal factors are considered: i) natural logarithm of total bank assets (size); ii) Equity-to-Asset ratio (capital adequacy); iii) ratio of loan loss reserves to gross loans (credit risk); iv) loans to customer deposits ratio (liquidity risk); v) cost-to-income ratio (efficiency); and vi) ratio of other operating income to the average bank assets (business mix indicator). Market concentration is proxied by the HHI while the chosen macroeconomic variables are inflation and GDP per capita growth. Bank-related yearly data was downloaded from Bankscope database, inputs for the HHI were derived from ECB Statistical Data Warehouse and macroeconomic variables were retrieved from the World Bank database. For this research, the authors' methodology engages a fixed effects estimation model after employing the Hausman test. Additionally, robust standard errors were estimated to safeguard the covariance estimator against possible heteroscedasticity. Bank size positively influences ROAA, although with a somewhat weak statistical significance[10]. Capital adequacy has a statistically significant upward effect on ROAA, albeit with a low coefficient value. Moreover, credit risk exhibits a negative, statistically significant impact on ROAE and ROAA. Regarding liquidity, the loan-to-deposit ratio negatively and significantly affects the dependent variables. The same applies to efficiency, however, the effect on ROAE is much larger. The diversification of business displays a positive relationship with profitability. In what concerns concentration, the higher the HHI the lower the profitability. GDP growth positively affects profitability as expected. Inflation on the other hand, seems to have no explanatory power on the left-hand side of the econometric equation, possibly because senior management isn´t able to properly forecast future inflation.

More recently, Borges and Tavares (2020) review the determinants of the profitability of the Portuguese banking sector over the period 2005-2011. To this end, the researchers built a balanced panel dataset of 126 observations composed of 18 banks which comprise 93% of the total consolidated assets of the Portuguese banking system. ROA and ROE are employed to capture bank performance. In terms of bank-specific factors, the following variables were selected: i) Equity-to-Asset ratio (capital); ii) cost-

[10] As one can infer from the reading thus far that there is no consensus regarding the effect size has on profitability. While some practitioners argue for scope economies, product differentiation, and too-big-to-fail status which provides larger banks with an implicit government guarantee (Iannotta et al., 2007), others suggest an inverse relationship (e.g., Stiroh and Rumble, 2006; Barros et al., 2007).

to-income ratio (efficiency); iii) ratio of net current assets over short-term liabilities and the ratio of loans granted over total assets (liquidity risk); iv) ratio of provisions over loans granted (credit risk); v) ratio of other operating income over total assets (diversification); and vi) total assets of the bank (size). In what concerns external factors, the HHI is used as a proxy for market concentration. Inflation, GDP growth, and the unemployment rate represent the macroeconomic variables. Yearly consolidated accounting data was made available by the Portuguese Banking Association (APB), inputs for the HHI were obtained from a database of the Banco de Portugal, and macroeconomic variables were extracted from the Eurostat and ECB databases. A fixed effects model is utilized after the Hausman test concluded it best fits the data. The EGLS method is employed to prevent possible heteroskedasticity or auto-correlation issues. Empirical findings indicate capital[11], efficiency and liquidity (ratio of net current assets over short-term liabilities) as being statistically significant and exhibiting the expected coefficient sign. Credit risk, inflation and GDP growth also displayed strong explanatory power but only concerning ROA. Diversification and size have a weak impact on profitability, whereas market concentration has strong significance regarding ROE with a positive sign.

An overview of the articles surveyed in the present section allows us to conclude that the most common bank-specific factors are (i) capitalization, (ii) liquidity, (iii) credit risk, (iv) size, and (v) efficiency, all of which are based on accounting data. While credit risk and efficiency suggest negative and positive effects on profitability, respectively, the remaining variables don´t exhibit a unique and consensual relationship with the dependent variables. Regarding external factors, market concentration can have distinct effects on bank performance; moreover, inflation and GDP growth coefficients usually display positive effects, although not always with p-values close to zero.

We contribute to the empirical literature by examining the effects of some prominent determinants of bank profitability within the largest full-service Portuguese commercial banks. To the best of our knowledge, this Dissertation arises as the first attempt to study bank performance in the Portuguese banking sector over a period that encompasses the Covid pandemic as well as the more recent inflationary episode, including full-year 2023 data.

[11] Capital positively affects ROA, and the opposite is verified for ROE.

{Page intentionally left blank}

4. Methodology and Data

This section describes how the proposed research question is addressed and encompasses an overarching framework for the econometric approach herein employed, as well as the data collection procedure. The description of variables and expected coefficient signs are also discussed.

4.1 Methodology

As mentioned earlier, this single-country study uses a panel data model. This type of longitudinal data was selected for this research given the several advantages offered over time series and cross-sectional data as pinpointed by Hsiao (1986) and Marques (2000). Econometric models resorting to panel data control for heterogeneity among cross-sections (i.e., banks), and allow for a larger number of observations, greater degrees of freedom, and diminished multicollinearity between independent variables, all of which contribute to more robust estimated parameters. Moreover, longitudinal data minimizes the so-called omitted-variable bias, which may result in endogeneity[12].

The econometric technique stems from the work of Molyneux and Thornton (1992), Bourke (1989), and most importantly Athanasoglou et al. (2006) covered in the previous section, wherein the general specification is as follows:

$$\pi_{i,t} = c + \beta_1 X_{it}^{b} + \beta_2 X_{it}^{n} + \beta_3 X_{t}^{m} + \varepsilon_{i,t}\,, \quad \varepsilon_{i,t} = v_i + u_{i,t} \quad (1)$$

where $\pi_{i,t}$ represents the dependent variable of bank *i* in period *t*; *c* is the constant term of the panel regression and can be interpreted as the expected profitability for all banks *i* in period *t* granted that every independent variable is null. The chosen set of explanatory variables is comprised of X_{it}^{b} which accounts for all bank-specific variables (or internal factors), X_{it}^{n} that is the industry-specific variable and X_{t}^{m} which constitutes the macroeconomic variables[13]. β_1, β_2 and β_3 are the coefficients of each independent variable X_{it}^{b}, X_{it}^{n} and X_{t}^{m}, respectively. In its turn, $\varepsilon_{i,t}$ corresponds to the random error term (or residual) described by the unobservable unit-specific effect v_i (time-invariant)

[12] In this specific case of endogeneity, the error term is correlated with a predictor that was omitted from the model (as opposed to being correlated with an independent variable within the model).
[13] Unlike the other explanatory variables, X_{t}^{m} does not vary with *i* as each value in period t is the same across all banks, hence the absence of *i* subscript.

and the idiosyncratic error $u_{i,t}$ (or time-varying error), both of which affect the dependent variable (Wooldrige, 2012).

Panel regression can be conducted with three distinctive approaches as pinpointed by Alexiou and Sofoklis (2009). The first is Pooled OLS wherein the parameters α and γ are the same across banks:

$$y_{it} = \tau\alpha + X_{it}\gamma + \varepsilon \quad (2)$$

where τ is a column of vectors *(TN × 1).* The GLS estimator is then reduced to Pooled OLS. The second approach is the Random Effects model wherein the intercept varies across banks and is time-invariant. The remaining coefficients on the other hand are the same across banks and are time-invariant as well:

$$y_{it} = \alpha + X_{it}\gamma' + \mu_i + \varepsilon_{it} \quad (3)$$

where $E(\mu_i=0)$, $E(\mu_i^2) = \sigma_\mu^2$, $E(\mu_i\mu_j) = 0$ for $i \neq j$ and $E(\varepsilon_{it}\mu_j) = 0$, for all *i, t* and *j*. The μ_i represents the random disturbance for each *i* observation. This model can be estimated by resorting to GLS. The third approach is the Fixed Effects model (or Least Squares Dummy Variables model) which assumes that differences between banks are captured by the constant-term α_i (time-invariant):

$$y_{it} = \alpha_i + X_{it}\gamma' + \varepsilon_{it} \quad (4)$$

where ε_{it} assumes $E(\varepsilon_{it}) = 0$ and variance σ_ε^2. This model controls for omitted-variable effects that vary across banks and remain constant through time.

The employed econometric methodology comprises the following steps: First, we apply the Fisher, Levin-Lin-Chu, and Im-Pesaran-Shin unit root tests to assess whether all variables are stationary in a balanced panel setting. Each variable is considered stationary if it rejects the null for at least two tests. Second, we examine which of the three approaches presented above is appropriate for the dataset. For this purpose, we argue against the case of Pooled OLS (POLS) and resort to the Hausman test to opt between Fixed Effects and Random Effects. Third, after concluding that Random Effects is the suitable approach (section 4), a specific regression is performed for each independent variable using cluster-robust standard errors (STATA command: vce(cluster i)) to control for collinearity, heteroscedasticity, and autocorrelation as we consider each bank *i* a different cluster and the observations of a given period are related to the ones of the

previous period for the same bank. Therefore, following Garcia and Guerreiro (2016) and Pires et al. (2021), we assume a more sensible assumption of intragroup correlation in contrast to the typical prerequisite that observations are independent. That said, observations are independent across clusters but not necessarily within them. Fourth, the baseline findings are subject to robustness checks to assess their overall stability.

4.2 Data

This empirical research utilizes a balanced panel dataset of 108 observations covering six of the largest Portuguese commercial banks (in terms of balance sheet's total assets): (i) Caixa Geral de Depósitos, (ii) Millennium BCP, (iii) Banco Santander Totta, (iv) Novo Banco, (v) Banco Português de Investimento, and (vi) Banco Montepio. The sample encompasses the period ranging from 2015 to 2023. It begins in 2015 as this was the first full year of Novo Banco's activity after the former Banco Espírito Santo entered into resolution, thus establishing its position as the current fourth largest bank in Portugal.

According to the Bank of Portugal, there are 62 domestic and foreign banks operating in Portugal as of August 2024. Nevertheless, the sample is assumed to be representative of the Portuguese banking sector since the chosen banks comprise 79% of total assets, 81% of loans granted to customers, and 82% of customer deposits held as of March 2024[14].

Each observation corresponds to a timespan of six months (one semester). The most relevant dependent and explanatory variables were chosen by critically analyzing the existing academic literature. Information regarding internal factors (including profitability) was retrieved from each bank's audited consolidated financial reports (half-yearly and end-of-year). Consolidated information was considered as opposed to standalone financial statements as the former is i) more readily available, ii) the complementary information is typically more comprehensive for consolidated figures, and iii) interest lies in examining different economic units rather than legal entities. Regarding control variables (or external factors), industry-specific data was collected from Bpstat (Bank of Portugal), whereas macroeconomic data were downloaded from Instituto Nacional de Estatística (INE). All input variables derive from own calculations based on publicly available information. The variables were then organized in long format

[14] The highest figures observed in the sample were actually registered in December 2015 wherein the six banks represented 84% of total assets, 90% of total loans granted, and 87% of total deposits.

and uploaded into STATA 18.5 statistical software to employ the model specification set forth herein.

4.2.1 Dependent variables

Analogously to many empirical studies, this research relies on some key profitability ratios, namely (i) Return on Average Equity (ROAE), (ii) Return on Average Assets (ROAA), and (iii) Profit Before Tax (PBT).

The first metric is computed by dividing the annualized net income of one semester (retrieved from the income statement) by average equity (made available in the balance sheet). Net income of each semester was multiplied by two since annualized figures were used in the numerator for all performance ratios. In addition, considering the Instrução n.º 6/2018 of Bank of Portugal as background (as well as the empirical literature), all denominators were averaged out, thus taking into consideration changes throughout each semester. ROAE is usually given more emphasis in banks' earnings releases and is oftentimes used as a metric for C-level Executive compensation[15].

Similarly, ROAA is determined by resorting to annualized net income, the difference being in the denominator which uses the average of book total assets (gross assets minus loan loss provisions and amortizations). This ratio is thus much smaller than the previous one and captures information regarding a higher risk component associated with high leverage (Athanasoglou et al., 2006).

On the other hand, PBT was also selected for this empirical study since it is part of the above-mentioned Instrução, albeit not so commonly considered in the academic literature[16]. Nonetheless, such metric is calculated by dividing profit before tax (income statement) over the average of book total assets. That said, the numerator is not impacted by i) distortions arising from the tax system (in the form of current and deferred corporate income tax), ii) results of associates and jointly controlled entities, and iii) results of subsidiaries held for sale. Including a third dependent variable provides added robustness when evaluating the regressor's effect on the left-hand side of the equations. All profitability measures are expressed in percentages following most research papers.

[15] Or a variation of this ratio such as return on tangible equity (ROTE).

[16] Although also covered in section 3, NIM was not studied in the present Dissertation due to the lack of publicly disclosed information necessary for the denominator: Interest Earning Assets. Some articles depart from the classic definition by resorting to Total Assets in the denominator (and net interest income in the numerator). For example, as of December 2023, Novo Banco's Interest Earning Assets comprises 94% of total assets; in the case of Millennium BCP this figure accounts for 88% of total assets; as for Banco Montepio this metric accounts for 94% of total assets.

4.2.2 Independent variables

The engaged econometric model outlines a set of five bank-specific characteristics and three external factors, the latter divided into sectorial and macroeconomic control variables. The following internal factors are analyzed as follows:

Capital ratio quantifies the capital of each financial institution and is calculated by dividing total equity (including non-controlling interests) by total assets according to each bank's balance sheet. We expect a positive estimated coefficient in line with the majority of the research articles.

Size is measured by taking the common logarithm (base 10) of total assets (EUR million) as depicted in the balance sheet. As shown in section 3, there has been considerable debate in the academic literature as to whether size has a positive or negative effect on the dependent variables, hence there is no clear-cut relationship between size and profitability. While some advocate in favor of economies of scope, others ascertain the adverse impact of higher agency costs underpinned by larger banks (Pasiouras and Kosmidou, 2007).

Loan-to-Deposit ratio (also known as the Transformation ratio), as the name suggests, is computed by dividing total loans granted to companies and households (after deducting any impairments, as described in Instrução n.º 6/2018) by total customer deposits. The former is described in the Assets side of the balance sheet whereas the latter is displayed in the liabilities side. Although increasing deposits may stimulate performance as it is a funding source with lower costs than external financing (García-Herrero et al., 2009) which is especially important for commercial banks to finance long-term loans with short-term deposits (maturity transformation), we expect a positive effect on profitability since the main source of a commercial bank's income derives from interest payments (especially during a hawkish stance carried out by the central bank).

Loan Loss Provisions (also known as the Impairment ratio) represents credit risk and is calculated by dividing loan loss provisions (written on the Asset side of the balance sheet) by total gross loans[17]. A higher ratio is usually tied to a higher stock of Non-Performing Loans (NPL) which comprise loans that entered default - or stage 3 classification under IFRS 9 - hence the expected negative effect on bank performance. It should be noted, however, that banks engage in forward-looking loan-loss provisioning (Peterson and Arun, 2018), meaning that a specific impairment can be allocated to a

[17] Loan loss provisions are either computed by the bank's collective impairment model (calibrated annually) or by an individual (case-by-case) impairment assessment of each debtor with significant exposures.

certain loan even before the debtor (i.e., borrower) fails to meet the payment schedule (a hallmark feature of IFRS 9).

Cost-to-income ratio represents efficiency and consists of the ratio between operating costs (i.e., employee costs, administrative expenses, and depreciation and amortization) and operating income (e.g., net interest income, non-interest income derived from services, fees, and commissions, among others). Both the numerator and denominator are available in the condensed income statement. This metric is frequently included in banks' official strategic plans (or guidance for the next years) given its importance in assessing and monitoring efficiency. Although each bank may disclose a recurrent cost-to-income ratio in its financial reports, the present Dissertation employs the definition outlined in the aforementioned Instrução. This, in turn, allows avoiding any distortions arising from each institution's definition of what constitutes non-recurring items that would be deducted from the recurrent ratio.

Regarding external factors, *Market concentration* has been chosen as the industry-specific variable. Such explanatory variable is proxied by the Herfindahl–Hirschman Index (HHI) which is obtained by summing the squared market share that each bank has in the industry they operate in. Most articles addressing similar research questions either compute the market share in terms of (i) total assets, (ii) loans granted, or (iii) deposits. The present research uses the first item since it's the most recurrent in the academic literature. Each bank's market share is retrieved by dividing its total assets by the total assets embedded in the Portuguese banking sector according to Bpstat[18]. According to APB (2021), an HHI of less than 1000 translates into a low-concentrated industry, between 1000 and 1800 constitutes a moderately concentrated industry, and above 1800 comprises a highly concentrated market. Taking the structure–conduct–performance (SCP) as background and the literature review, we expect that the higher the level of market concentration the higher the performance, as greater market power may lead to collusive behavior. However, it is worthwhile to point out that not all articles advocate for undeniable evidence supporting the SCP paradigm. For instance, Gilbert (1984) delves into 45 empirical studies, of which only 27 suggest that the SCP holds.

Concerning macroeconomic variables, the *Real GDP growth* has been chosen to retrieve the relationship between country output (adjusted for inflation) and bank profitability. Total output in each semester is compared against the same period of the

[18] BPstat series ID 12504494.

previous year (based on constant prices). It is expected to have an impact on various factors related to the demand and supply of bank deposits and loans (Anbar and Alper, 2011). Moreover, economic growth may accommodate lower provisioning levels (Athanasoglou et al., 2014). Following the academic literature review, we expect a positive influence on the independent variables.

The second macroeconomic variable employed is *Inflation*, which is based on the monthly Consumer Price Index (CPI) for all goods and services published by INE. The average monthly CPI in each semester is compared against the same period of the previous year. The link between inflation and profitability may present either a positive or negative correlation depending on whether it is anticipated or unanticipated (Perry, 1992). If inflation is properly forecasted, banks can adjust interest rates to increase revenues at a faster pace than costs. It should be noted, however, that companies and households may register an excessive debt burden if interest rates increase significantly as part of a contractionary monetary policy, which in turn leads to higher loan provisioning and lower profitability. Nevertheless, we expect the estimated coefficient to exhibit a positive sign as the majority of existing studies point in this direction. It is relevant to include these control variables since the dataset comprises periods of turmoil in this particular sector.

The table below presents a summary of both independent and dependent variables.

Table 1 – Definition of the variables and expected effect

Dependent variables	Definition	Expected Effect
ROAE	Net income (annualized) over Average Total Equity (%)	
ROAA	Net income (annualized) over Average Total Assets (%)	
PBT	Profit before tax (annualized) over Average Total Assets (%)	
Independent variables		
Bank-specific		
Capital Ratio	Total Equity over Total Assets (%)	+
Size	Logarithm of Total Assets (EUR Million)	+/-
Loan-to-Deposit ratio	Total Loans over Total Deposits (%)	+
Loan Loss Provisions	Loan loss provisions over Total Gross Loans (%)	-
Cost-to-income ratio	Operating costs over total operating income (%)	-
Industry-specific		
Market concentration	Herfindahl–Hirschman Index (total assets used)	+
Macroeconomic		
Real GDP growth	GDP growth rate in % (year-over-year)	+
Inflation	Consumer Price Index variation in % (year-over-year)	+

Source: own elaboration

After showcasing an overview of all variables and their respective definitions, it is worth mentioning some key takeaways from the summary statistics displayed in the table below. On average, each bank attains a ROAE of 2.38% during the sample period. For comparison purposes, the cost of equity – compensation that market participants demand for investing in and holding banks' equity – at the end of the first semester of 2023 was 13.2% for European banks according to the ECB, whereas the ROAE for our sample was 12% for the same period. The standard deviation for this profitability ratio is 15.2%, thus translating into considerable heterogeneity among financial institutions and between periods. As expected, bank size reveals a low standard deviation, followed by the capital ratio. Regarding the HHI, the sample suggests moderate concentration throughout the chosen timespan[19]. Both the maximum values of GDP growth and inflation were recorded during 2022. Furthermore, even though the normality of the variables is not a strict requirement for most cases, the Shapiro-Wilk test was conducted as a complementary step (STATA command: swilk). The null hypothesis is that the data are not normally distributed. Furthermore, except for *Capital Ratio*, all variables safely reject the null and thus do not follow a normal distribution. The frequently chosen Jarque-Bera test – based on skewness and kurtosis of the sample data – was not conducted in the present Dissertation as it best applies to large samples, whereas the test employed is more effective for small to medium sample sizes.

Table 2 - Descriptive Statistics

Dependent variables	Mean	Median	Std. Dev	Min	Max	Shapiro-Wilk
ROAE	2.384	5.055	15.199	-62.775	32.780	0.000
ROAA	0.185	0.415	1.203	-4.669	2.320	0.000
PBT	0.364	0.693	1.267	-5.021	3.173	0.000
Independent variables						
Bank-specific						
Capital Ratio (CAP)	8.159	8.113	1.286	4.151	10.802	0.079
Size	10.691	10.707	0.233	10.249	11.028	0.000
Loan-to-Deposit ratio (LDR)	92.271	89.725	14.347	60.470	123.324	0.019
Loan Loss Provisions (LLP)	5.521	4.406	3.675	1.747	18.070	0.000
Cost-to-income ratio (COST)	57.404	51.010	30.997	22.387	238.985	0.000
Industry-specific						
Market concentration (HHI)	1444.457	1397.079	106.653	1293.115	1645.952	0.000
Macroeconomic						
Real GDP growth (GDP)	2.162	2.692	4.235	-10.201	9.725	0.000
Inflation (INF)	1.914	0.796	2.571	-0.089	9.495	0.000

No. of observations for each variable: 108
Source: Own elaboration based on STATA 18.5 results

[19] These concentration figures were derived using the market share of each of the six banks. If the market shares of the remaining banks were also taken into account, then the HHI would be slightly higher.

In addition, Pearson's Correlation matrix was built for the independent variables to assess any multicollinearity issues. As one can infer from the table below, all coefficients are below absolute 0.6, with the highest being 0.550 related to *loan-to-deposit ratio* and *market concentration*. Many authors consider a coefficient of above absolute 0.8 as an indication of serious collinearity concerns. That said, there is no need to remove any of the previously selected independent variables due to multicollinearity.

Table 3 - Correlation matrix for the explanatory variables

	CAP	SIZE	LDR	LLP	COST	HHI	GDP	INF
CAP	1							
SIZE	0.036	1						
LDR	-0.292*	-0.433*	1					
LLP	0.165	0.017	0.198*	1				
COST	-0.051	-0.235*	0.143	0.247*	1			
HHI	-0.180	-0.019	0.550*	0.490*	0.247*	1		
GDP	-0.035	0.025	-0.036	-0.018	-0.070	0.004	1	
INF	-0.022	0.038	-0.318*	-0.291*	-0.234*	-0.453*	0.419*	1

Note: * indicates 5% significance level

Source: Own elaboration based on STATA 18.5 results

{Page intentionally left blank}

5. Empirical Findings and Discussion

The present section describes the main empirical results of the proposed econometric methodology following the steps enumerated in subsection 4.1. These empirical findings are then critically reviewed and compared against other articles' findings.

5.1 Stationarity

We start by establishing whether both dependent and independent variables are stationary to determine whether profitability and corresponding explanatory variables' mean and variance are constant over time (i.e., time-independent). There is a multitude of unit root tests to assess stationarity. Despite the fact that most seminal works employ a single test, the present Dissertation resorts to three first-generation nonstationary tests for enhanced robustness, all of which take into consideration the autoregressive structure of the time series and assume cross-sectional independence across units (Barbieri, 2006). Stationary variables allow for the identification of trends embedded in the time series, ultimately underpinning more robust estimated coefficients. Each test was conducted using the time trend option. The first unit root test employed is the Fisher-type Augmented Dickey-Fuller (Dickey and Fuller, 1981) whose null hypothesis implies that all any given time series is nonstationary, whereas the alternative is that at least one time series is stationary (STATA command: xtunitroot fisher). This is a broadly used test as it doesn´t require for the dataset to be balanced (Maddala e Wu, 1999). The second is the Levin-Lin-Chu test (Levin, Lin and Chu, 2002) whose null hypothesis considers identical autoregressive parameters across the panel (STATA command: xtunitroot llc). The third is the Im-Pesaran-Shin test (Im, Pesaran and Shin, 2003) which relaxes the assumption of identical first-order correlation under the alternative, thus allowing for simultaneous stationary and nonstationary series (STATA command: xtunitroot ips). That said, we have chosen tests with different degrees of heterogeneity considered under the alternative hypothesis. Table 4 presents the results of the tests mentioned above.

Table 4 - Unit root tests, at level (including time trend)

	ROAE	ROAA	PBT	CAP	SIZE	LDR	LLP	COST	HHI	GDP	INF
Fisher-ADF	0.016**	0.090*	0.143	0.931	0.906	0.928	0.617	0.033**	0.171	0.047**	0.000***
Levin-Lin-Chu	0.153	0.293	0.028**	0.514	0.052*	0.213	0.108	0.042**	0.318	0.000***	0.000***
Im-Pesaran-Shin	0.220	0.358	0.354	0.889	0.738	0.864	0.662	0.278	0.103	0.034**	0.000***

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

Source: Own elaboration based on STATA 18.5 results

Real GDP growth and *Inflation* are the only variables that reject all null hypotheses considering a significance level of 5% or lower. As such, we repeat each test for the whole panel by resorting to the first difference technique (STATA commands: d.dependent_variable; d.independent_variable) which can be summarized as follows:

$$\Delta y_{it} = y_{it} - y_{it-1}\ ,\ i=1,\ldots,6;\ t=2,\ldots,18 \quad (5)$$

with i and t denoting the usual representation of the number of banks and periods, respectively. Table 5 exhibits the results for the first difference technique.

Table 5 - Unit root tests with first difference (including time trend)

	ΔROAE	ΔROAA	ΔPBT	ΔCAP	ΔSIZE	ΔLDR	ΔLLP	ΔCOST	ΔHHI	ΔGDP	ΔINF
Fisher-ADF	0.000***	0.000***	0.000***	0.018**	0.069*	0.029**	0.000***	0.000***	0.000***	0.002***	0.000***
Levin-Lin-Chu	0.003***	0.005***	0.001***	0.002***	0.032**	0.049**	0.000***	0.001***	0.000***	0.000***	0.002***
Im-Pesaran-Shin	0.000***	0.000***	0.001***	0.043**	0.163	0.126	0.000***	0.000***	0.000***	0.004***	0.000***

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

Source: Own elaboration based on STATA 18.5 results

As one can infer from Table 5, all variables reject the hypothesis of the existence of unit roots for at least two tests with a significance level of 5% or lower, thus rendering the dataset into a stationary panel dataset used in the context of the present empirical research. Although the first-order difference diminishes the degrees of freedom by one, it ensures that the selected variables are stationary. This in turn avoids spurious regressions and contributes to consistent and reliable estimates (Engle and Granger, 1987). The three regression models are thus described hereunder:

Model 1:

$$\Delta ROAE = c + \beta_1 \Delta CAP_{it} + \beta_2 \Delta SIZE_{it} + \beta_3 \Delta LDR_{it} + \beta_4 \Delta LLP_{it} + \beta_5 \Delta COST_{it} + \beta_6 \Delta HHI_t + \beta_7 \Delta GDP_t + \beta_8 \Delta INF_t + \varepsilon_{i,t} \quad (6)$$

Model 2:

$$\Delta ROAA = c + \beta_1 \Delta CAP_{it} + \beta_2 \Delta SIZE_{it} + \beta_3 \Delta LDR_{it} + \beta_4 \Delta LLP_{it} + \beta_5 \Delta COST_{it} + \beta_6 \Delta HHI_t + \beta_7 \Delta GDP_t + \beta_8 \Delta INF_t + \varepsilon_{i,t} \quad (7)$$

Model 3:

$$\Delta PBT = c + \beta_1 \Delta CAP_{it} + \beta_2 \Delta SIZE_{it} + \beta_3 \Delta LDR_{it} + \beta_4 \Delta LLP_{it} + \beta_5 \Delta COST_{it} + \beta_6 \Delta HHI_t + \beta_7 \Delta GDP_t + \beta_8 \Delta INF_t + \varepsilon_{i,t} \quad (8)$$

with i=1,…,6; t=2,…,18 as previously mentioned.

5.2 The Hausman specification test

The second step described in subsection 4.1 is to assess which regression model variant best suits the dataset. The Pooled OLS (POLS) approach is deemed inappropriate as it ignores the panel structure of the data (Schmidheiny and Basel, 2011), thus not considering how observations are grouped into units (Clark and Linzer, 2015). Each bank is an independent unit with autonomous management oversight, therefore the POLS technique – which treats panel data as if it were a simple cross-sectional dataset – is excluded. Hence, the Breusch-Pagan Lagrange multiplier and the *F-test* were not performed[20]. The Hausman specification test - originally described by Hausman (1978) - is thus applied to choose between the Random Effects (RE) variant model and the Fixed Effects (FE) variant model. These two models yield distinct estimation results for the same dataset. This is particularly true when the number of individuals is large, and the number of periods is small (Baltagi, 1998). The null hypothesis of the Hausman test is that the preferred model is the Random Effects model, meaning there is no correlation between the individual errors and the independent variables. Rejecting the null hypothesis translates into a more efficient Fixed Effects estimator. Firstly, it is necessary to estimate all three regressions (one for each of the profitability measures addressed within the present Dissertation); this is accomplished using both models (STATA command: xtreg) to run the Hausman test (STATA command: hausman). The results are summarized in the table below.

[20] The former to compare Pooled OLS against Random Effects and the latter to help choose between POLS and Fixed Effects.

Table 6 - Results from the Hausman specification test

Dependent variable	Chi-Sq. Statistic	p-value
ΔROAE	0.83	0.999
ΔROAA	0.89	0.999
ΔPBT	0.18	1.000

Source: Own elaboration based on STATA 18.5 results

As can be discerned from the preceding specification test, all statistics failed to reject the null for the chosen regression models (Appendix A, B, and C). That said, the estimator employed herein is the Random Effects.

5.3 The Random Effects Regression Results

After conducting the Hausman test and concluding that the Random Effects estimator is more efficient than the Fixed Effects model, the third step consists of applying the former to the three regression model specifications described earlier.

Table 7 showcases the estimated coefficients using Panel EGLS, the corresponding p-values as well as the goodness-of-fit of each regression.

Table 7 - Random Effects baseline results

Variables	ΔROAE	ΔROAA	ΔPBT
ΔCapital Ratio	9.490*** (3.152)	0.591*** (0.139)	0.773*** (0.257)
ΔSize	264.488*** (60.097)	17.879*** (4.083)	16.892*** (4.851)
ΔLoan-to-Deposit ratio	0.073 (0.371)	0.022 (0.029)	-0.029* (0.016)
ΔLoan Loss Provisions	-2.116*** (0.753)	-0.126** (0.056)	-0.268*** (0.057)
ΔCost-to-income ratio	-0.141*** (0.025)	-0.011*** (0.002)	-0.009*** (0.002)
ΔMarket concentration (HHI)	-0.086* (0.050)	-0.007* (0.004)	-0.003 (0.004)
ΔReal GDP growth	0.409 (0.402)	0.037 (0.030)	0.038 (0.029)
ΔInflation	0.206 (0.470)	-0.002 (0.031)	-0.001 (0.031)
Constant	-1.050 (1.164)	-0.025 (0.078)	-0.121 (0.075)
R-Squared (overall)	0.464	0.412	0.480
No. of banks (clusters)	6	6	6
No. of observations	102	102	102

*** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors in parentheses

Source: Own elaboration based on STATA 18.5 results

According to the regression output (Appendix D, E, and F), the first difference of the chosen independent variables explains 46.4% of the overall variation verified in ΔROAE throughout the sample period, and 41.2% of the variation registered in ΔROAA is explained by the regressors. In turn, 48% of the variation observed in ΔPBT is due to a variation in the predictors. The p-value associated with the *F-statistic* is equal to zero for all regressions, meaning that all coefficients in each model are jointly different from zero. Regarding the estimated coefficients' meaning, their interpretation slightly differs from those deriving from the usual lin-lin and log-linear functional forms. Accordingly, in the case of variable *Capital ratio,* a (unitary) percentage point increase in the variation of the predictor will lead to a percentage increase in the delta of each dependent variable equal to the corresponding estimated coefficient, *ceteris paribus*. Regarding *Loan Loss Provisions* and *Cost-to-income ratio*, a percentage point increase in the variation of each explanatory variable translates into a percentage decrease in the variation of ROAE, ROAE, and PBT equal to the corresponding estimated coefficient. Further, since we have initially applied the logarithm of total bank assets to compute *Size*, the meaning of the estimated coefficient is dissimilar, that is a 1% variation in asset value growth results in an increment in the left-hand side of each equation equal to the respective estimated coefficient[21]. As Table 7 demonstrates, all coefficient signs with statistical relevancy (p-values lower than 0.05) are the same across all models for each predictor (with distinct coefficient magnitudes as expected), indicating the overall stability of the empirical findings, as represented by the three equations.

Capital ratio presents a statistically significant positive effect on all dependent variables, wherein the estimated coefficients of the predictor showcase a p-value lower than 0.01. These results come as no surprise since the existing empirical evidence strongly suggests that well-capitalized banks face lower funding costs (Abreu and Mendes, 2001; Kosmidou et al., 2005; Athanasoglou et al., 2006; García-Herrero et al., 2009; Petria et al., 2015). Moreover, a robust capitalization is usually tied to higher solvency ratios (e.g., Common equity Tier I) above the absolute minimum requirements typically imposed by the Basel III protocol, which in turn diminishes banks' dependency on external funding, increases the capacity of exogenous shock absorption, signals an expectation of better performance in the foreseeable future, and allows for more rating upgrades provided by Credit Rating Agencies. In light of the above, the baseline results suggest an absence of

[21] The brief numerical explanation of the coefficients' meaning strictly applies to statistically significant independent variables.

overly cautious risk management frameworks that could potentially constrain higher profitability levels. Furthermore, the present empirical research's findings do not support the argument that high-capitalized banks may indicate overly conservative risk management policies that lead to low profitability, as indicated by Goddard et al. (2004). Moreover, the hypothesis presented by Berger (1995) - in which a high capital ratio leads to lower returns as investors demand lower returns given the perceived diminished risk on equity held – also does not hold for the chosen dataset.

Similarly to the previous regressor, *Size* also exhibits a positive estimated coefficient with relevant statistical significance, thus implying a relatively strong size-efficiency relationship, the existence of economies of scale for the Portuguese banking sector (as in Bourke, 1989; Molyneux and Thornton, 1992; Akhavein et al., 1997; Bikker and Hu, 2002; Athanasoglou et al., 2006; Anbar and Alper, 2011) as well as the benefits associated to product and geographic diversification (scope economies), thus providing access to markets that a small bank cannot entry (see Hughes et al., 1999; Goddard et al., 2001; Alexiou and Sofoklis, 2009). Another size-related advantage is that larger banks are able to exert market power through a stronger brand image, thus generating non-competitive profits as suggested by the Relative-Market-Power hypothesis (see Berger, 1995; Goddard et al., 2004). Furthermore, large institutions can secure financing for their operations at a lower cost than their smaller counterparts (Alexiou and Sofoklis, 2009). Additionally, the results seem to ascertain the premise set forth by Athanasoglou et al. (2006): smaller banks emphasize increasing their market share, rather than improving profitability. Given that the adopted representative dataset comprises the largest banks operating in Portugal, there was an initial concern about not having enough difference in terms of total assets between the six banking groups encompassed by the sample, as all of these banks are of significant size. The baseline results suggest otherwise, meaning there is sufficient heterogeneity among the selected banks in this specific regard. Notwithstanding, including smaller banks in the sample could perhaps yield a conclusion in a different direction in terms of statistical significance, as the benefit of increasing size may hold up to a certain threshold, beyond which the effect of size could become negative due to excess bureaucracy, inertia, and rigidities (Eichengreen and Gibson, 2001). However, including more cross-sections might also bring about further distortions related to bank competition issues (or lack thereof).

The *Loan-to-Deposit* ratio on the other hand demonstrates a somewhat significant effect in model 3 only. The predictor doesn´t significantly impact the commonly used

ROAE and ROAA, thus entailing an innovation for the Portuguese banking sector. However, the negative coefficient achieved in model 3 is in line with Molyneaux and Thornton (1992), Alexiou and Sofoklis (2009), Petria et al. (2015), and Mota et al. (2019), thus contrasting with the works of Bashir (2003), Sufian and Habibullah (2009), and Samad (2015). Intuitively, we expected positive and significant estimated coefficients for the transformation ratio across all models (see Bourke, 1989), although at the expense of higher liquidity risk since a sufficiently high ratio adverts to the bank's deficient liquidity to cover any unforeseen fund requirements (Islam and Rana, 2017). Further, as pointed out by Menicucci and Paolucci (2016), loans granted are a measure of asset efficiency and offer higher expected returns in comparison with other asset classes. This observation is especially relevant for the sampled Portuguese banks, which are essentially commercial banks with significant credit lending portfolios. Most loans are funded by deposits (and, to a lesser extent, interbank borrowings), and the more deposits the more loan opportunities a commercial bank seeks to incur, which in turn increases the corresponding profit levels (Lee and Hsieh, 2013). As shown in section 2, the loan-to-deposit ratio has entered a downward phase in the wake of the GFC, which alongside soaring profits since 2018 may help explain the negative coefficient attained in model 3. In fact, Portuguese banks have been displaying excess liquidity throughout recent years and thus have been trying to increase the transformation ratio as deposits constitute a source of liabilities which diminishes net interest margin when these deposits are not adequately transformed into strongly performing assets (Kashyap et al., 2002).

With respect to *Loan Loss Provisions*, which is the chosen proxy for credit risk and embedded asset quality, the expected negative coefficient sign is highly significant for all baseline models. This relationship has also been confirmed by Athanasoglou et al. (2006), Alexiou and Sofoklis (2009), Dietrich and Wanzenried (2011), Trujillo-Ponce (2013), Petria et al. (2015), Menicucci and Paolucci (2016), and Borges and Tavares (2020), although most research articles refer to either ROA or ROE. Hence, the present Dissertation provides the existing academic literature with another insight by resorting to ΔPBT in this particular regard. Moreover, riskier loans may generate higher interest income while the debtor is performing (Iannotta et al., 2007; Adelopo et al., 2018) but somewhat frequently these loans start to default quicker than other counterparts with lower credit risk which translates into higher provisioning. The latter effect seems to outweigh the former according to the academic literature and the present empirical research. This observation highlights the increasing importance of high collateralization

levels to minimize provisions once a given debtor and the corresponding loan becomes non-performing.

In turn, the *Cost-to-income ratio* displays a highly statistically significant negative relationship with the dependent variables across all regression models. Such an inverse link was also derived by Alexiou and Sofoklis (2009), García-Herrero et al. (2009), Petria et al. (2015), among others. This is the most widely used ratio to assess and monitor bank efficiency and is frequently referred to by practitioners when addressing similar research questions. Moreover, this independent variable exhibits the most consensual (and statistically significant) effect on bank profitability within the existing literature. We have computed this metric as opposed to simply retrieving it from each bank's report as some institutions present their efficiency ratios by offsetting extraordinary or one-off events that would otherwise hinder this variable. That is, our approach has the innovative added-value of addressing the adequate figures.

Market concentration (in terms of total assets) proxied by the Herfindahl-Hirschman Index presents statistically weak negative coefficients for models 1 and 2 which differs from our initial expectation depicted in subsection 4.2.2., hence the interesting implied positive effect that competition has on performance. This may be because market concentration has been decreasing during the past years while overlapping with escalated bank profitability, in the specific case of the Portuguese banking industry. The finding is consistent with that of Clarke et al. (2000), Kosmidou (2008), Petria et al. (2015), and Borroni et al. (2016). It is worth noting that Clarke et al. (2000) have put forward the premise that the entry of larger foreign banks into the domestic market (e.g. Santander, Bankinter, Abanca, and Caixabank regarding the Portuguese case) may force domestic banks to invest more in enhancing their innovative services, thus increasing their operating costs.

As for the macroeconomic variables, both *Real GDP growth* and *Inflation* are not statistically significant in any of the three proposed regressions as in Naceur (2003), Alexiou and Sofoklis (2009), and Anbar and Alper (2011) in the cases of Tunisia, Greece, and Turkey, respectively. Both time series display some outliners throughout the sample period, especially during the novel coronavirus pandemic as portrayed in section 2, which may impede an efficient regression against the dependent variables. Regarding inflation, one reason that has been put forward for the lack of statistical power is that senior management cannot efficiently foresee future inflation, despite the typical stability in the monetary policy of the European Union. On the other hand, the regression results also

imply that institutions are not sluggish in adjusting interest rates during inflationary environments. Failing to do so, coupled with the possibility of costs rising faster than revenues, could adversely affect bank profitability. Moreover, and taking into consideration the adopted time frame (2014-2023), inflation has only become a more recent problem in the aftermath of the pandemic episode and, therefore, has not been a troubling issue for the Portuguese banking industry throughout the adopted time frame as a whole. Concerning real GDP growth – used as a factor to control for general economic conditions – it is found not to be statistically relevant, even though the sample encompasses both dovish and hawkish monetary stances. Furthermore, economic growth does not capture key aspects of banking regulations and technology advances in the financial industry omitted from the dataset (as pinpointed by Naceur, 2003).

5.4 Robustness Checks

In addition to the baseline Random Effects model discussed earlier, the present Dissertation also conducts three (3) robustness checks. The first is related to the estimation of the Fixed Effects, which is also applied to the three equations following as in Mota et al. (2019). As can be seen in Table 8, the general direction of all estimated coefficients remains throughout the models. As expected, considering the Hausman test, there is a slight decrease in the statistical power of some independent variables, namely ΔCapital Ratio for Models 1 and 3, ΔSize for Models 2 and 3, ΔLoan-to-Deposit ratio for Model 3, ΔLoan Loss Provisions for Models 1 and 2, and ΔMarket concentration ratio for Models 1 and 2. Similarly to the Random Effects model, the remaining independent variables are deemed insignificant.

Table 8 - Fixed Effects regression results

Variables	ΔROAE	ΔROAA	ΔPBT
ΔCapital Ratio	9.606**	0.598***	0.779**
	(3.199)	(0.142)	(0.260)
ΔSize	280.764***	19.230**	17.384**
	(68.230)	(5.042)	(5.239)
ΔLoan-to-Deposit ratio	0.106	0.025	-0.029
	(0.412)	(0.032)	(0.017)
ΔLoan Loss Provisions	-1.755*	-0.097	-0.257***
	(0.836)	(0.068)	(0.060)
ΔCost-to-income ratio	-0.136***	-0.011***	-0.009***
	(0.023)	(0.001)	(0.002)
ΔMarket concentration (HHI)	-0.092	-0.007	-0.004
	(0.053)	(0.004)	(0.004)
ΔReal GDP growth	0.412	0.038	0.038
	(0.400)	(0.030)	(0.029)
ΔInflation	0.208	-0.002	-0.001
	(0.471)	(0.031)	(0.031)
Constant	-0.922	-0.013	-0.119*
	(0.939)	(0.057)	(0.059)
R-Squared (overall)	0.463	0.411	0.480
No. of banks (clusters)	6	6	6
No. of observations	102	102	102

*** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors in parentheses

Source: Own elaboration based on STATA 18.5 results

Moreover, a second robustness test is also applied as the independent variable ΔCapital Ratio is replaced by the variable ΔSolvency Ratio which is computed by dividing total Equity by Total Liabilities and was found to be stationary at the first difference. Practitioners frequently overlook this newly employed variable and thus the present Dissertation further contributes to the existing empirical literature. The Random Effects estimator is applied to all models (Table 9) with ΔSolvency Ratio exhibiting robust statistical power with the corresponding p-values lower than 0.01 across all models. Only one of these variables can be used at the time because including both in a given model would raise multicollinearity issues as they are highly correlated.

Table 9 - Random Effects regression results with ΔSolvency Ratio

Variables	ΔROAE	ΔROAA	ΔPBT
ΔSolvency Ratio	7.981***	0.500***	0.648***
	(2.771)	(0.124)	(0.227)
ΔSize	265.868***	18.008***	16.983***
	(62.612)	(4.183)	(5.031)
ΔLoan-to-Deposit ratio	0.070	0.021	-0.030*
	(0.376)	(0.029)	(0.017)
ΔLoan Loss Provisions	-2.154***	-0.129**	-0.271***
	(0.786)	(0.057)	(0.061)
ΔCost-to-income ratio	-0.141***	-0.011***	-0.009***
	(0.026)	(0.002)	(0.002)
ΔMarket concentration (HHI)	-0.085*	-0.007*	-0.003
	(0.051)	(0.004)	(0.004)
ΔReal GDP growth	0.401	0.037	0.037
	(0.406)	(0.030)	(0.029)
ΔInflation	0.204	-0.002	-0.001
	(0.468)	(0.031)	(0.031)
Constant	-1.063	-0.026	-0.122
	(1.196)	(0.079)	(0.079)
R-Squared (overall)	0.456	0.409	0.469
No. of banks (clusters)	6	6	6
No. of observations	102	102	108

*** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors in parentheses

Source: Own elaboration based on STATA 18.5 results

Lastly, a third robustness test is also conducted by estimating a Random Effects model applied to the period ranging from 2015 to 2019, thus exclusively encompassing the pre-pandemic period only. The goal is to evaluate whether the estimated coefficients' sign holds throughout the business cycle's upward trajectory, which presents fewer outliers than the baseline scenario. As shown in Table 10, the coefficient signs for the variables with an initial p-value lower than 0.1 remain the same for the new models, with the key differences of ΔLoan Loss Provisions not being significant for models 1 and 2, and ΔCost-to-income not being significant for model 3.

Table 10 - Random Effects regression results for 2015-2019 period

Variables	ΔROAE	ΔROAA	ΔPBT
ΔCapital Ratio	11.656***	0.702***	0.917***
	(3.750)	(0.169)	(0.310)
ΔSize	276.433***	18.389***	19.259***
	(71.862)	(4.094)	(7.010)
ΔLoan-to-Deposit ratio	0.136	0.029	-0.039*
	(0.420)	(0.037)	(0.020)
ΔLoan Loss Provisions	-1.508	-0.080	-0.231***
	(0.968)	(0.096)	(0.072)
ΔCost-to-income ratio	-0.104***	-0.009***	-0.006
	(0.039)	(0.002)	(0.004)
ΔMarket concentration (HHI)	-0.058	-0.003	-0.007
	(0.067)	(0.004)	(0.007)
ΔReal GDP growth	2.619	0.309	-0.136
	(4.854)	(0.367)	(0.235)
ΔInflation	-8.460*	-0.885*	0.245
	(4.696)	(0.525)	(0.309)
Constant	-1.871	-0.075	-0.216**
	(1.560)	(0.114)	(0.104)
R-Squared (overall)	0.565	0.497	0.469
No. of banks (clusters)	6	6	6
No. of observations	102	102	108

*** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors in parentheses

Source: Own elaboration based on STATA 18.5 results

All in all, as indicated by both the baseline model and the three robustness checks herein conducted, the results for the strongly significant explanatory variables are generally consistent with the existing empirical literature and provide an innovative contribution to the contemporary debate on bank profitability, thus bridging the gap between the existing studies and the case of the Portuguese banking system, especially in view of the fact that the research on the latter is rather scant.

To maintain sufficient levels of profitability, financial institutions should continue focusing, *inter alia*, on growing their business either via organic strategies (e.g., expanding the range of products and services provided to existing customers, and acquiring new customers by offering more competitive price options) or inorganic opportunities (i.e., consolidation). Increasing size must be accomplished with enough prudence so that loan origination does not translate into losses for the banks. This is specially the case when business cycle conditions rapidly deteriorate, leading assets to become non-performing on account of a potentially severe macroeconomic downturn.

Accordingly, measures that reduce the stock of NPL (i.e., restructuring and cure, collateral foreclosure, write-offs, loan sales, and assignment of the contractual position to Corporate Restructuring Funds) and that foster operating efficiency should be highly prioritized. Although the so-called transformation ratio didn´t display a strong significant effect on profitability, banks should still strike a careful balance between this metric and the inherent liquidity risk. Dividend distribution should not hinder adequate levels of capitalization as the latter is strongly correlated with profit generation. Indeed, since all banks composing the dataset are classified by the EBA as Other Systemically Important Institutions (O-SIIs) given their systemic importance at the national level, they are subject to tighter regulation and additional buffer requirements to account for the externalities that they could exert on the domestic economy in the case of failure or distress (Andries et al., 2020). Such regulation allowed banks to face the pandemic with greater resilience to shock given the reforms implemented after the GFC (Berger et al., 2021).

{Page intentionally left blank}

6. Conclusion

The present Dissertation aims to identify and critically analyze a proposed set of key determinants that influence the profitability of Portuguese banks. Through the academic literature review of the existing research articles focusing on the topic addressed herein, it was possible to pinpoint the main internal and external factors tied to bank performance.

Concerning profitability indicators, the three dependent variables of choice are i) Return on Average Equity (ROAE), ii) Return on Average Assets (ROAA), and less commonly used iii) Profit Before Tax (PBT). Hence, three distinctive models are built.

In what concerns key determinants of bank profitability, the bank-specific characteristics (internal factors) employed in this Dissertation are i) Capital Ratio; ii) Size, iii) Loan-to-Deposit ratio, iv) Loan Loss Provisions, and v) Cost-to-income ratio. Additionally, three external factors are used, specifically i) Market Concentration, proxied by the HHI (industry-specific characteristic), ii) Real GDP growth, and iii) Inflation (macroeconomic characteristics).

The dataset comprises the six largest Portuguese commercial banks, namely Caixa Geral de Depósitos, Millennium BCP, Banco Santander Totta, Novo Banco, Banco Português de Investimento, and Banco Montepio, for the period 2015-2023. A total of 108 half-yearly observations were derived exclusively from publicly available information from these institutions' official financial accounts, thus establishing a balanced panel dataset that encompasses the upward and the ensuing downward phases of the business cycle related to the pandemic.

The Pearson correlation matrix is first applied to all regressors. After arguing for the absence of multicollinearity issues among the selected determinants, the whole panel is subject to the first difference technique to achieve stationarity and avoid spurious regressions as not all variables are stationary at level. Further, the Hausman specification test is conducted for all three models, allowing us to conclude that the Random Effects model is more efficient than the Fixed Effects model. The Panel EGLS estimator is then implemented across the regression models using cluster-robust standard errors to control for collinearity, heteroscedasticity, and autocorrelation. Moreover, three robustness checks are carried out to account for i) the Fixed Effects regression model as an alternative to the RE model, ii) a different capital adequacy predictor, and iii) a different sample period covering the pre-pandemic phase.

To a great extent, the findings related to significant bank-specific determinants are overall in line with those described in seminal articles and other empirical studies, whereas control variables do exhibit weak to no explanatory power. The baseline results advocate for four highly significant internal factors across all regression models with corresponding p-values lower than 0.05. The first is the *Capital Ratio*, wherein the higher the percentage of assets being financed through equity the higher the bank's profitability, suggesting that well-capitalized institutions face lower funding costs and reduce their dependency on external funding. The second is bank *Size* captured by the common logarithm of total consolidated assets, for which there is an indication of economies of scale and scope economies within the Portuguese banking industry. Thirdly, the *Loan Loss Provisions* highlight the tradeoff between riskier loans and profitability. The fourth is the *Cost-to-income ratio* which displays the most consensual effect on profitability among practitioners. As expected, operating efficiency has a robust and negative impact on net income and profit before tax. On the other hand, the *Loan-to-Deposit ratio* and the external factors do not demonstrate any profound and significant repercussions on profitability, similar to other research articles.

As in virtually all research, the present Dissertation has its limitations. First and foremost, bank-specific data collection was restricted to each institution's financial reports instead of resorting to a proprietary database, thus restricting the research to publicly available information and obliging to a time-consuming data collection procedure. This provides, however, enhanced tractability and enables any replication studies to be conducted. Secondly, the number of observations is rather small compared to the typical single-country studies, which is a natural reflection of choosing the Portuguese banking industry as the center of the research question. Nevertheless, the present critical analysis engages a panel data model and resorts to half-yearly data to circumvent the previous concern and offer an innovative approach to studying the Portuguese case. Lastly, banks may proceed to income smoothing as part of an earnings management strategy, particularly in the second half of the year, which may distort the true capacity of generating profits.

Regarding the desideratum for future research, it is worth investigating any persistence of profit (POP) among Portuguese banks (following Berger et al., 2000; Goddard et al., 2004; Athanasoglou et al., 2008) as well as extending the sample period and including additional banks in the dataset. Further, it may be interesting to depart from

the empirical literature by including the Consumer confidence index as an explanatory variable.

Banks are the centerpiece of a bank-based economy such as Portugal. It is of paramount importance to achieve a sound and profitable Portuguese banking sector, indelibly marked by increasingly demanding requirements introduced by the third Basel Accord. That said, both academics and senior bank Managers may find in this and other research articles alike an opportunity to revisit their strategic plans and set goals to leverage key aspects that foster efficiency and sustainably tackle the determinants responsible for a subdued performance.

{Page intentionally left blank}

References

Abreu, M., & Mendes, V. (2001). Commercial bank interest margins and profitability: evidence for some EU countries. In *Pan-European Conference Jointly Organised by the IEFS-UK & University of Macedonia Economic & Social Sciences, Thessaloniki, Greece, May* (Vol. 34, No. 2, pp. 17-20).

Adelopo, I., Lloydking, R. & Tauringana, V. (2018). Determinants of bank profitability before, during, and after the financial crisis. *International Journal of Managerial Finance, 14*(4), 378-398.

Akhavein, J., Berger, A. and D. Humphrey (1997). "The effects of megamergers on efficiency and prices: Evidence from a bank profit function." *Finance and Economic Discussion Series No. 9*, Board of Governors of the Federal Reserve System, Washington, D.C.

Alexiou, C. e Sofoklis, V. (2009) «Determinants of bank profitability: Evidence from the Greek banking sector», *Economic Annals*, 54(182), pp. 93–118.

Anbar, A., & Alper, D. (2011). Bank specific and macroeconomic determinants of commercial bank profitability: Empirical evidence from Turkey. Business and Economics Research Journal, 2(2), 139.

Andries, A. M., Nistor, S., Ongena, S., & Sprincean, N. (2020). On becoming an O-SII ("Other Systemically Important Institution"). *Journal of Banking & Finance, 111*, 105723.

Associação Portuguesa de Bancos (2021) *Boletim Informativo.*

Athanasoglou, P.P., Brissimis, S.N., Delis, M.D. (2005), Bank-Specific, Industry-Specific and Macroeconomic Determinants of Bank Profitability, Bank of Greece, Working Paper No. 25.

Athanasoglou, P. P., Brissimis, S. N., Delis, M. D. (2008). Bank-specific, industry-specific and macroeconomic determinants of bank profitability. *Journal of international financial Markets, Institutions and Money, 18*(2), 121-136.

Athanasoglou, P. P., Daniilidis, I. & Delis, M. D. (2014). Bank procyclicality and output: Issues and policies. *Journal of Economics and Business*, 72, 58-83.

Athanasoglou, P.P., Delis, M.D., Staikouras, C.K. (2006). Determinants of Bank Profitability in the South Eastern European Region, Bank of Greece, Working Paper No. 47.

Augusto, F., and S. Félix. (2014). "The Impact of Bank Recapitalization on Firms' Access to Credit: Evidence from Portugal." *Financial Stability Paper No. 2* (November), Banco de Portugal.

Baltagi B. H. (1998), 'In Hand book of Applied Economics Statistics', New York.

Barbieri, L. (2006), "Panel unit root tests: a review", *Serie Rossa: Economia – Quadernoi*, Vol. 43, Ottobre, pp. 1-53.

Barros, Carlos Pestana, Candida Ferreira, and JonathanWillians. (2007). Analysing the determinants of performance of best and worst European banks: A mixed logit approach. *Journal of Banking and Finance 31*: 2189–203.

Barros, Fátima and Leonor Modesto (1999). "Portuguese banking sector: a mixed oligopoly?" International Journal of Industrial Organization, 17(6): 869-886.

Bashir, A. H. M. (2003). Determinants of profitability in Islamic banks: Some evidence from the Middle East. *Islamic economic studies, 11*(1).

Berger, A. (1995). "The relationship between capital and earnings in banking." *Journal of Money, Credit and Banking* 27, 432-456.

Berger, A.N., Bonime, S.D., Covitz, D.M., Hancock, D. (2000). Why are bank profits so persistent? The roles of product market competition, information opacity and regional macroeconomic shocks. *Journal of Banking & Finance 24*, 1203–1235.

Berger, A., Demirguc-Kunt, A., Levine, R. and J. Haubrich (2004). "Bank concentration and competition: An evolution in the making." *Journal of Money, Credit and Banking* 36, 433-451.

Berger, A. N., Demirgüç-Kunt, A., Moshirian, F., & Saunders, A. (2021). The way forward for banks during the COVID-19 crisis and beyond: Government and central bank responses, threats to the global banking industry. *Journal of Banking and Finance, 133.*

Bhimjee, D. (2013). THREE ESSAYS ON THE GLOBAL FINANCIAL CRISIS OF 2007- 2008. Thesis (PhD), ISCTE-IUL.

Bikker, J. and H. Hu (2002). "Cyclical patterns in profits, provisioning and lending of banks and procyclicality of the new Basel capital requirements." *BNL Quarterly Review 221*, 143-75.

Borges, M.R., and Tavares, A.S. (2020). "Determinants of bank performance in the context of crisis: A panel data analysis for Portugal." *European Research Studies*, 23: 674–686.

Borroni, M., Piva, M., & Rossi, S. (2016). Determinants of Bank Profitability in the Euro Area: Has Anything Changed? (No. dises1619). Università Cattolica del Sacro Cuore, Dipartimenti e Istituti di Scienze Economiche (DISCE).

Boucinha M, Ribeiro N (2009). An assessment of competition in the Portuguese banking system in the 1991–2004 period. Working paper 1/2009, Banco de Portugal

Bourke, P. (1989). Concentration and other determinants of bank profitability in Europe, North America and Australia. *Journal of Banking & Finance*, *13*(1), 65-79.

Canhoto, A. (2004). Portuguese banking: A structural model of competition in the deposits market. *Review of financial economics, 13*(1-2), 41-63.

Cipollini, A., Fiordelisi, F. (2012). Economic value, competition and financial distress in the European banking system. *J. Bank. Finance 36*, 3101–3109.

Clark, T. S., & Linzer, D. A. (2015). Should I use fixed or random effects? *Political Science Research and Methods, 3*, 399 – 408.

Clarke, G., Cull, R., D'Amato, L., Molinari, A. (2000). On the Kindness of Strangers? The Impact of Foreign Entry on Domestic Banks in Argentina. In Claessens, S., Jansen, M. (eds), *The Internationalization of Financial Services: Issues and Lessons for Developing Countries*. Dordrecht and Boston: Kluwer Academic Publishers.

Crosignani, M., Faria-e-Castro, M. and Fonseca, L. (2015), "The Portuguese Banking System During the Sovereign Debt Crisis", *Banco de Portugal Economic Studies*, Vol. 1, No 2, pp. 43-80.

Dickey, D. A. & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica 49*, 1057-72

Dietrich, A., & Wanzenried, G. (2011). Determinants of bank profitability before and during the crisis: Evidence from Switzerland. *Journal of International Financial Markets, Institutions and Money, 21*(3), 307-327.

Eichengreen, B., Gibson, H.D. (2001). Greek banking at the dawn of the new millennium. CERP Discussion Paper 2791, London.

Elzinga, K. G. and D. E. Mills (2011). The lerner index of monopoly power: Origins and uses. *American Economic Review 101* (3), 558–64.

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 251–276.

Ercegovac, R., Klinac, I., & Zdrilić, I. (2020). Bank-specific determinants of EU banks profitability after the 2007 financial crisis. *Journal of Contemporary Management Issues, 25*(1), 89–102. European Banking

García-Herrero, A., Gavilá, S., & Santabárbara, D. (2009). What explains the low profitability of Chinese banks? *Journal of Banking & Finance, 33*(11), 2080-2092.

Garcia, M. T. M., & Guerreiro, J. P. S. M. (2016). Internal and external determinants of banks' profitability: The Portuguese case. *Journal of Economic Studies*, 43(1), 90-107.

Gilbert, R. (1984), ''Bank market structure and competition – a survey'', *Journal of Money Credit and Banking*, Vol. 16, pp. 617-45.

Goddard, J. A., Molyneux, P., & Wilson, J. O. S. (2001). European banking: Efficiency, technology, and growth. John Wiley.

Goddard, J., Molyneux, P., Wilson, J.O.S. (2004). The profitability of European banks: A cross-sectional and dynamic panel analysis. Manchester School 72, 363–381.

Hausman JA (1978). Specification tests in econometrics. *Econometrica*; 46(6):1251.

Hsiao, Cheng (1986). Analysis of Panel Data. Cambridge: Cambridge University Press.

Hughes, J. P., Lang, W. W., Mester, L. J. and Moon, C. G. (1999). 'The Dollars and Sense of Bank Consolidation', *Journal of Banking and Finance*, Vol. 23, No. 2–4, pp. 291–324

Iannotta, G., Nocera, G., Sironi, A. (2007). Ownership structure, risk and performance in the European banking industry. *J. Bank. Finance* 31 (7), 2127–2149.

Im, K.S., Pesaran, M.H., Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of econometrics 115* (1), 53–74.

Islam, M. A., & Rana, R. H. (2017). Determinants of bank profitability for the selected private commercial banks in Bangladesh: a panel data analysis. *Banks and Bank Systems, 12*(3), 1–10.

Kashyap, A. K., Rajan, R., & Stein, J. C. (2002). Banks as liquidity providers: An explanation for the coexistence of lending and deposit-taking. *The Journal of Finance, 57*(1), 33-73.

Kosmidou, K. (2008). The determinants of banks' profits in Greece during the period of EU financial integration. *Managerial Finance*, *34*(3), 146-159.

Kosmidou, K., Tanna, S. e Pasiouras, F. (2005) «Determinants of profitability of domestic UK commercial banks: panel evidence from the period 1995-2002. *In Money Macro and Finance (MMF) Research Group Conference* (VOL. 45, pp 1-27).

Lee, C. C., & Hsieh, M. F. (2013). The impact of bank capital on profitability and risk in Asian banking. *Journal of international money and finance, 32*, 251-281.

Levin, A., Lin, C. F., & Chu, C. S. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of econometrics, 108*(1), 1-24.

Lindenberg, E.B. and S.A. Ross (1981) 'Tobin's q Ratio and Industrial Organization', *Journal of Business, 54*: 1-32.

Maddala, G. and Wu Shaoen (1999), A Comparative Study of Unit Root Tests with panel Data and New Simple Test. *Oxford Bulletin of Economics and Statistics, Special Issue*. 0305-9049.

Marques, L. D. (2000). Modelos Dinâmicos com Dados em Painel: revisão de literatura. Faculdade de Economia da Universidade do Porto.

Mendes, V., & Rebelo, J. (1999). Productive efficiency, technological change and productivity in Portuguese banking. *Applied Financial Economics, 9*(5), 513-521.

Menicucci, E., & Paolucci, G. (2016). The determinants of bank profitability: empirical evidence from European banking sector. *Journal of financial reporting and Accounting.*

Molyneux, P. and J. Thornton (1992). “Determinants of European bank profitability: A note.” *Journal of Banking and Finance* 16, 1173-1178.

Mota, C., Silva, B., & Silva, E. (2019). Determinantes da rentabilidade bancária: Evidências para os maiores bancos portugueses. *European Journal of Applied 51 Business Management, 5*, 78–96.

Naceur, B. S. (2003). “The Determinants of the Tunisian Banking Industry Profitability: Panel Evidence.” Université Libre de Tunis: Department of Finance.

Pasiouras, F., & Kosmidou, K. (2007). Factors influencing the profitability of domestic and foreign commercial banks in the European Union. *Research in International Business and Finance, 21*(2), 222-237.

Perry, P. (1992). Do Banks Gain or Lose from Inflation? *Journal of Retail Banking, 14*(2), 25-35.

Peterson, O. K., & Arun, T. G. (2018). Income smoothing among European systemic and nonsystemic banks. *The British Accounting Review*, 50(5), 539-558.

Petria, N., Capraru, B., & Ihnatov, I. (2015). Determinants of banks' profitability: evidence from EU 27 banking systems. *Procedia economics and finance, 20*, 518-524.

Pires, C., Basílio, M., & Borralho, C. (2021). Determinants of Portuguese banks' profitability: an update. *Tourism & Management Studies*, 17(3), 63-70.

Rhoades, S.A. (1995). Market Share Inequality, the HHI, and Other Measures of the Firm-Composition of a Market. in *Review of Industrial Organization, 10*(6), 657-74.

Ribeiro, N. (2007). "Debt growth: Factors, institutional issues and implications — The Portuguese case". In Enoch, C. and Ötker-Robe, I. (eds.), *"Rapid credit growth in central and eastern Europe: Endless boom or early warning?"* International Monetary Fund.

Samad, A. (2015). Determinants bank profitability: Empirical evidence from Bangladesh commercial banks. *International Journal of Financial Research, 6*(3), 173–179.

Schmidheiny, K., Basel, U. (2011). Panel data: fixed and random effects. Short Guides to Microeconometrics.

Short, B. K. (1979). The relation between commercial bank profit rates and banking concentration in Canada, Western Europe, and Japan. *Journal of banking & Finance*, 3(3), 209-219.

Stadheim, V. (2020). Public banking in pandemic times: Portugal's Caixa Geral de Depósitos. In D. McDonald, T. Marois, and D. Barrowclough (eds*.), Public banks and COVID-19: Combatting the pandemic with public finance* (pp. 311–332). Municipal Services Project, UNCTAD and Eurodad.

Stiroh, K.J., Rumble, A. (2006). The dark side of diversification: The case of US financial holding companies. *J. Bank. Financ*. 30 (8), 2131–2161.

Sufian, F., & Habibullah, M. S. (2009). Bank specific and macroeconomic determinants of bank profitability: Empirical evidence from the China banking sector. *Frontiers of Economics in China, 4*(2), 274-291.

Tan, Y. (2016) «The impacts of risk and competition on bank profitability in China», *Journal of International Financial Markets, Institutions and Money*, 40, pp. 85–110.

Trujillo-Ponce, A. (2013). What determines the profitability of banks? Evidence from Spain. *Accounting & Finance, 53*(2), 561-586

Wooldridge, J.M. (2012). *Introductory Econometrics: A Modern Approach*. South-Western, Cengage Learning, 5th Edition, printed in USA. ISBN:978-1-111-53104-1.

{Page intentionally left blank}

Appendix

A. The Hausman specification test for model 1

```
. hausman ROAE_FEd ROAE_REd
```

	Coefficients (b) ROAE_FEd	(B) ROAE_REd	(b-B) Difference	sqrt(diag(V_b-V_B)) Std. err.
CAPd	9.606335	9.490269	.116066	.3361975
SIZEd	280.7644	264.4877	16.27672	20.40332
LDRd	.1059067	.0733122	.0325945	.0883125
d HHIad GDPd INFd, re	-1.754748	-2.116444	.3616965	.4407386
COSTd	-.1356716	-.1414582	.0057867	.0092645
HHIad	-.0915583	-.0858858	-.0056725	.0106317
GDPd	.4120615	.4089042	.0031574	.0506637
INFd	.2079908	.2055201	.0024707	.1264708

b = Consistent under H0 and Ha; obtained from **xtreg**.
B = Inconsistent under Ha, efficient under H0; obtained from **xtreg**.

Test of H0: Difference in coefficients not systematic

chi2(**8**) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= **0.83**
Prob > chi2 = **0.9991**

B. The Hausman specification test for model 2

```
. hausman ROAA_FEd ROAA_REd
```

	Coefficients (b) ROAA_FEd	(B) ROAA_REd	(b-B) Difference	sqrt(diag(V_b-V_B)) Std. err.
CAPd	.5982114	.590572	.0076394	.0263888
SIZEd	19.22989	17.87929	1.350605	1.599218
LDRd	.0247875	.0216718	.0031157	.006925
LLPd	-.0973102	-.1263385	.0290283	.0345538
COSTd	-.01095	-.0114076	.0004576	.0007268
HHIad	-.007118	-.0066596	-.0004585	.0008343
GDPd	.0376881	.0374225	.0002656	.0039784
INFd	-.0019097	-.0020593	.0001496	.0099311

b = Consistent under H0 and Ha; obtained from **xtreg**.
B = Inconsistent under Ha, efficient under H0; obtained from **xtreg**.

Test of H0: Difference in coefficients not systematic

chi2(**8**) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= **0.89**
Prob > chi2 = **0.9988**

C. The Hausman specification test for model 3

```
. hausman PBT_FEd PBT_REd
```

	Coefficients (b) PBT_FEd	(B) PBT_REd	(b-B) Difference	sqrt(diag(V_b-V_B)) Std. err.
CAPd	.7789719	.7727774	.0061945	.0263453
SIZEd	17.38401	16.89243	.4915728	1.525084
LDRd	-.0294861	-.0292992	-.0001869	.0066998
LLPd	-.2565418	-.2677778	.011236	.0332282
COSTd	-.0085451	-.00875	.000205	.0007148
HHIad	-.0035975	-.00341	-.0001875	.0008255
GDPd	.0380006	.0379163	.0000843	.0040234
INFd	-.001005	-.00106	.000055	.010041

b = Consistent under H0 and Ha; obtained from **xtreg**.
B = Inconsistent under Ha, efficient under H0; obtained from **xtreg**.

Test of H0: Difference in coefficients not systematic

chi2(**8**) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= **0.18**
Prob > chi2 = **1.0000**

D. The Random Effects model baseline results for model 1

```
. xtreg ROAEd CAPd SIZEd LDRd LLPd COSTd HHIad GDPd INFd, re vce(cluster i)

Random-effects GLS regression                   Number of obs     =       102
Group variable: i                               Number of groups  =         6

R-squared:                                      Obs per group:
     Within  = 0.4693                                         min =        17
     Between = 0.0290                                         avg =      17.0
     Overall = 0.4635                                         max =        17

                                                Wald chi2(5)      =         .
corr(u_i, X) = 0 (assumed)                      Prob > chi2       =         .

                                  (Std. err. adjusted for 6 clusters in i)
------------------------------------------------------------------------------
             |               Robust
      ROAEd  | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
        CAPd |   9.490269   3.151859     3.01   0.003     3.312738    15.6678
       SIZEd |   264.4877   60.09661     4.40   0.000     146.7005    382.2749
        LDRd |   .0733122   .3706035     0.20   0.843    -.6530573    .7996817
        LLPd |  -2.116444   .7530171    -2.81   0.005    -3.592331   -.640558
       COSTd |  -.1414582   .0254595    -5.56   0.000    -.1913579   -.0915585
       HHIad |  -.0858858   .0495895    -1.73   0.083    -.1830795    .0113079
        GDPd |   .4089042   .4015596     1.02   0.309    -.3781382    1.195947
        INFd |   .2055201   .4700555     0.44   0.662    -.7157718    1.126812
       _cons |  -1.050272   1.163618    -0.90   0.367    -3.330922    1.230378
-------------+----------------------------------------------------------------
     sigma_u |          0
     sigma_e |    11.3331
         rho |          0   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

E. The Random Effects model baseline results for model 2

```
. xtreg ROAAd CAPd SIZEd LDRd LLPd COSTd HHIad GDPd INFd, re vce(cluster i)

Random-effects GLS regression                   Number of obs     =       102
Group variable: i                               Number of groups  =         6

R-squared:                                      Obs per group:
     Within  = 0.4183                                         min =        17
     Between = 0.0538                                         avg =      17.0
     Overall = 0.4121                                         max =        17

                                                Wald chi2(5)      =         .
corr(u_i, X) = 0 (assumed)                      Prob > chi2       =         .

                                  (Std. err. adjusted for 6 clusters in i)
------------------------------------------------------------------------------
             |               Robust
      ROAAd  | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
        CAPd |    .590572   .1394213     4.24   0.000     .3173113    .8638327
       SIZEd |   17.87929    4.08329     4.38   0.000     9.876184    25.88239
        LDRd |   .0216718   .0288164     0.75   0.452    -.0348073    .0781509
        LLPd |  -.1263385   .0563842    -2.24   0.025    -.2368495   -.0158275
       COSTd |  -.0114076   .0016226    -7.03   0.000    -.0145877   -.0082274
       HHIad |  -.0066596   .0036981    -1.80   0.072    -.0139077    .0005886
        GDPd |   .0374225   .0301673     1.24   0.215    -.0217043    .0965493
        INFd |  -.0020593   .0306833    -0.07   0.946    -.0621974    .0580788
       _cons |  -.0246222   .0782012    -0.31   0.753    -.1778937    .1286492
-------------+----------------------------------------------------------------
     sigma_u |          0
     sigma_e |  .88761481
         rho |          0   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

F. The Random Effects model baseline results for model 3

```
. xtreg PBTd CAPd SIZEd LDRd LLPd COSTd HHIad GDPd INFd, re vce(cluster i)

Random-effects GLS regression                   Number of obs     =       102
Group variable: i                               Number of groups  =         6

R-squared:                                      Obs per group:
     Within  = 0.4820                                         min =        17
     Between = 0.0014                                         avg =      17.0
     Overall = 0.4799                                         max =        17

                                                Wald chi2(5)      =         .
corr(u_i, X) = 0 (assumed)                      Prob > chi2       =         .

                                  (Std. err. adjusted for 6 clusters in i)
------------------------------------------------------------------------------
             |               Robust
        PBTd |      Coef.   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
        CAPd |   .7727774   .2569184     3.01   0.003     .2692265    1.276328
       SIZEd |   16.89243   4.850693     3.48   0.000     7.385251    26.39962
        LDRd |  -.0292992    .016468    -1.78   0.075    -.0615759    .0029774
        LLPd |  -.2677778   .0571188    -4.69   0.000    -.3797285   -.1558271
       COSTd |    -.00875   .0016772    -5.22   0.000    -.0120373   -.0054628
       HHIad |    -.00341   .0036634    -0.93   0.352      -.01059    .0037701
        GDPd |   .0379163   .0286499     1.32   0.186    -.0182365     .094069
        INFd |    -.00106   .0314078    -0.03   0.973    -.0626181     .060498
       _cons |  -.1209885   .0754696    -1.60   0.109    -.2689062    .0269291
-------------+----------------------------------------------------------------
     sigma_u |          0
     sigma_e |  .82442477
         rho |          0   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```