U.PORTO

FEP FACULDADE DE ECONOMIA UNIVERSIDADE DO PORTO

The Impact of Macroeconomic Variables, Bank-Specific, Institutional and Market Level Factors on Non-Performing Loans

An Analysis in the Portuguese Context Pre- and Post-COVID-19

Márcia Patrícia Fernandes da Silva

Dissertation

Master in Finance

Supervised by

José Abílio de Oliveira Matos

2025

Acknowledgments

In this section, I would like to express my sincere gratitude to all the people who, in some way, allowed me and gave me the motivation to cross this academic finish line, despite all the challenges.

First, to my advisor, Professor José Abílio de Oliveira Matos, who, despite all his responsibilities and academic challenges, always found the time and availability to guide me through this academic work and for sharing his knowledge, which facilitated my journey. I also want to emphasise that the professor was always available and committed to balancing his schedule with mine, and for that I am very grateful.

At last, to all my friends and family who stood by me, offering support and advice.

Abstract

In the Portuguese economy, banks are the main institutions in the financial system, and loans remain the primary source of income, despite the growing interest in diversification. NPLs - traditionally classified as loans overdue for more than three months - pose material risks to the stability of the banking system, limiting liquidity and profitability, and increasing systemic risk for banks. This article analyses the determinants of NPLs in the Portuguese banking industry, focusing on the pre- and post-COVID-19 eras. It is organised into six parts, ranging from theory to empirical analysis. Section 1 presents the study, its objectives, and its relevance for the Portuguese banking system and the European Monetary Union. The literature on NPL determinants is reviewed in Section 2, and Section 3 presents the dataset and Section 4 methodology, which includes ARDL, Elastic Net, OLS, ECM, and VAR techniques. Section 5 discusses the empirical results, including detailed COVID-19 dynamics, and Section 6 concludes with a summary of the main findings, policy implications, and scope for future research. The main results show that NPL dynamics in Portugal are fundamentally explained by persistence and bank-specific factors (liquidity and size), rather than macroeconomic variables and COVID-19 shocks, which exert short-term and context-dependent impacts. Historical credit risk highly conditions current NPL levels; this dynamic clearly reflects the structural diversities of Portuguese banks. Policy recommendations call for strong liquidity regulation, capital buffers, and targeted assistance during major shocks.

Keywords: NPL, Portuguese banking sector, macroeconomic determinants, bank-specific factors, COVID-19, ARDL, financial stability.

Resumo

Na economia portuguesa, os bancos ocupam uma posição de destaque no sistema financeiro, tendo principalmente os empréstimos como sua principal fonte de rendimento, embora o interesse em diversificar se torne cada vez mais evidente atualmente. O crédito malparado (NPL) - geralmente definidos como empréstimos cujo principal ou juros estão em atraso há mais de 90 dias - representam uma ameaça significativa à solidez do sistema bancário, reduzem a disponibilidade de crédito, lucratividade e aumentam os riscos sistémicos para os bancos e o sistema financeiro em geral. Este estudo examina os determinantes dos NPLs no setor bancário português, antes e após a pandemia de COVID-19. O estudo está dividido em 6 seções: os objetivos do estudo são delineados e sua relevância para o sistema bancário português e a União Económica e Monetária é discutida na secção 1. A pesquisa de literatura sobre os determinantes dos NPLs é fornecida na seção 2, e a secção 3 descreve as variáveis e as respetivas fontes utilizadas, a secção 4 apresenta a metodologia que envolve ARDL, Elastic Net, OLS, ECM e VAR. A seção 5 exibe os resultados empíricos, incluindo uma análise específica das dinâmicas da COVID-19, enquanto a secção 6 conclui com uma síntese dos principais resultados, limitações encontradas ao longo do estudo e as implicações políticas e futuras abordagens de pesquisa. Os principais resultados revelaram que a persistência e fatores específicos dos bancos (liquidez e tamanho) são os determinantes mais influentes para a dinâmica dos NPLs em Portugal, e não as variáveis macroeconômicas e choques da COVID-19, que têm efeitos de curto prazo e dependentes do contexto. O risco de crédito histórico é, portanto, um fator condicionante muito forte para os níveis atuais de NPL, uma relação que também demonstra claramente a diferenciação estrutural dos bancos portugueses. As prescrições políticas sugerem uma regulamentação sólida da liquidez e das reservas de capital, bem como assistência direcionada durante grandes choques.

Table of Contents

Acknowledgments i

Abstract ii

Resumo iii

List of Figures v

List of Tables vi

List of Acronyms vi

Introduction 1

Literature Review 4

Data Description and Methodology 7

Data Source and processing/preparation 7

Exploratory analysis of the data 10

Comparative Analysis of Pre- and Post-COVID-19 Trends 13

Methodology used in the study 15

Results and Discussion 19

ARDL Estimation Results 20

Conclusions and implications 26

References 28

Appendices 37

Appendix A: Cleaning and Frequency-Harmonised Data 37

Appendix B: Descriptive Analysis of the Dataset 39

B1) Analysis of all 4 variable categories with COVID-19 split 39

B2) Correlation matrix heatmap for all variables 56

B3) Scatter Plot Analysis of NPL Determinants 60

B4) Variables correlated pairs 63

Appendix C: Data Testing and Modelling 67

C1) Test the data 67

C2) Vector Autoregression (VAR) Framework 71

C3) Autoregressive Distributed Lag (ARDL) 74

C4) ARDL Estimation Results – Elastic Net, ARDL-based 75

C5) ARDL Estimation Results – OLS simple regression, ARDL-based 77

C6) ARDL Estimation Results – Dynamic ECM estimation ARDL-based 79

Appendix D: Covid dummy variable extended treatment 82

D1) Covid dummy variable: Binary Specification 82

D2) Covid dummy variable: Phase-by-Phase Specification 93

List of Figures

Figure 1 - Variables plot 10
Figure 2 - ARDL variables trend 17
Figure 4 - Residuals versus Fitted Values 23
Figure 5 - Missing values 38
Figure 6 - Time series CAR with covid 39
Figure 7 - Time series Bank Size with covid 40
Figure 8 - Time series FSI with covid 40
Figure 9 - Time series NPL with covid 41
Figure 10 - Time series Personal loan with covid 42
Figure 11 - Time series ROA with covid 42
Figure 12 - Times series LDR with covid 43
Figure 13 - Time series business freedom with covid 44
Figure 14 - Times series investment freedom with covid 44
Figure 15 - Time series judicial effectiveness with covid 45
Figure 16 - Time series property rights with covid 46
Figure 17 - Time series PSI-20 with covid 46
Figure 18 - Time series PGB 10Y with covid 47
Figure 19 - Time series US/EU exchange rate with covid 47
Figure 20 - Time series of GNI with covid 48
Figure 21 - Time series house pricing with covid 49
Figure 22 - Time series interest rates with covid 50
Figure 23 - Time series inflation rate with covid 50
Figure 24 - Time series PCRH-D with covid 51
Figure 25 - Time series exports with covid 52
Figure 26 - Time series real GDP with covid 52
Figure 27 - Time series unemployment rate with covid 53
Figure 28 - Time series public debt with covid 54
Figure 29 - Time series firm loan with covid 54
Figure 30 - Correlation matrix of all variables 56
Figure 31 - Scatter plot of NPL determinants 60
Figure 32 - Variable pairs with high correlation 64
Figure 33 - ADF Statistic 67
Figure 48 - Residuals Vs Fitted values - Binary specification 92
Figure 49 - Residuals Vs Fitted values - Phase-by-phase specification 104

List of Tables

Table 1 - Description of variables and source 7
Table 2 - Variables of raw sample 9
Table 3 - ADRL model output (AIC & BIC) 20
Table 4 - Variables quarterly harmonised 37
Table 5 - Variance Inflation Factor (VIF) output 68
Table 6 – Elastic Net results 70
Table 7 - Vector Autoregression (VAR) results 73
Table 8 - ARDL variable selection 75
Table 9 – Elastic Net ARDL-based results 76
Table 10 - OLS simple regression, ARDL-based results 77
Table 11 - Dynamic ECM estimation ARDL-based results 79
Table 12 - First differenced ECM and lagged ECT 81
Table 13 - VIF results - Binary Specification 83
Table 14 – Elastic Net results - Binary Specification 83
Table 15 - Vector Autoregression (VAR) results - Binary Specification 86
Table 16 - ARDL model results - Binary Specification 87
Table 17 – Elastic Net ARDL-based - Binary Specification 88
Table 18 - OLS regression results ARDL-based – Binary Specification 89
Table 19 - Dynamic ECM estimation ARDL-based results - Binary Specification 90
Table 20 - First differenced ECM and lagged ECT - Binary Specification 91
Table 21 - VIF results - Phase-by-phase Specification 93
Table 22 – Elastic Net results - Phase-by-phase Specification 94
Table 23 - Vector Autoregression (VAR) results - Phase-by-phase Specification 97
Table 24 - ARDL Model results - Phase-by-phase Specification 99
Table 25 – Elastic Net ARDL-based - Phase-by-phase Specification 100
Table 26 - OLS regression model ARDL-based - Phase-by-phase Specification 101
Table 27 - Dynamic ECM estimation ARDL-based results - Phase-by-phase Specification 102
Table 28 - First differenced ECM and lagged ECT - Phase-by-phase Specification 103

List of Acronyms

ADF - Augmented Dickey-Fuller

AIC - Akaike Information Criterion

ARDL - Autoregressive Distributed Lag

BIC - Bayesian Information Criterion

CAR - Solvency / Capital Adequacy

EBA - European Banking Authority

ECB – European Central Bank

ECM - Error Correction Model

ECT - Error Correction Term

EU – European Union

FRED - Federal Reserve Economic Data

FSI - Financial Soundness Indicator

GDP – Gross Domestic Product

GNI – Gross National Income

IMF - International Monetary Fund

INE - Statistics Portugal

LDR – Loan-to-deposit ratio

NPL – Non-performing Loan

OLS - Ordinary Least Squares

PCRH-D - Private consumption of resident households by durability M€

PGB 10Y - 10-year Portugal Government Bond

PSI - Portuguese stock index

ROA – Return on Assets

SME - Small and Medium-sized Enterprises

VAR - Vector Autoregression model

VIF - Variance Inflation Factor

Introduction

In some economies, such as the Portuguese one, the financial system is predominantly bank-centred, an opinion validated in Mirović et al. (2024). These organisations, led by governance and shareholder agency, are looking for cost-effective ways to allocate and manage their assets to enhance shareholder value (Floros, 2020; Horvat et al., 2022). While the importance of revenue diversification has increased over the years, those lending activities continue to be the bedrock of a bank, with the highest profit potential. Nevertheless, for loans, the credit risk is complete, so that no matter how strong the credit approval is (Casu & Girardone, 2006). Efficient control of this risk is one of the most important indicators of a bank's efficiency of work.

The rapid growth of NPL, typically defined as loans due unpaid over 90 days, as set out by Washington (2005). Higher capital requirements for NPL will carry on that staircase of demands; profits are curtailed, and liquidity is restricted. The credit supply, in principle, will be choked (Chiesa and Mansilla 2020). For these reasons, they are often used as measures of financial strength and systemic risk (Klein 2013). Higher levels of NPLs that result in an increased likelihood of default make additional work for administration and reduce cost-effectiveness (Berger & DeYoung, 1997), all of which consumes banking stability - a matter particularly pressing in dynamic economies (Ghosh, 2017).

Defaults are also responsible for loss in the value of assets, which weakens bank capital and can, in turn, destabilise linked financial institutions (Morttinen et al., 2005). A high volume of NPLs will produce deficits in liquidity, cause financial intermediation to slow down and dissuade investment. In the final analysis, this will damage economic growth. Historical data additionally proves that they feature in financial crises (Kunt & Detragiache, 1998; Gonzalez-Hermosillo, 1999) and excludes one of the essential things about recovering trust as well as functionality (Klein, 2013).

The 2008-2009 global financial crisis and the 2010 European sovereign debt crisis highlighted both the instability generated by illusory credit expansion and the need for stricter supervision, along with larger capital buffers and less mistrust between banks and governments. To meet this challenge, the ECB has focused on reducing NPLs, giving

guidance in 2017. This was followed by the EBA's 2018 guidelines and Bank of Portugal disclosure standards for smaller institutions in 2019. These measures reflected the recognition that high NPL levels reduce interest earned while eroding both capital and resources available for productive use, at the same time impeding market access (Bank of Portugal, 2025).

Portuguese banking is a case in point. Scandals involving Bank *Espírito Santo* in 2014 and *Banif* in 2016 revealed scandalous balance sheet impairment due to toxic assets and defaults (Mendanha, 2021). Empirically, banks that are less profitable and more interest rate-exposed are observed to be under higher default risk in downturns, while stability is found only for the continued effects of the NPL ratio on the real growth of the GDP. The LDR, however, is expected to carry a positive relationship with NPLs, indicating liquidity risk vulnerabilities (Mendanha, 2021).

While these were also examined in previous crises, the COVID-19 pandemic represented a new kind of shock. On top of that, it mixed public health restrictions, loan moratoria and extraordinary fiscal and monetary operations. This unprecedented situation changed borrower conduct and bank risk exposures, which poses the questions of the continuity and the causes of NPLs in the post-COVID-19 scenario. Econometric modelling of NPLs, according to Klein (2013), provides an additional way to assess the longer-term outlook and stress-test such figures, which is motor oil for macroprudential oversight. Such tools are not only ubiquitous, but they also serve to light up hidden recesses within the system (Beck et al., 2013). In Portugal, high-frequency availability of data provides room for more refined modelling to track ongoing changes in NPL dynamics.

Portuguese banks have significantly reduced their NPL ratios over the past few years. Since April 2018, institutions with significant exposures have been required to submit strategic reduction plans, supported by EU and national initiatives to develop secondary NPL markets (Mendanha, 2021). Although there has been some progress, ongoing research of Portugal's banking sector lags larger eurozone countries; the influence of COVID-19 is a good example. In addition, existing studies generally emphasise either macroeconomic factors or bank-specific elements; few try to deal with them simultaneously while also considering pandemic effects.

The purpose of this study is to examine the effects of macroeconomic variables and bank-specific factors on NPLs in the Portuguese context, with special emphasis given to both pre- and post-COVID-19 periods. The objectives are threefold: (i) to assess the relationship between macroeconomic conditions and NPLs, (ii) to evaluate the effect of bank-specific characteristics, institutional and market level, and (iii) to see whether the pandemic itself changed what determines NPLs in Portugal's banking system.

Literature Review

The business cycle has an impact on the determinants of NPLs (Climent & Serrano 2019). To supervise the financial sector effectively, these dynamics must be monitored. Stress testing has become a useful tool for assessing the resilience of the banking system in adverse circumstances (Milenković et al, 2024). Many studies have identified external (macroeconomic) and internal (bank-specific) determinants of bad loans. Naili and Lahrichi (2022) found, for example, that external factors have a more significant impact, suggesting that banks are highly vulnerable to macroeconomic conditions despite risk management strategies that are focused on their own internal affairs.

A large amount of research confirms that macroeconomic variables like GDP growth, unemployment rates, inflation, and interest rates significantly influence the level of NPLs. Higher GDP growth generally reduces NPLs by increasing borrowers' ability to repay their debts. However, unemployed people and the high interest rates deter them from paying back loans (Salas & Saurina, 2002; Mendanha, 2021). In addition, Klein (2013) discovered that NPL risks go up when inflation rises and exchange rates depreciate, especially in emerging economies. When NPL drivers were examined by Kanapickienė et al. (2023), they were grouped into macroeconomic, financial market and institutional factors. It was shown that economic growth and a favourable stock market and a bad foreign exchange market reduce the incidence of NPL; conversely, indicators of the credit market and bond market often increase it. During crises like the COVID-19 pandemic, macroeconomic volatility becomes an important factor. Interestingly, Kanapickienė et al (2023) pointed out that in the CEE region during the outbreak of COVID-19, there was a reduced level of risk from NPLs caused by the pandemic, likely due to state support programmes and temporary moratoria on loan repayments. This means that crisis-response policies can partly decouple economic contractions from loan defaults.

From 2008 to 2013, both the EU and most of its member states experienced dislocation on the economic front. In 2009, GDP contracted by over 4% (Pjanić et al., 2020). In the Eurozone, the average employment growth was only 1.72% (European

Central Bank). During this period, the banking sector's NPL ratio averaged 8.06%; its ROA and ROE were 0.51% and 5.54% respectively (Milenković et al., 2024).

In terms of bank-specific indicators, including positions of minimum regulatory capital rates, net profit and LDR, NPLs are on an upward trend. According to the moral hazard hypothesis (Berger and DeYoung, 1997), banks with lower capital tend to lend more recklessly and have higher NPL ratios. This is substantiated by evidence from Spain, which links low capital ratios to higher NPLs (Salas & Saurina, 2002). Governance considerations, well-capitalised and profitable banks also have a slightly greater capacity to withstand the impact of rising NPLs (Thornton & Di Tommaso, 2020). Loans from the banking system grew as the criteria for obtaining credit were eased, compared with advances on the same terms acquired by borrowers (Keeton, 1999). Sinkey and Greenawalt (1991) found that strong banks are less affected by losses, thus underscoring the importance of extending or completing Basel III rules (Milojević & Redžepagić, 2021). Under conditions of stress, capital adequacy must be thoroughly stress-tested for both insolvency and risk exposure analyses, as best to have (Vunjak et al., 2015). Liquidity and profitability at a bank are determined by the degradation of macroeconomic and institutional factors (Radovanov et al., 2023), highlighting the importance of closely following their attributes to effectively handle NPL levels.

In assessing banking stability and performance, NPLs remain a crucial field. For different types of NPL, therefore, the forces driving them operate quite distinctively and call for diverse policies. As for households, growth in GDP and the rise in house prices can reduce NPLs, but unemployment and interest rates will have an opposite effect (Bofondi & Ropele, 2011). In the case of companies, defaults tend to be more cyclical as well as investment-driven (Bofondi & Ropele, 2011). Similar findings are reported in Greece (Louzis et al., 2012) and Serbia (Otašević, 2013). In Portugal, the role of these distinctions becomes particularly important because COVID-19 has had very unequal consequences across sectors: service-based SME's had a significant strain placed on them by temporary factors, but to a much slighter extent, households experienced gains from receiving three-month allowances for not working.

Loan dynamics at the firm level are also reflected in NPL patterns as well as the overall economic situation. For instance, *"zombie loans"* - loans provided at subsidies to failing businesses - reduce productivity, waste financial resources, and make the banking sector more vulnerable (Acharya et al., 2024; Constâncio, 2015). Empirical study ARDL co-integration, Bayesian Model Average example shows that GDP, public debt, construction of education while unemployment and domestic credit all have persisting effects on NPLs. But household credit as well as personal loans are crucial to the accumulation of bank sector risk factors (Foglia, 2022; Bolognesi et al., 2020; Maltritz & Molchanov, 2014; Khan et al., 2020). Recapitalisation was absent after weakly capitalised banks began defaulting on their own debts beginning in 2008, or else became *"zombie"* creditors for industries that either went bankrupt or survived in inefficient forms. This backwardness affects not only performance at the company level but also credit risk (Acharya et al., 2024).

The nature of NPL patterns is also affected by the institutional and structural settings. As for Betti et al. (2007), Rahman & Howlader (2012), Dillard et al. (2004), the development of the financial market, credit accessibility, governance quality, accounting integrity and individual company nature differentials were all factors that contributed to normal-risk and over-indebtedness levels. Regulatory responses in Portugal also involved such interventions as standardised EBA financial reporting templates and selling off assets to reduce exposure to NPLs.

All in all, macroeconomic variables, especially GDP growth, unemployment, inflation, and interest rates, are recognised as the main factors driving NPLs. Nevertheless, their impact on different borrower types or within different institutional environments is not yet fully understood. Despite the thousands upon thousands of papers and books that have been written on corporate finance in recent years, there are still many questions for scholars interested in public finances and particular circumstances. Firm-level loan effects have not been thoroughly researched in Portugal during the pre- and post-COVID-19 period.

Data Description and Methodology

Data Source and processing/preparation

Portugal's banking industry and economy study incorporates a quarterly data set covering early 2005 to early 2025. This data set contains four main variable categories: indicators specific to banks, market and institutional variables, and key macroeconomic time series. The dataset was collected from widely accepted and official sources, ensuring a high level of reliability and comparability.

Category	Variable	Impact on NPL	References	Source
Bank-Specific Variables	Solvency / Capital Adequacy Ratio (CAR)	Mixed, can lower household NPLs, but raise corporate NPLs due to risk-taking	Salas & Saurina (2002); Berger & DeYoung (1997); Thornton & Di Tommaso (2020)	Banck Portugal stat
	Profitability (ROA)	Negative (more resilient banks)	Thornton & Di Tommaso (2020)	Banck Portugal stat
	Bank Size	Mixed: larger banks may engage in riskier behaviour ("too big to fail")	Giannoccolo & Mansilla-Fernandez (2017); Berger et al. (2013)	Banck Portugal stat, *own calculations*
	PSI-20 (Portuguese Stock Index)	Negative	Kanapickienė et al. (2023); Abusharbeh (2020)	investing.com
	Firm Loans Portugal	Positive	Rahman & Howlader, (2024); Acharya et al., (2024)	Banck Portugal stat
	Financial Soundness Indicator (Credit at Risk / Gross Credit) - FSI	*Anlalysed along side with NPL*	*Implementation of the EBA's harmonised reporting framework 2015 to NPL*	Banck Portugal stat
	Personal Loans Portugal	Positive	Kanapickienė et al. (2023); Foglia, (2022)	Banck Portugal stat
	Loan-to-Deposit Ratio (LDR)	Positive	Maltritz & Molchanov (2014); Gounder, C., & Bajaj, V. (2022); Magdalinos, P., & Tsakalos, I. (2021)	ECB Data Portal
Institutional / Market-Level Variables	Portuguese 10y Governament Bond Yield (PGB 10Y)	Positive	Kanapickienė et al. (2023); Lehmann and Manz (2006)	investing.com
	Property Rights	Negative	Kanapickienė et al. (2023); Maltritz and Molchanov (2014)	The Heritage Foundation
	Investment Freedom	Negative	Kanapickienė et al. (2023); Maltritz and Molchanov (2014)	The Heritage Foundation
	Business Freedom	Positive	Kanapickienė et al. (2023); Maltritz and Molchanov (2014)	The Heritage Foundation
	Judicial Efectiveness	Negative	Kanapickienė et al. (2023); Maltritz and Molchanov (2014)	The Heritage Foundation
Macroeconomic Variables	Real GDP PT	Negative	Salas & Saurina (2002); Louzis et al. (2010); Bofondi & Ropele (2011); Otasevic (2013); Kanapickienė et al. (2023); (Mendanha, 2021)	FRED Economic Data
	Unemployment Rate	Positive	Louzis et al. (2010); Bofondi & Ropele (2011); Espinoza & Prasad (2010); (Mendanha, 2021)	FRED Economic Data
	Inflation Rate	Positive	Otasevic (2013); Klein (2013); Kanapickienė et al. (2023)	FRED Economic Data
	Interest Rate	Positive	Bofondi & Ropele (2011); Louzis et al. (2010); Otasevic (2013); Rinaldi & Sanchis-Arellano (2006); Bahruddin et al. (2018)	FRED Economic Data
	USD/EU Exchange Rate	Positive	Otasevic (2013); Ilievska et al. (2012); Kanapickienė et al. (2023); Klein (2013)	FRED Economic Data
	House Pricing PT	Negative	Bofondi & Ropele (2011)	FRED Economic Data
	Exports PT M€	Negative	Ilievska et al. (2012)	Banck Portugal stat
	Private consumption of resident households by durability M€ (PCRH-D)	Negative	Bofondi & Ropele (2011)	INE - Statistics Portugal
	Gross National Income (GNI)	Negative	Ilievska et al. (2022)	INE - Statistics Portugal
	Public Debt PT	Positive	Foglia, (2022); Bolognesi et al., (2020); Maltritz & Molchanov, (2014); Khan et al., (2020)	Federal Reserve Bank of St. Louis (FRED)
	COVID-19 Crisis (dummy/period)	Mixed, often negative due to policy support	Kanapickienė et al. (2023); (Mendanha, 2021)	DGS Oficial Site

Table 1 - Description of variables and source

Financial data were obtained from the Bank of Portugal, providing an easily accessible and reliable source of data about various facets of the domestic banking sector. The bank-specific indicators employed in this study include CAR ratio, bank size (measured as the logarithm of total assets), ROA, and the LDR, sourced from the ECB Data Portal. The data also includes the volume of personal and firm loans, and the FSI, which as credit at risk relative to total gross credit has heterogeneous effects on NPLs: increase when capitalization and profitability grow (Thornton & Di Tommaso, 2020), more loan exposure and a higher LDR level, however, result in the frequency of NPLs rising (Maltritz & Molchanov, 2014; Gounder & Bajaj, 2022) (Appendix A).

NPL information is only available from 2015 onwards, following improvements to the EBA's harmonised reporting framework, which ensures comparability across banks. To maintain consistency and extend the time horizon, the FSI variable was retained in the dataset, as it provides a proxy for asset quality from 2010 onwards and provides similar information to the NPL ratio according to the harmonised definition (see Appendix A).

The institutional and market environment was measured by a series of indices issued by the Heritage Foundation, comprising business freedom, investment freedom, judicial effectiveness and property rights. The strength of the institution and the protection of property rights appear to mitigate NPLs, but the greater freedom to undertake business, incidentally, may be increasing NPLs (Maltritz & Molchanov, 2014; Kanapickienė et al., 2023). In addition to those measures, the market data reflected in this article is taken from Investing.com: PSI-20 and PGB 10y yield. Stock market performance tends to lower financial fragility; by comparison, a higher sovereign bond yield is indicative of an aggravation of credit quality (Lehmann & Manz, 2006) (Appendix A).

A broad array of macroeconomic indicators was collected from the FRED and INE databases. These encompass such things as real GDP, GNI, exports, and real estate prices. In general, they tend to have negative correlations with NPLs (Bofondi & Ropele, 2011; Ilievski, 2022). In contrast, unemployment, inflation rates and the exchange rate of USD/EUR will tend to push up NPLs by making borrowers have less repayment capacity and having higher costs for servicing debt (Louzis et al., 2012; Otasevic, 2013). In

addition to this, public debt (obtained from FRED) is associated with higher NPLs as a result of sovereign (Foglia, 2022; Khan et al., 2020) - Appendix A.

The dummy variable represents the COVID-19 crisis, which is causing a structural rupture. Initially, it was thought that, with NPL ratios decreasing, policy support measures such as loan moratoria, state guarantees, or ECB interpositions would be able to mitigate the impact of the crisis, but the benefits have been mixed (Mendanha, 2021); Kanapickienė et al., 2023). In the baseline specification, the dummy variable takes the value of 0 before the onset of the pandemic and 1 from the beginning of the COVID-19 period until the end of the sample. An extended specification is also considered, whereby the dummy variable assumes the value of 0 until the emergence of the pandemic, 1 during the pandemic period, and reverts to 0 once the "end" of the pandemic is reached, defined as the moment when all contingency measures were lifted in Portugal (Appendix A).

By analysing NPLs in the Portuguese context, this dataset can be used to provide insights into how NPLs are influenced by factors such as finance, macroeconomics, institutions, and the market. With this increasing data coverage, it becomes possible to analyse what determines NPLs in Portugal in the pre- and post-COVID-19 periods, thus answering the main research question.

Category	Variable	Units	Frequency (original)	Raw Observations	Sample Period
Dependent variable	NPL	Ratio (%)	Quarterly (Dec)	37	2015Q4 – 2024Q4
Bank-specific Variables	CAR	Ratio (%)	Quarterly (Dec)	60	2010Q1 – 2024Q4
	Bank Size	€ (log/absolute)	Quarterly (Dec)	68	2008Q1 – 2024Q4
	Firm Loan PT	€ million	Monthly	243	2005M1 – 2025M3
	FSI	Ratio (%)	Quarterly (Dec)	60	2010Q1 – 2024Q4
	Personal Loan PT	€ million	Monthly	243	2005M1 – 2025M3
	ROA	Ratio (%)	Quarterly (Dec)	68	2008Q1 – 2024Q4
	LDR	Ratio (%)	Quarterly (Dec)	83	2004Q4 – 2025Q2
Institutional & Market Variables	Business Freedom	Index (0–100)	Annual (Jan)	21	2005 – 2025
	Investment Freedom	Index (0–100)	Annual (Jan)	21	2005 – 2025
	Judicial Effectiveness	Index (0–100)	Annual (Jan)	21	2005 – 2025
	Property Rights	Index (0–100)	Annual (Jan)	21	2005 – 2025
	PSI-20	Index	Monthly (Start)	245	2005M1 – 2025M5
	PGB 10y	Yield (%)	Monthly (Start)	245	2005M1 – 2025M5
	Exchange rate US/EU	USD/EUR	Daily (Business)	5305	2005M1 – 2025M5
Macroeconomic Variables	GNI	€ million	Quarterly (Oct)	81	2005Q1 – 2025Q1
	House Pricing PT	Index (2015=100)	Quarterly (Oct)	81	2005Q1 – 2025Q1
	Inflation Rate	% change (y-o-y)	Quarterly (Oct)	81	2005Q1 – 2025Q1
	Interest Rate	%	Quarterly (Oct)	81	2005Q1 – 2025Q1
	PCRH-D	€ million	Quarterly (Oct)	81	2005Q1 – 2025Q1
	Exports PT	€ million	Annual (Dec)	20	2005 – 2024
	Real GDP PT	€ million	Quarterly (Oct)	81	2005Q1 – 2025Q1
	Unemployment Rate	% labour force	Quarterly (Oct)	81	2005Q1 – 2025Q1
	Public Debt PT	€ million / % GDP	Annual (Dec)	21	2004 – 2024

Table 2 - Variables of raw sample

Exploratory analysis of the data

This section provides a quantitative analysis of Portuguese variables, emphasising their economic and financial importance. Covering the period 2005 to 2025, the graphical representation is divided into four categories: banking sector indicators, institutional quality and freedom indices, financial market indicators, and macroeconomic variables.

The whole aim is to find patterns, trends and swings in Portugal's economy and banking system, from periods of financial strain that might indicate other trend changes such as the world-wide financial crisis, the sovereign debt crash or even the COVID-19 pandemic. The trends observed here set the context for subsequent econometric analysis and suggest possible links between NPLs, macro-financial variables and institutions.

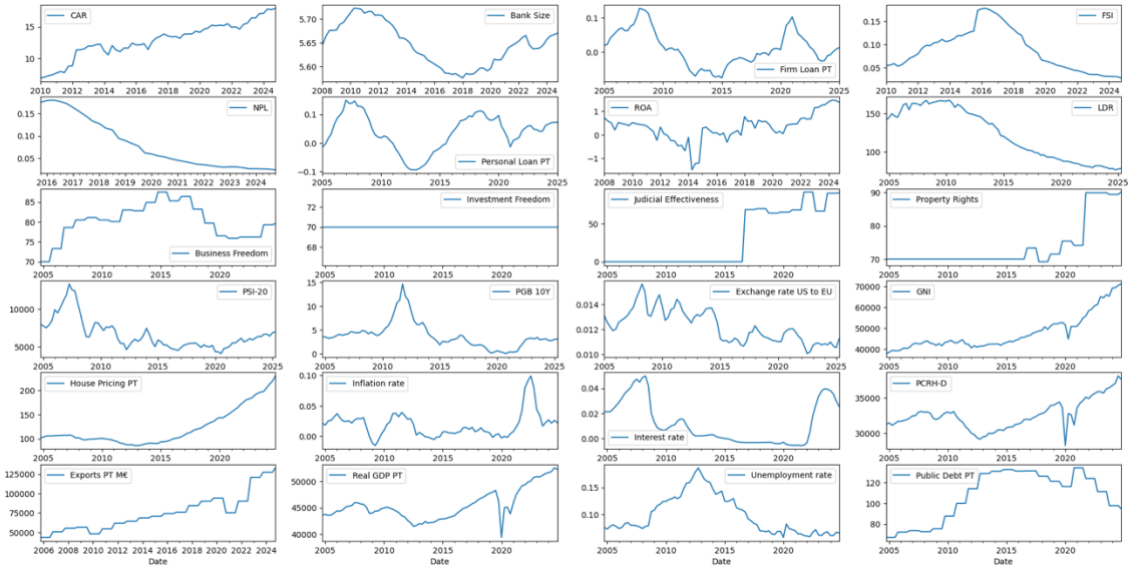


Figure 1 - Variables plot

On most of these measures of resilience, banking system indicators show significant gains after the sovereign debt crisis, as the capital adequacy ratio, in effect, almost doubled following 2012 due to regulatory pressure and Basel III requirements. Bank size declined slightly until 2017, reflecting deleveraging, before recovering as the sector gradually consolidated. Bank size overall fell slightly until 2017 due to deleveraging (though recovered with a slow grind in that year), and NPLs dropped sharply from 15% in 2016 to below 2% by 2024, showing large improvements in asset quality, recapitalisation and filled holes made by supervisors. In the meantime, ROA seemed

erratic over time, with negative readings in crisis years and 2020 when the COVID-19 shock hit, but recovered after 2021. In 2016, FSI was highest as NPLs, then fell drastically; as to the LDR, it went down from very high (nearly 160%) in 2008 to below 80% by 2025, showing a sustainable funding structure. Firm loans rose and soared around 2008, fell between 2010 and 2013, and stabilised at a significantly lower level during the recovery era. They spiked during the 2020 temporary upturn, mainly from COVID-related support and rose on a steady incline again in 2021 - 2024. Personal loans followed the same pattern: a pre-2008 rise, a sharp fall from 2009 to 2012, and then, very gradually, they both decreased and recovered from 2013. In 2020 an especially pronounced decline in lending (probably reflecting increased uncertainty during the pandemic) as creditors tightened risk assessment rules and credit access was restricted. Following this period, growth finally stabilised at modest levels.

Freedom and institutional quality indices were also developed over this period. Judicial effectiveness improved sharply from 2015, which might have been linked to structural reforms and major adjustments after the bailout. Property rights also remained steady and slightly improved after 2018, showing a solid institutional framework.

Financial market indicators reveal a volatile movement for the PSI-20, peaking in 2007- 2008 but then declining during the crisis. It really stayed relatively flat throughout the 2010's but rebounded somewhat post-2020 compared to growth in actual economic indicators, hinting at a slower capital market. Following the sovereign debt crisis, PGB 10y yields soared thanks to ECB intervention and subsequently stayed low, reflecting restored investor confidence. Moderately fluctuating, the USD/EUR exchange rate is as one would expect with euro membership.

Macroeconomic variables demonstrate steady GNI growth, interrupted by a COVID-19-related dip in 2020, followed by recovery. House prices - it is known that they were flat until 2013 and rose thereon - soared later because of factors like tourism, foreign investment and housing demand. During most of the period, inflation was low; it hit a peak around 2021-2022 due to rising global prices. Correspondingly, interest rates declined steadily until 2015, remained near zero until 2022, and then increased sharply as the ECB tightened monetary policy in response to rising inflation, aiming to stabilise the

economy and prevent potential macroeconomic disruption. Exports grew steadily, there was a small dip related to the COVID-19 recession in 2020, and Real GDP picked up where the foreign-exchange crisis left off. PCRH-D soared step by step, experienced crisis-induced declines, and unemployment rose to more than 16% in 2013 and decreased to 6% by 2025. After 2010, the public debt soared and went over 120%, then after 2017 dropped slightly. During COVID-19, some of the debt gradually increased again (Appendix B).

In summary, it is possible to find that between 2008 and 2013, the global financial and sovereign debt crises gave rise to high NPLs, spiking bond yields, peak unemployment, rising public debt, and contractions in both firm and personal lending. From 2014 to 2019, the recovery phase witnessed declining NPLs, revitalised bank capital, restored economic growth across the country, a reduction in the unemployment rate, and lively popular markets in both exports and housing. While the shock caused by COVID-19 continued through 2021, it left temporary contractions in GDP, consumption and employment, negative ROA for the banks, sudden but temporary shifts of bank lending direction, accumulated new debt and unemployment. Further, for the period between 2022 and 2025, the economy showed rapid recovery, with declining NPLs for example, stable growth, a gradual reestablishment of credit channels to normal operation, temporary rises in the rate of inflation and interest, and modest reductions in public debt (Appendix B).

Comparative Analysis of Pre- and Post-COVID-19 Trends

The post-COVID-19 period is characterised by a marked strengthening of the Portuguese banking sector, significant improvements in institutional indicators and a pronounced recovery in macroeconomic performance, although accompanied by new challenges related to inflation and financial markets.

Category	Variable	Pre-COVID Mean	Post-COVID Mean	Δ +/-	Pre-COVID Std	Post-COVID Std	Δ +/-
Bank-specific Variables	CAR	11.351	15.801	39.20%	2.212	1.182	-46.56%
	Bank Size	5.648	5.642	-0.10%	0.051	0.020	-60.78%
	Firm Loan PT	0.007	0.025	257.14%	0.055	0.038	-30.91%
	FSI	0.110	0.043	-60.90%	0.038	0.011	-71.05%
	NPL	0.132	0.037	-71.97%	0.040	0.011	-72.5%
	Personal Loan PT	0.034	0.047	38.24%	0.072	0.026	-63.89%
	ROA	0.123	0.778	532.52%	0.497	0.495	-0.40%
	LDR	136.900	81.718	-40.31%	26.749	3.750	-85.98%
Institutional & Market Variables	Business Freedom	81.051	76.975	-5.03%	4.555	1.426	-68.69%
	Investment Freedom	70.000	70.000	0,00%	0.000	0.000	0,00%
	Judicial Effectiveness	14.407	78.045	441.72%	27.936	12.383	-55.67%
	Property Rights	70.351	84.475	20.08%	1.131	7.405	554.73%
	PSI-20	6972.448	5761.407	-17.37%	2210.373	850.801	-61.51%
	PGB 10y	4.507	1.981	-56.05%	2.880	1.309	-54.55%
	Exchange rate US to EU	0.013	0.011	-15.38%	0.001	0.001	0,00%
Macroeconomic Variables	Gross National Income	43872.980	59640.400	35.93%	3591.564	7817.378	117.65%
	House Pricing PT	102.760	180.217	75.37%	12.818	25.994	102.79%
	Inflation rate	0.014	0.033	135.71%	0.014	0.031	121.43%
	Interest rate	0.011	0.014	27.27%	0.016	0.019	18.75%
	PCRH-D	31799.908	34896.320	9.74%	1275.535	2331.683	82.80%
	Exports PT (M€)	64629.317	103536.154	60.20%	13945.988	21286.248	52.63%
	Real GDP PT	44416.684	48884.590	10.06%	1647.154	3446.875	109.26%
	Unemployment rate	0.109	0.066	-39.45%	0.035	0.006	-82.85%
	Public Debt PT	104.455	115.551	10.62%	25.673	13.424	-4.77%

Table 3 - Comparative Analysis of Pre- and Post-COVID-19 Trends

Post-COVID can be observed a significant structural change in the banking sector, banks, financial markets, and the real economy. The banking system experienced an increase in CAR of close to 40%, ROA by a factor of more than 5, and in the case of NPLs, a decrease of almost 72%, bolstered by a reduction in FSI of 61%. This suggests cleaner balance sheets and higher profitability. Lending activity grew for firms (+257%) and households (+38%) despite the LDR shrinking by 40%, which indicated a healthier deposit base and prudency in granting credit.

Institutional factors also improved dramatically, with the effectiveness of the judiciary increasing over fourfold, and property rights strengthening by 20%, though business freedom was still weak, slightly lower (–5%) than in 2018. In financial markets, sovereign risks were reduced with a decline in PGB 10y yields of 56%, but equity results deteriorated, with a 17% drop in the PSI-20. The value of the exchange rate also depreciated slightly.

From a macroeconomic point of view, GNI increased by 36% GDP by 10%, exports by 60%, and household consumption by 10%. Unemployment dropped steeply (-39%), indicating a more robust recovery in the labour market. Meanwhile, house prices increased by 75%, inflation sped up by 136% and interest rates jumped by 27%, similar to the post-pandemic demand pressures and rate hikes. The public debt also increased by 11%, emphasising the fiscal costs of the crisis.

On balance, the post-COVID era is one in which bank resilience is stronger, institutional quality is better, macroeconomic expansion is robust, although inflation is higher, housing market appreciation has been rapid, equity sentiment is weaker and fiscal challenges still linger.

Methodology used in the study

This study investigates the drivers of NPLs in the Portuguese banking sector, with NPLs serving as the dependent variable. The methodological framework follows similar approaches employed by Foglia (2022) and Kjosevski, Petkovski, and Naumovska (2019). Additional details are provided in Appendix C.

Given the time-series nature of the data, it was first necessary to assess the stationarity of each variable. According to Gujarati & Porter (2009), if a series has a stable mean and variance over time, as well as its covariance is only affected by the time, it can be regarded as stationary. The The Augmented Dickey–Fuller (ADF) test, also in the tradition of Kjosevski et al. (2019), Foglia (2022), and Castro and Santos (2010), accepts the hypothesis that NPL data had a stationary I(0) value, making it suitable for inclusion in an ARDL framework accommodating I(0) and I(1) variables (Appendix C).

To identify relevant predictors and mitigate multicollinearity, Pearson correlation analysis was conducted (Cotovio, 2023; Kanapickienė et al., 2023; Gounder and Bajaj, 2022). High correlation between variables was particularly closely scrutinised. Consequently, PCRH-D was chosen to represent macroeconomic indicators, Judicial Effectiveness represented institutional quality, and NPL was the dependent variable, while excluding redundant measures such as FSI. A COVID dummy variable was included to account for pandemic-related structural breaks. Further analysis was carried out to test for multicollinearity. When excessive VIF values were found, these aspects of the model were removed to avoid unstable or difficult-to-interpret coefficient estimates (Appendix C).

The VAR framework was then applied to capture dynamic interdependencies between NPLs, bank-specific indicators, and macroeconomic variables. NPLs were treated as the dependent variable. This approach allows bi-directional relationships for all variables irrespective of their position in any given list (Marques et al., 2020). Prior studies (Espinoza and Prasad, 2010; Klein, 2013) demonstrate that NPL shocks have a large effect on credit creation and macroeconomic conditions. This underpins the decision to use VAR. The hypothetical model was well supported in its collective form by the data

analysis performed upon it. Consequently, found dynamics such as unemployment that have continuity without being sustained by any single cause; Judicial Effectiveness (polarity negative); significant lagged CAR in solvency and Interest Rate behaviour showing both autoregressive patterns as well as a certain sensitivity to macroeconomic shocks (see Appendix C).

On the other hand, the core of analysis lies in ARDL methodology, which gives a flexible and robust way to distinguish between the immediate impact of an event and its long-term outcomes. ARDL is especially suited to small samples and can handle different orders of integrated variables (I(0) or I(1)), unlike traditional co-integration methods that require them all to be integrated uniformly (Kjosevski et al., 2019; Foglia, 2022). The short-run dynamic adjustment is captured by lags of difference terms on both dependent and independent variables, while lags of levels reveal long-run equilibrium relationships. The framework also allows for the derivation of a dynamic ECM and to quantify the rate at which deviations from long-term equilibrium are reversed (Banerjee et al., 1993; Pesaran & Shin, 1995). Optimal lag selection is guided by information criteria such as AIC and BIC to ensure both statistical validity and parsimony.

The ARDL model is formally specified as follows:

$$\Delta y_t = \beta_0 + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \sum_{j=0}^{q} \delta_j \Delta x_{t-j} + \lambda_1 y_{t-1} + \lambda_2 x_{t-1} + \varepsilon_t$$

where y_t is the dependent variable (NPL), x_t represents the independent determinants, and Δ_{yt-i} and Δ_{xt-j} are lagged differences capturing short-term effects with coefficients β_i and δ_j. The lagged levels y_{t-1} and x_{t-1} capture the long-run equilibrium, with λ_1 and λ_2 representing long-term coefficients, while ε_t is the error term accounting for unexplained variation. It is worth mentioning that the equation simplifies the model to include only one independent variable. If more than one independent variable is present, the summation terms for short-run dynamics and long-run equilibrium would be expanded to include each of these additional determinants.

To refine variable selection and address multicollinearity, Elastic Net regression was applied, combining LASSO and Ridge penalties to retain informative predictors

while controlling for correlated variables (Zou & Hastie, 2005; Tibshirani, 1996). After this, OLS estimation of the coefficients on the ARDL model provided more robust coefficients. This way, by incorporating ARDL into Elastic Net and ECM analysis, the methodology effectively distinguishes short-term and long-term effects. Thus, it is possible to understand the real meanings of these phenomena rather than simply observing them. Additionally, the model yields both economic insight and policy-relevant information on drivers of NPLs in the Portuguese banking sector (Appendix C).

Variables were retained based on their predictive importance, as identified by Elastic Net regression coefficients, while VIF diagnostics ensured that multicollinearity was within acceptable limits. Furthermore, the significance of lagged effects in the VAR analysis informed the inclusion of variables with relevant dynamic interactions. The graphical (Figure 2) shows the variables recommended for inclusion in the ARDL model, distinguishing between main determinants and optional variables for robustness checks. Variables such as ROA, Property Rights, Personal Loan PT, and the NPL ratio itself were excluded due to insignificant predictive power or overlap with the dependent variable.

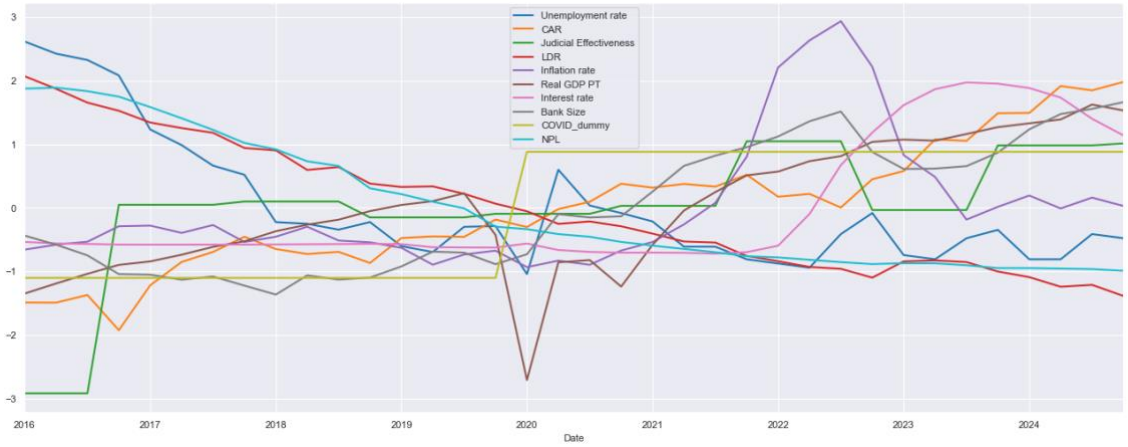


Figure 2 - ARDL variables trend

Using the selected variables and the period of 2016–2024, it is possible to observe a few interesting patterns. In the pre-COVID period (2016–2019), NPLs and the unemployment rate continuously decreased, signalling economic improvement. LDR went down a bit, and Inflation stayed near its historical average. Real GDP collapsed during the COVID shock in 2020, and interest rates dipped due to ultra-loose monetary policy. Crucially, NPLs did not soar overnight, suggesting the success of policy responses

such as loan moratoria. During the recovery period from COVID (2021–2022), GDP recovered along with bank size and CAR, whereas inflation stayed low. Inflation escalated in the post-COVID period, and interest rates rose. Despite the macro-financial headwinds, NPLs were still below their historical average, indicating the adequacy of policy buffers as well as enhanced bank fundamentals. Some key co-movements were seen, like in NPLs and unemployment pre-COVID, GDP growth with bank size expansion, and the post-2022 separation of inflation and interest rates, signalling a change in the policy environment.

Results and Discussion

The ADF test confirmed, i.e. accepts the hypothesis, that the NPL series is stationary at levels (ADF = -4.893, $p < 0.001$), supporting its inclusion as an I(0) variable in the ARDL framework. Correlation analysis and VIF diagnostics identified and removed redundant predictors, while Elastic Net regression further refined the variable set. This process ensured that the ARDL model could be estimated with a stable and interpretable set of determinants. It also provides a solid foundation for VAR forecasting, allowing the analysis of the dynamic relationships between NPLs, bank-specific factors, and macroeconomic conditions. Detailed supporting statistical results are presented in Appendix C.

The VAR model specification (Espinoza & Prasad, 2010; Klein, 2013; Marques et al., 2020), derived from Elastic Net and VIF diagnostics, retained LDR, Unemployment Rate, CAR, Judicial Effectiveness, Interest Rate, the PGB 10y yield, and Inflation. On the other hand, COVID-19 was removed as a dummy variable from the dynamic specification: corresponding to a shift effect capturing an unprecedented shock to the system rather than a typical explanatory variable that evolves just like any of the others in the VAR. The model had a strong fit to the data (AIC = -32.65, BIC = -30.16, log-likelihood = 279.74), and economic theory gave plausible results. Unemployment persisted, but this was mitigated by Judicial Effectiveness ($p = 0.012$). CAR was highly autocorrelated but adversely affected by high LDR ($p = 0.008$) and inflation ($p = 0.016$). Interest Rates were mainly autoregressive, but fell with CAR and PGB 10y yields, rising with inflation if anything. Judicial Effectiveness itself was persistent, but responded to interest-rate shocks by decreasing, which interest rates being negative, would mean less than zero income. Inflation was positively influenced by lagged CAR and negatively by interest rates. In the end, the results emphasise the key role played by institutional quality, liquidity and solvency in shaping NPLs macro-financial dynamics. Appendix C provides more detailed statistical output and full regression tables to enable a closer look at the diagnostics of this model, familiarisation with the coefficients.

ARDL Estimation Results

The estimated ARDL model with lag length selected by AIC or BIC. The results revealed that the model is highly compelling with a strong performance. The AIC-selected model is an ARDL (1,1,0,0,1), and the BIC-selected model is shaped like ARDL (1,0,0,0,1). Both models showed that NPL is a sticky problem, with NPL.L1 (the lagged dependent variable) highly significant in both (AIC: 0.8535, BIC: 0.8074, $p < 0.01$), indicating that credit defaults accumulate over time.

```
Dependent var in columns: True
Independent vars in columns: True
Optimal ARDL order using AIC: (1, 1, 0, 0, 1)
Summary of ARDL model selected by AIC:
                              ARDL Model Results
==============================================================================
Dep. Variable:                      NPL   No. Observations:                   36
Model:             ARDL(1, 1, 0, 0, 1)   Log Likelihood                166.343
Method:                Conditional MLE   S.D. of innovations              0.002
Date:                 Sun, 07 Sep 2025   AIC                           -314.686
Time:                         15:06:30   BIC                           -300.688
Sample:                     06-30-2016   HQIC                          -309.854
                          - 12-31-2024
==============================================================================
                    coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const            -0.8552      0.256     -3.338      0.002      -1.381      -0.330
NPL.L1            0.8535      0.064     13.416      0.000       0.723       0.984
LDR.L0            0.0024      0.000      4.967      0.000       0.001       0.003
LDR.L1           -0.0007      0.001     -1.375      0.180      -0.002       0.000
Real GDP PT.L0  1.131e-06   5.01e-07      2.256      0.032    1.02e-07    2.16e-06
Bank Size.L0      0.1172      0.047      2.470      0.020       0.020       0.215
COVID_dummy.L0    0.0168      0.005      3.218      0.003       0.006       0.028
COVID_dummy.L1   -0.0087      0.004     -2.463      0.020      -0.016      -0.001
==============================================================================
Optimal ARDL order using BIC: (1, 0, 0, 0, 1)
Summary of ARDL model selected by BIC:
                              ARDL Model Results
==============================================================================
Dep. Variable:                      NPL   No. Observations:                   36
Model:             ARDL(1, 0, 0, 0, 1)   Log Likelihood                165.158
Method:                Conditional MLE   S.D. of innovations              0.002
Date:                 Sun, 07 Sep 2025   AIC                           -314.316
Time:                         15:06:34   BIC                           -301.874
Sample:                     06-30-2016   HQIC                          -310.021
                          - 12-31-2024
==============================================================================
                    coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const            -0.8101      0.258     -3.139      0.004      -1.339      -0.281
NPL.L1            0.8074      0.055     14.697      0.000       0.695       0.920
LDR.L0            0.0020      0.000      4.992      0.000       0.001       0.003
Real GDP PT.L0  1.485e-06   4.37e-07      3.399      0.002     5.9e-07    2.38e-06
Bank Size.L0      0.1010      0.047      2.163      0.039       0.005       0.197
COVID_dummy.L0    0.0193      0.005      3.897      0.001       0.009       0.030
COVID_dummy.L1   -0.0097      0.003     -2.769      0.010      -0.017      -0.003
==============================================================================
```

Table 3 - ADRL model output (AIC & BIC)

Bank-specific determinants had an effect on NPL dynamics. LDR was positive for NPLs (positive coefficients under both AIC and BIC), suggesting that when topping up their credit reserves bank will purchase risk insurance with deposits. NPLs will therefore increase sharply. Bank size also had a positive association (AIC: 0.1172, BIC: 0.1010). Larger banks are typically not staked in just one type of loan but have all sorts of loans. The COVID dummy picked up an immediate rise in NPLs (AIC: 0.0168, BIC: 0.0193), followed by a negative lagged effect (AIC: -0.0087, BIC: -0.0097), again reflecting temporary credit stress related to the pandemic and some relief through policy steps.

Real GDP PT was the only macroeconomic determinant that had a significant negative impact (AIC: 1.131e-06, BIC: 1.485e-06, $p < 0.05$) on credit risk, affirming that credit risk is pro-cyclical. Unemployment and interest rates were not statistically significant, indicating an indirect effect or being overshadowed by other determinants under Portugal conditions.

In terms of model selection, AIC implied ARDL (1,1,0,0,1) and BIC implied the less complex ARDL (1,0,0,0,1). Given this slightly higher complexity, the AIC-selected model was retained as the preferred model, as it can capture short-term dynamics, such as the lagged effect of LDR on NPLs, which is relatively theoretically important for the understanding of the bank liquidity adjustment and the credit risk propagation. The two specifications provided qualitatively similar results on NPLs main determinants, but the AIC model covers more immediate responses with good explanatory power (Log-likelihood: 166.343, AIC: -314.686).

Elastic Net applied to the ARDL results reinforced the predominance of bank-level determinants. Lagged NPL and bank size exerted the strongest effects, with LDR and COVID shocks also significant. Real GDP exhibited context-specific dynamics, and the COVID dummy variable confirmed the transitory effects of the pandemic. Additionally, the Ordinary Least Squares (OLS) estimation of the ARDL (1,1,0,0,1) model produced very high explanatory power ($R^2 = 0.998$), with all core variables statistically significant (see the output in Appendix C).

To distinguish between short-run and long-run dynamics, the ECM was estimated. In the short run, fluctuations were mainly the result of ΔLDR, highlighting Liquidity's prompt effect on NPLs. Other short-term shocks, such as ΔGDP and ΔCOVID, were weak or a non-starter. The ECT was negative and large, implying that deviations from the long-run equilibrium would be corrected in time. It was marginally significant at 10%. Diagnostic examinations confirmed that the data were approximately normal and that multicollinearity was low, but there were slight signs of both serial correlation and heteroskedasticity.

ARDL, Elastic Net, OLS, ECM and VAR models each point to the robust persistence of NPLs as the most salient feature of Portuguese credit risk dynamics. In all specifications, lagged NPLs remain not only significant but highly so, illustrating that these defaults camouflaged and aggregated together over time so much as to represent de facto enormous proportions of the total pool. Bank-specific determinants - especially LDR or bank size - lie at the heart of short-term sustainability, while economic variables with national scope, such as GDP or institutional structure, play out more in an enduring COVID-related shocks were severe but temporary, reflecting the cushioning influence of policy interventions such as moratoria and capital injections (Appendix C).

When different modelling approaches are combined, more powerfully can reinforce these results thus, the ARDL can capture not only immediate but also long-run causes. Elastic Net controlling multicollinearity while preserving parsimony makes sure that the outcome is still in focus; OLS makes clear that the robustness of parameter estimations holds; ECM also makes a distinction between transient shocks and long-run convergence. Lastly, VAR finds both endogenous terms linked with each other in a bidirectional way. This yields the result where institutional quality and bank solvency are found to be stabilising forces. Together, these approaches demonstrate that prudent liquidity management, resilient capital buffers, and effective institutional frameworks are critical to reducing NPL vulnerability.

To determine whether there is an adequate fit of the ECM, a Residuals versus Fitted values plot was generated (Figure 3). The diagnostic is useful for evaluating model

specification, testing the presence of heteroskedasticity, and testing for systematic errors in predictions (Banerjee, et al., 1998; Gujarati and Porter, 2009).

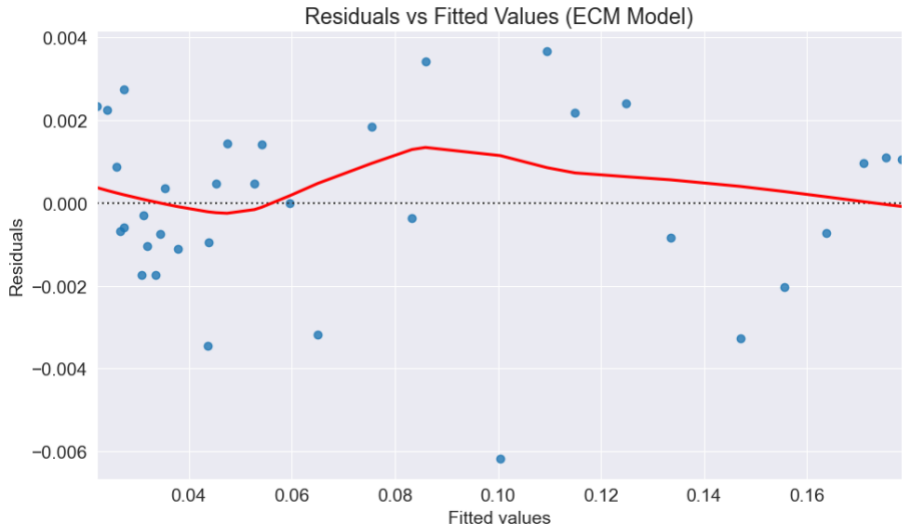


Figure 3 - Residuals versus Fitted Values

They are mostly centred around the zero, which shows that the model does not have much of a systematic over- or under-prediction. The LOESS smoother reveals slight curvature, indicating small non-linearities or omitted dynamics in the short-run relationships. At higher fitted values it is possible to observe a bit wider residual spread, suggesting the presence of mild heteroskedasticity. Some outliers are also observed that may influence the fit of the model.

In general, the residuals appear to resemble white noise, which confirms that the model could capture the short and long-run dynamics. However, despite the above comments, these findings suggest that some more lags, interaction terms, or robust standard errors might make the robustness of the specification and the test of the inferences more reliable (Banerjee et al., 1998; Gujarati & Porter, 2009).

Examining the different dataset specifications provides further insights. In the original specification, lagged NPL, LDR, bank size, and GDP consistently emerge as the key determinants, while the COVID dummy captures short-term shocks. The AIC-selected ARDL (1,1,0,0,1) model economy and Elastic Net fully demonstrate the domination of bank-specific factors. The OLS results reveal an extremely high degree of explanatory power (R^2~99.8%), while the ECM outputs suggest that short-run variability

in controlled variables is mainly caused by changes in LDR. Finally, ECT reflects partial long-term declaring independence at a slow rate (Appendix C).

In light of the previous analysis is essential now to reconsider the treatment of the COVID-19 dummy variable. In the first specification, the dummy equalled 0 before and 1 during the pandemic and until de end of the sample period. However, these methods do not take into account the post-pandemic period, which might be analytically relevant if the consequences of COVID-19 wane when the actual crisis is over. To control for this, two robust specifications were considered: binary specification and phase-by-phase specification.

In the binary specification, the COVID dummy variable took the value 1 from March 1, 2020[1], to September 30, 2022[2], thus isolating the epidemic period. In the phase-by-phase specification, time was divided into four periods: Pre-COVID (January 1, 2008, to February 29, 2020), COVID-2020, COVID-2021, COVID-2022, and Post-COVID (October 1, 2022[3], to December 31, 2024), capturing year-specific and transition effects for more detailed insights into the evolution of credit risk (Appendix D).

In the Binary part of the COVID specification, both interest rates and the COVID dummy are significant in ARDL and Elastic Net models. These have shrinking effects in OLS and ECM, estimated that their primary significance is when delayed NPL is included, indicating short-term dynamics mainly driven by NPL persistence. The ECM coefficient is negative but not statistically significant, implying only a jumping adjustment towards the long-run equilibrium. The phase-based approach uses years before the pandemic outbreak era to record time-varying dynamics. NPL persistency remains the dominant factor, with short-run effects that also hint at LDR changing in phases, where found delayed unemployment impacts emerging from time pollution or other such causes. The ECM output once again demonstrates liquidity as an important

[1] *The first Situation Report on COVID-19 in Portugal was published by the Directorate-General of Health (DGS) on 3 March 2020 (Relatório de Situação No. 1, available at: https://www.dgs.pt/em-destaque/boletim-informativo-n-5-01032020-pdf.aspx)*

[2] *First daily bulletin on 5 March 2020 (Boletim Informativo No. 5, available at: https://www.dgs.pt/em-destaque/relatorio-de-situacao-n-001-03032020-pdf.aspx).*

[3] *The official conclusion of the pandemic in Portugal is marked as 30 September 2022, when all contingency measures were lifted (Portugal Government, 2022, available at: https://www.portugal.gov.pt/pt/gc23/comunicacao/noticia?i=fim-do-estado-de-alerta&utm_source.com).*

short-run driver, and recent NPL variations. However, high collinearity between variables limits the robustness of any results this year (Appendix D).

Combining results obtained from all specifications, there are key features that appear throughout. NPL persistency is the strongest determinant, following the previously established pattern of entrenched credit risk. Bank-specific factors, especially liquidity, decide whether short-term or long-term developments show any signs at all. Macroeconomic variables show fewer stable effects: GDP was relevant in the base specification, unemployment has delayed effects in phase-based models, and interest rates matter mainly under the binary COVID specification. Pandemic shocks exert immediate but transient impacts, largely absorbed once the lagged NPL is accounted for. Across all models, ECM results indicate partial, often weak, adjustment toward long-run equilibrium (Appendix D).

There are certain common findings across all the models. The NPL entrenchment is the strongest driving factor, while bank-specific factors, such as LDR and size, are corrective factors; and macroeconomic (or health crisis) shocks are felt immediately but are transitory. Overall, it emerges that the original specification offers the most robust findings, whereas the binary and phase-based measures offer further and less robust insights. In general, NPL dynamics in Portugal are dominantly driven by persistence and bank-specific fundamentals, and COVID-related impacts have been far from persistent. Appendix D provides detailed results for all models and specifications.

Conclusions and implications

The study answers the following research question: "What is the impact of macroeconomic variables and bank-specific factors on non-performing loans in Portugal, before and after COVID-19?" The macroeconomic, bank-level, and unprecedented COVID-19 shock factors had been examined to determine the determinants of NPL. Findings reveal that NPLs continue to be an important indicator of financial well-being, restricting liquidity, reducing the profitability and the availability of credit. In the Portuguese case, influenced as it is by both historical crises and the coronavirus, strong banking institutions and close supervision are necessary for maintaining the stability of the whole system.

The results show that macroeconomic variables have a moderating effect on NPLs, but that they are less stable than those of bank-specific ones. GDP showed a pro-cyclical relationship with credit risk, and unemployment seemed to be particularly relevant under COVID-specific circumstances. The results further illustrate the sensitivity of NPLs to economic cycles and confirm that macroeconomic shocks can exacerbate financial instability, albeit in a non-persistent manner. Bank-specific features were found to be pivotal in determining NPL processes. Past defaults had been the strongest determinant (indicating the depth of default risk and how historical events shape future expectations). Liquidity management, determined by the LDR, also had an immediate stabilising effect, while larger bank size increased the risk, in line with evidence of moral hazard and capital adequacy considerations. Overall, bank-level fundamentals emerged as significant determinants of both immediate and sustained NPL behaviour. There are a few more shocks and shocks were introduced by the COVID-19 pandemic, rather than on the NPLs themselves. Short-term increases were largely absorbed when lagged NPLs were considered, implying that policy measures (loan moratoria, government subsidies and capital injections) mitigated system weakness. This supports the notion that the pandemic has temporarily reshaped borrowers' behaviour and bank exposures to vulnerability, but that it has not significantly changed the longer-term forces behind the development of defaults.

Methodologically, ARDL models distinguished short- and long-run effects, Elastic Net identified major predictors, OLS ascertained robustness, ECM explored adjustment dynamics, and VAR emphasised institutional quality's stabilising role. Combined, these approaches give an exhaustive picture of NPL dynamics in Portugal.

These considerations have policy implications, including strong liquidity management, robust capital buffers, and institutional resilience. Stemming from the sovereign debt crisis, frameworks set out in regulatory regimes emerged that have continued to be useful, but additional refinements, including stress-testing and monitoring of household and business exposures, have potentially helped to smooth credit risk.

The study faces limitations, notably the use of aggregate NPL data, the restricted post-pandemic horizon, and the need for deeper governance indicators. Future research should include household and business sectors, analyse even the *"zombie credit"*, integrate data on fiscal variables, extend sample decay, and ideally involve a cross-country comparison within the European monetary union. Across Portugal, a general assessment of NPL behaviour yields the following results: persistence and Bank-specific fundamentals first; then macroeconomic and pandemic shocks affect in non-persistent ways. As most components can be made subject to the regime switches, it becomes clear that of these, only non-seasonal (intermittent) factors should be treated as standalone units. A key contribution of this study is that it integrates multiple modelling approaches in one place, providing a more complete understanding of the risks involved in lending to small investment central economies. The findings provide evidence-based recommendations for effectively supervising and stabilising financial institutions.

References

Acharya, V. V., Crosignani, M., Eisert, T., & Eufinger, C. (2024). Zombie credit and (dis-) inflation: evidence from Europe. The Journal of Finance, 79(3), 1883-1929. https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.13342

Balgova, M., Plekhanov, A., & Skrzypinska, M. (2017). Reducing non-performing loans: Stylized facts and economic impact. Ms., EBRD and Oxford, December. https://mariabalgova.github.io/Balgova_Plekhanov_SkrzypinskaNPL2.pdf

Banco de Portugal, 2025. Bank Size: Total Assets (log). https://bpstat.bportugal.pt/serie/12504493

Banco de Portugal, 2025. Financial Stability Indicator (Asset Quality). https://bpstat.bportugal.pt/dados/series?mode=graphic&svid=_M2-AAAAAAA.!l!A!C!10!!!False!!!!!!!%3A1uE8ak%3ABpZ6DeNz9VRJGGMA5FAACObfCHQ&series=12504572&redirectTo=%2Fdados%2Fselecao

Banco de Portugal, 2025. Loans to Firms in Portugal. https://bpstat.bportugal.pt/serie/12530265.

Banco de Portugal, 2025. Loans to Households in Portugal. https://bpstat.bportugal.pt/serie/12558228

Banco de Portugal, 2025. Non-Performing Loans (NPLs). https://bpstat.bportugal.pt/serie/12504544

Banco de Portugal, 2025. Return on Assets (ROA). https://bpstat.bportugal.pt/dados/selecao

Banco de Portugal, 2025. Solvency / Capital Adequacy Ratio (CAR). https://bpstat.bportugal.pt/serie/12504543

Bandiera, L., Cuaresma, J. C., & Vincelette, G. A. (2010). Unpleasant Surprises: Sovereign Default Determinants and Prospects, Policy Research Working Papers, 5401.

Banerjee, A., et al. (1993). Co-integration, error correction, and the econometric analysis of non-stationary data. Oxford university press. https://books.google.com/books?hl=pt-PT&lr=&id=LfVQEAAAQBAJ&oi=fnd&pg=PR7&dq=Banerjee,+A.,+Dolado,+J.+J.,+Galbraith,+J.+W.,+%26+Hendry,+D.+F.+(1993).+Cointegration,+Error+Correction,+and+the+Econometric+Analysis+of+Non-Stationary+Data,+Oxford:+Oxford+University+Press.&ots=-8WAXNAqPU&sig=rP9s7hvT9ocKIsEGY65v22iMf_w

Banerjee, A., et al. (1998). Error-correction mechanism tests for cointegration in a single-equation framework. Journal of time series analysis, 19(3), 267-283. https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-9892.00091

Basto, R. B. (2013). Uma política macroprudencial para a estabilidade financeira. Lisboa: Banco de Portugal, Relatório de Estabilidade Financeira, 77-92. https://www.bportugal.pt/sites/default/files/anexos/papers/ar201304_p.pdf

Beck, R., Jakubik, P., & Piloiu, A. (2013). Non-performing loans: What matters in addition to the economic cycle? (No. 1515). ECB Working Paper. https://www.econstor.eu/handle/10419/153948

Berger, A. N., & DeYoung, R. (1997). Problem loans and cost efficiency in commercial banks. Journal of banking & finance, 21(6), 849-870. https://www.sciencedirect.com/science/article/pii/S0378426697000034

Betti, G., Dourmashkin, N., Rossi, M., & Ping Yin, Y. (2007). Consumer over-indebtedness in the EU: measurement and characteristics. Journal of economic studies, 34(2), 136-156. https://www.emerald.com/insight/content/doi/10.1108/01443580710745371/full/html

Bijsterbosch, M., & Falagiarda, M. (2015). The macroeconomic impact of financial fragmentation in the euro area: Which role for credit supply?. Journal of International Money and Finance, 54, 93-115. https://www.sciencedirect.com/science/article/pii/S0261560615000285

Biyanwila, B. P. U. K., & Anuradha, P. A. N. S. (2023). Behavioral factors affecting household over-indebtedness: A systematic review. Journal of Business Research and Insights (former Vidyodaya Journal of Management), 9(I). http://journals.sjp.ac.lk/index.php/vjm/article/view/6373

Bofondi, M., & Ropele, T. (2011). Macroeconomic determinants of bad loans: evidence from Italian banks. Bank of Italy Occasional Paper, (89).

Bolognesi, E., Compagno, C., Miani, S., & Tasca, R. (2020). Non-performing loans and the cost of deleveraging: The Italian experience. Journal of Accounting and Public Policy, 39(6), 106786. https://doi.org/10.1016/j.jaccpubpol.2020.106786

Bonfim, D., Farinha, L., & Félix, S. (2019). Economics Synopsis Credit and the economy: lessons from a decade of research at Banco de Portugal. Volume V, 73. https://www.bportugal.pt/sites/default/files/anexos/papers/re201909_e.pdf

Castro, G., & Santos, C. (2010). Determinantes das taxas de juro e do crédito bancário. Boletim Económico. https://www.bportugal.pt/sites/default/files/anexos/papers/ab201002_p.pdf

Castro, V. (2013). Macroeconomic determinants of the credit risk in the banking system: The case of the GIPSI. Economic modelling, 31, 672-683. https://www.sciencedirect.com/science/article/pii/S0264999313000308

Casu, B., & Girardone, C. (2006). Bank competition, concentration and efficiency in the single European market. The Manchester School, 74(4), 441-468. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9957.2006.00503.x

Chiesa, G., & Mansilla-Fernández, J. M. (2021). The dynamic effects of non-performing loans on banks' cost of capital and lending supply in the Eurozone. Empirica, 48(2), 397-427. https://link.springer.com/article/10.1007/s10663-020-09475-5

Cincinelli, P., & Piatti, D. (2017). Non-performing loans, moral hazard & supervisory authority: The Italian banking system. Journal of Financial Management, Markets and Institutions, (1), 5-34. https://www.rivisteweb.it/doi/10.12831/87058

Constancio, V. (2015, August). Understanding inflation dynamics and monetary policy. In Speech at the Jackson Hole Economic Policy Symposium (Vol. 29, pp. 455-483). https://www.kansascityfed.org/Research/documents/7077/ConstancioPanelist_JH2015.pdf

Cotovio, C. I. B. (2023). Non-Performing Loans in Portugal During Crisis Periods 2007 Financial Crisis and the Covid-19 Pandemic Crisis (Master's thesis, Universidade NOVA de Lisboa (Portugal)). https://search.proquest.com/openview/5236531e138d8c7d59eee2ede8fcbe60/1?pq-origsite=gscholar&cbl=2026366&diss=y

Da Silva, F. B. P. M. (2022). Determinantes do Crédito Malparado: Evidência do setor bancário europeu (UE27) (Master's thesis, Universidade do Porto (Portugal)). https://search.proquest.com/openview/b90e7cacb5c79df08792c0590a67a523/1?pq-origsite=gscholar&cbl=2026366&diss=y

Dillard, J. F., Rigsby, J. T., & Goodman, C. (2004). The making and remaking of organization context: Duality and the institutionalization process. Accounting, Auditing & Accountability Journal, 17(4), 506-542. https://www.emerald.com/insight/content/doi/10.1108/09513570410554542/full/html

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. Econometrica: journal of the Econometric Society, 251-276. https://www.jstor.org/stable/1913236

Espinoza, M. R. A., & Prasad, A. (2010). Nonperforming loans in the GCC banking system and their macroeconomic effects. International Monetary Fund. https://books.google.com/books?hl=pt-PT&lr=&id=KpUYEAAAQBAJ&oi=fnd&pg=PA3&dq=Espinoza+and+Prasad+(2010),&ots=PUudaiOV3n&sig=0rZ8VUCF1klTlCllZN2xaDK3MW8

European Central Bank, 2025. Loans-to-Deposits Ratio (MFIs excluding ESCB, Portugal). https://data.ecb.europa.eu/data/datasets/GFS/GFS.A.N.PT.W0.S13.S1.C.L.LE.GD.T._Z.XDC_R_B1GQ._T.F.V.N._T

Farinha, L., & Noorali, S. (2004). Indebtedness and wealth of Portuguese households. Economic bulletin and financial stability report articles, Banco de Portugal. https://www.bportugal.pt/sites/default/files/anexos/papers/ar200402_e.pdf

Ferreira, C. (2022). Determinants of non-performing loans: A panel data approach. International Advances in Economic Research, 28(3), 133-153. https://link.springer.com/article/10.1007/s11294-022-09860-9

Floros, C. (2020). Banking development and economy in Greece: Evidence from regional data. Journal of Risk and Financial Management, 13(10), 243. https://www.mdpi.com/1911-8074/13/10/243

Foglia, M. (2022). Non-performing loans and macroeconomics factors: The Italian case. Risks, 10 (1), 21. https://www.mdpi.com/2227-9091/10/1/21

Frade, C., Lopes, C. A., Jesus, F., & Ferreira, T. (2008). Um perfil dos sobreendividados em Portugal. https://estudogeral.uc.pt/bitstream/10316/87662/1/Um%20perfil%20dos%20sobreendividados%20em%20Portugal.pdf

FRED, 2025. Exports of Goods and Services (GDP by Expenditure, Portugal). https://fred.stlouisfed.org

FRED, 2025. General Government Gross Debt (% of GDP, Portugal). https://fred.stlouisfed.org/series/GGGDTAPRT188N

FRED, 2025. Interest Rates: 3-Month Interbank, Portugal. https://fred.stlouisfed.org/series/IR3TIB01PTQ156N

FRED, 2025. Real Gross Domestic Product, Portugal. https://fred.stlouisfed.org/series/CLVMNACSCAB1GQPT

FRED, 2025. Real Residential Property Prices, Portugal. https://fred.stlouisfed.org/series/QPTR368BIS

FRED, 2025. U.S. Dollars to Euro Spot Exchange Rate. https://fred.stlouisfed.org/series/DEXUSEU

FRED, 2025. Unemployment Rate (15–74 years, Portugal. https://fred.stlouisfed.org/series/LRUN74TTPTQ156N

Friedman, J. H., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. Journal of statistical software, 33, 1-22 https://www.jstatsoft.org/article/view/v033i01

Friedman, R. (2020). Use of Machine Learning Algorithms to Predict Illness Severity and Diagnosis Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) From Patient Data. https://dash.harvard.edu/entities/publication/2f7e92ac-23a9-47f7-ba7d-f3983afdac9f

Gandrud, C., & Hallerberg, M. (2017). How not to create zombie banks: lessons for Italy from Japan (No. 2017/6). Bruegel Policy Contribution. https://www.econstor.eu/handle/10419/173102

Garcia, T., & Vieira, P. (2017). O mercado de crédito especializado ao consumo em Portugal. https://ap1.sib.uc.pt/bitstream/10316.2/93559/1/o_mercado_de_credito.pdf

Ghosh, A. (2017). Sector-specific analysis of non-performing loans in the US banking system and their macroeconomic impact. Journal of Economics and Business, 93, 29-45. https://www.sciencedirect.com/science/article/pii/S0148619517300140

Gounder, C., & Bajaj, V. (2022). Do Microeconomic Determinants of Country Impact the Level of Non-Performing Loans? Evidence from International Data. https://www.researchgate.net/profile/Chitra-Gounder-2/publication/363799336_Do_Microeconomic_Determinants_of_Country_Impact_the_Level_of_Non-Performing_Loans_Evidence_from_International_Data/links/632e8bb286b22d3db4d9e1f4/Do-Microeconomic-Determinants-of-Country-Impact-the-Level-of-Non-Performing-Loans-Evidence-from-International-Data.pdf

Gujarati, D. N., & Porter, D. C. (2009). Basic econometrics. McGraw-hill. https://cbpbu.ac.in/userfiles/file/2020/STUDY_MAT/ECO/1.pdf

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1849872

https://research.wu.ac.at/en/publications/unpleasant-surprises-sovereign-default-determinants-and-prospects

https://www.sciencedirect.com/science/article/pii/0167487095000134

Investing.com, 2025. Portugal Government 10-Year Bond Yield – Historical Data. https://pt.investing.com/rates-bonds/portugal-10-year-bond-yield-historical-data

Investing.com, 2025. PSI-20 Historical Data. https://pt.investing.com/indices/psi-20-historical-data

Kanapickienė, R., Keliuotytė-Staniulėnienė, G., Vasiliauskaitė, D., Špicas, R., Neifaltas, A., & Valukonis, M. (2023). Macroeconomic factors of consumer loan credit risk in central and eastern European countries. Economies, 11(4), 102. https://www.mdpi.com/2227-7099/11/4/102

Kjosevski, J., Petkovski, M., & Naumovska, E. (2019). Bank-specific and macroeconomic determinants of non-performing loans in the Republic of Macedonia: Comparative analysis of enterprise and household NPLs. Economic research-Ekonomska istraživanja, 32(1), 1185-1203. https://hrcak.srce.hr/228754

Klein, N. (2013). Non-performing loans in CESEE: Determinants and impact on macroeconomic performance. International Monetary Fund. https://books.google.com/books?hl=pt-PT&lr=&id=LT8ZEAAAQBAJ&oi=fnd&pg=PA3&dq=Klein+(2013)+npl&ots=XvCDXh-ZM&sig=sPY84CninhlCyvBvqD8xMX7GpNw

Korzeb, Z., & Niedziółka, P. (2020). Condition of banks listed on the Warsaw Stock Exchange during the first 3 months of the pandemic in Poland. Journal of Banking and Financial Economics, (2 (14), 5-20. http://www.scielo.org.co/scielo.php?pid=S0123-59232021000200226&script=sci_arttext&tlng=em

Lea, S. E., Webley, P., & Walker, C. M. (1995). Psychological factors in consumer debt: Money management, economic socialization, and credit use. Journal of economic psychology, 16(4), 681-701. https://www.sciencedirect.com/science/article/pii/0167487095000134

Louzis, D. P., Vouldis, A. T., & Metaxas, V. L. (2012). Macroeconomic and bank-specific determinants of non-performing loans in Greece: A comparative study of mortgage, business and consumer loan portfolios. Journal of banking & finance, 36(4), 1012-1027. https://www.sciencedirect.com/science/article/pii/S0378426611002895

Magdalinos, P., & Tsakalos, I. (2021). Macroeconomic and financial determinants of non-performing loans: Evidence from PIIGS. Bulletin of Applied Economics, 8(2), 97-110. https://scholar.archive.org/work/rjshmyaiszh6phtzhnilz2vxwi/access/wayback/https://www.risk market.co.uk/bae/journals-articles/issues/macroeconomic-and-financial-determinants-of-non-performing-loans-evidence-from-piigs/?download=attachment.pdf

Mah, J. S. (2000). An empirical examination of the disaggregated import demand of Korea—the case of information technology products. Journal of Asian economics, 11(2), 237-244. https://www.sciencedirect.com/science/article/pii/S1049007800000531

Makri, V., Tsagkanos, A., & Bellas, A. (2014). Determinants of non-performing loans: The case of Eurozone. Panoeconomicus, 61(2), 193-206. http://panoeconomicus.org/index.php/jorunal/article/view/72

Maltritz, D., & Molchanov, A. (2014). Country credit risk determinants with model uncertainty. International Review of Economics & Finance, 29, 224-234. https://www.sciencedirect.com/science/article/pii/S1059056013000531

Marques, C., Martinho, R., & Silva, R. (2020). Empréstimos não produtivos e oferta de crédito: Evidência para Portugal. Revista de Estudos Económicos, 6, 1-26. https://www.bportugal.pt/sites/default/files/anexos/papers/re202003_pt_0.pdf

Matos, C. C. (2014). Determinantes da concessão de crédito bancário: aplicação a Portugal (Master's thesis, Universidade dos Acores (Portugal)). https://search.proquest.com/openview/04afee68c1cb9092563842b35452fe0f/1?pq-origsite=gscholar&cbl=2026366&diss=y

Mendanha, M. A. R. (2021). Evolução dos non performing loans (NPL) no sector bancário português. https://repositorio.ipl.pt/entities/publication/3326e2bc-3da8-400d-ad48-8161a0c0b46b

Milenković, N., Kalaš, B., Mirović, V., & Andrašić, J. (2024). Static and Dynamic Modeling of Non-Performing Loan Determinants in the Eurozone. Mathematics, 12(21), 3323. https://www.mdpi.com/2227-7390/12/21/3323

Mina, F. C. (2016). Determinants of household debt refinancing in Portugal (Master's thesis, Universidade de Lisboa (Portugal)). https://search.proquest.com/openview/7ed429491e3ee9bbfd75eb1072a44094/1?pq-origsite=gscholar&cbl=2026366&diss=y

Mixão, H. M. P. (2020). Non-performing loans: the Portuguese case (Master's thesis, Universidade de Lisboa (Portugal)). https://search.proquest.com/openview/6a671fc5ae61189ac2fbff3df7f91a32/1?pq-origsite=gscholar&cbl=2026366&diss=y

Otaševic, D. (2013). Macroeconomic determinants of the quality of banks' loan portfolio in Serbia (No. 27). https://ideas.repec.org/p/nsb/wppnbs/27.html

Pattichis, C. A. (1999). Price and income elasticities of disaggregated import demand: results from UECMs and an application. Applied Economics, 31(9), 1061-1071. https://www.tandfonline.com/doi/abs/10.1080/000368499323544?casa_token=MLllypUt5bkAAAAA:lEDAWZYVvbLtGGZSAoWLW7LlKQSYkBOzu2B2uFJw79uNWMTCmhoxruNsf_ZDHzH_-4hvW2f6KYbRng

Pesaran, M. H., & Shin, Y. (1995). An autoregressive distributed lag modelling approach to cointegration analysis (Vol. 9514, pp. 371-413). Cambridge, UK: Department of Applied Economics, University of Cambridge. http://request-attachments.storage.googleapis.com/bRv1Dv9b8djCBcOc9hAnvz9gHg1eA4HF1gOGySUCokMEpfXnVvGzMvfj3Hu6YWtrdDaYEeP7BAVQP0FTkZs8JQKRIh6HNaElqtPV/An_Autoregressive_Distributed_Lag_Modeling_Approac.pdf

Plikas, J. H., Kenourgios, D., & Savvakis, G. A. (2024). COVID-19 and Non-performing loans in Europe. Journal of Risk and Financial Management, 17(7), 271. https://www.mdpi.com/1911-8074/17/7/271

Prodregosa, F. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research. vol, 12, 2825-2830. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.ElasticNetCV.html?utm_source=chatgpt.com

Rahman, M. A., & Howlader, M. S. (2025). The impact of accounting environment, firm and loan attributes on non-performing loan ratios of countries: the moderating role of good governance. International Journal of Disclosure and Governance, 22(1), 219-243. https://link.springer.com/article/10.1057/s41310-024-00252-5

Roman, A., & Irina, B. (2015). An empirical analysis of the macroeconomic determinants of non-performing loans in EU28 banking sector. Revista economica, 67(2), 108. http://oldeconomice.ulbsibiu.ro/revista.economica/archive/6729roman_bilan.pdf

Salas, V., & Saurina, J. (2002). Credit risk in two institutional regimes: Spanish commercial and savings banks. Journal of financial services research, 22(3), 203-224. https://link.springer.com/article/10.1023/a:1019781109676

Santos, B. M. (2017). Proposed Solution for the Non-Performing Loans in the Portuguese Banking System (Master's thesis, Universidade NOVA de Lisboa (Portugal)). https://search.proquest.com/openview/318924eff3dda819e562b809e47687fa/1?pq-origsite=gscholar&cbl=2026366&diss=y

Sinkey Jr, J. F., & Greenawalt, M. B. (1991). Loan-loss experience and risk-taking behavior at large commercial banks. Journal of Financial services research, 5(1), 43-59. https://link.springer.com/article/10.1007/BF00127083

Soares, E. A. C. (2017). Determinants of nonperforming loans in Portugal (Master's thesis, Universidade de Lisboa (Portugal)). https://search.proquest.com/openview/b3d742ec82926e9c4edaf4042e13343d/1?pq-origsite=gscholar&cbl=2026366&diss=y

Statistics Portugal (INE), 2025. Gross National Income, Portugal. https://www.ine.pt/ngt_server/attachfileu.jsp?look_parentBoui=215821289&att_display=n&att_download=y

Statistics Portugal (INE), 2025. Private Consumption of Resident Households (Durable Goods). https://www.ine.pt/ngt_server/attachfileu.jsp?look_parentBoui=661706543&att_display=n&att_download=y

Suradi, D., Siregar, H., & Sartono, B. (2021). Non-Performing Loan Determinants during COVID-19 Pandemic. https://www.academia.edu/download/77785303/IJRR037.pdf

The Heritage Foundation, 2025. Business Freedom https://www.heritage.org/index/pages/all-country-scores

The Heritage Foundation, 2025. Investment Freedom. https://www.heritage.org/index/pages/all-country-scores

The Heritage Foundation, 2025. Judicial Effectiveness. http://www.heritage.org/index

The Heritage Foundation, 2025. Property Rights https://www.heritage.org/index/pages/all-country-scores

Thornton, J., & Di Tommaso, C. (2021). The effect of non-performing loans on credit expansion: Do capital and profitability matter? Evidence from European banks. International journal of

finance & economics, 26(3), 4822-4839. https://onlinelibrary.wiley.com/doi/abs/10.1002/ijfe.2042

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B: Statistical Methodology, 58(1), 267-288. https://academic.oup.com/jrsssb/article-abstract/58/1/267/7027929

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society Series B: Statistical Methodology, 67(2), 301-320. https://academic.oup.com/jrsssb/article-abstract/67/2/301/7109482

Pjanić, M., Milenković, N., Andrašić, J., Kalaš, B., & Mirović, V. (2020). Public debt's predictors in EU: Evidence from members and non-members of European Monetary Union. *Economic research-Ekonomska istraživanja*, *33*(1), 3562-3579. https://hrcak.srce.hr/file/370120

Naili, M., & Lahrichi, Y. (2022). Banks' credit risk, systematic determinants and specific factors: recent evidence from emerging markets. *Heliyon*, *8*(2). https://www.cell.com/heliyon/fulltext/S2405-8440(22)00248-1

Climent-Serrano, S. (2019). Effects of economic variables on NPLs depending on the economic cycle. *Empirical Economics*, *56*(1), 325-340. https://link.springer.com/article/10.1007/s00181-017-1362-y

Lehmann, H., & Manz, M. (2006). *The exposure of Swiss banks to macroeconomic shocks-an empirical investigation* (No. 2006-04). https://ideas.repec.org/p/snb/snbwpa/2006-04.html

Ilievski, A. (2022). Non-Performing Loans in North Macedonia-Lessons Learned?. *Acta Polytechnica Hungarica*, *19*(8), 91-109. https://www.researchgate.net/profile/Andrej-Ilievski-2/publication/364172981_Non-Performing_Loans_in_North_Macedonia_-_Lessons_Learned/links/633fde99ff870c55ce0a4da5/Non-Performing-Loans-in-North-Macedonia-Lessons-Learned.pdf

Khan, M. A., Siddique, A., & Sarwar, Z. (2020). Determinants of non-performing loans in the banking sector in developing state. *Asian Journal of Accounting Research*, *5*(1), 135-145. https://www.emerald.com/insight/content/doi/10.1108/ajar-10-2019-0080/full/html

Appendices

Appendix A: Cleaning and Frequency-Harmonised Data

To ensure equality in dates across each data set, Dates were unified, and the indices recalculated for studies. A summary part is given in the variable output file of each variable name, several observation types and frequency start-end dates for line charts.

	Name	Size	Type	Freq	Start	End
1	CAR	60	float64	Q-DEC	2010-03-31	2024-12-31
2	Bank Size	68	float64	Q-DEC	2008-03-31	2024-12-31
3	Firm Loan PT	81	float64	Q-DEC	2005-03-31	2025-03-31
4	FSI	60	float64	Q-DEC	2010-03-31	2024-12-31
5	NPL	37	float64	Q-DEC	2015-12-31	2024-12-31
6	Personal Loan PT	81	float64	Q-DEC	2005-03-31	2025-03-31
7	ROA	68	float64	Q-DEC	2008-03-31	2024-12-31
8	LDR	83	float64	Q-DEC	2004-12-31	2025-06-30
9	Business Freedom	81	float64	Q-DEC	2004-12-31	2024-12-31
10	Investment Freedom	81	int64	Q-DEC	2004-12-31	2024-12-31
11	Judicial Effectiveness	81	float64	Q-DEC	2004-12-31	2024-12-31
12	Property Rights	81	float64	Q-DEC	2004-12-31	2024-12-31
13	PSI-20	83	float64	Q-DEC	2004-12-31	2025-06-30
14	PGB 10Y	83	float64	Q-DEC	2004-12-31	2025-06-30
15	Exchange rate US to EU	82	float64	Q-DEC	2005-03-31	2025-06-30
16	GNI	81	float64	Q-DEC	2004-12-31	2024-12-31
17	House Pricing PT	81	float64	Q-DEC	2004-12-31	2024-12-31
18	Inflation rate	81	float64	Q-DEC	2004-12-31	2024-12-31
19	Interest rate	81	float64	Q-DEC	2004-12-31	2024-12-31
20	PCRH-D	81	float64	Q-DEC	2004-12-31	2024-12-31
21	Exports PT M€	77	float64	Q-DEC	2005-12-31	2024-12-31
22	Real GDP PT	81	float64	Q-DEC	2004-12-31	2024-12-31
23	Unemployment rate	81	float64	Q-DEC	2004-12-31	2024-12-31
24	Public Debt PT	81	float64	Q-DEC	2004-12-31	2024-12-31

Table 4 - Variables quarterly harmonised

Since the data didn't have the same frequency, it was harmonised to a quarterly frequency. The monthly data were then averaged to create daily averages, and the annual numbers were filled in where they didn't match the quarterly ones. Finally, these numbers were merged into a single dataset. Missing values must be retained to visualise this type of coverage over time. Year after year, coverage improved. Between 2016 and 2024, there were no missing values, with two isolated gaps in 2025, which can be attributed to reporting delays due to time constraints (see Figure 4).

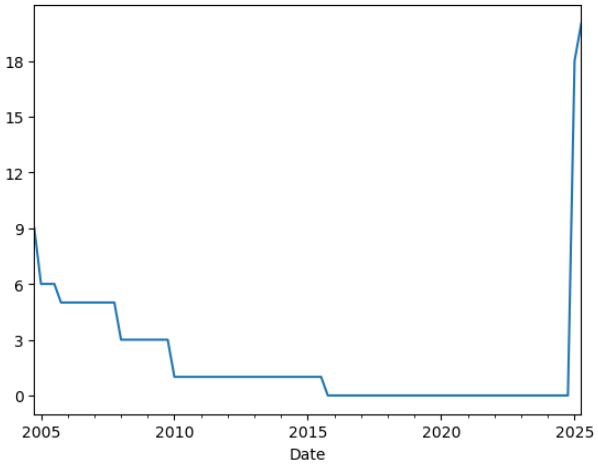


Figure 4 - Missing values

By doing this preparation, all variables were being brought into step and were *"comparable"* over time. A dummy variable was introduced to represent COVID-19. Although the official date of the first confirmed COVID-19 cases in Portugal was 2 March 2020, the period that is referred to as *"Pandemic"* began on 1 March 2020. This was the date on which DGS began issuing daily publications about COVID-19 across all media. Therefore, all COVID-19 dummy observations before this date were coded as 0, while those from 1 March 2020 onwards were coded as 1. In this way, the analysis can exclude pre-pandemic trends and segregate the effects of COVID-19 on economic and financial indicators. This variable permits the analysis to separate the economic and financial indicators of the pandemic from those originating earlier.

Appendix B: Descriptive Analysis of the Dataset

B1) Analysis of all 4 variable categories with COVID-19 split

All variables are represented to distinguish pre- and post-COVID dynamics using the dummy variable in a graph presentation.

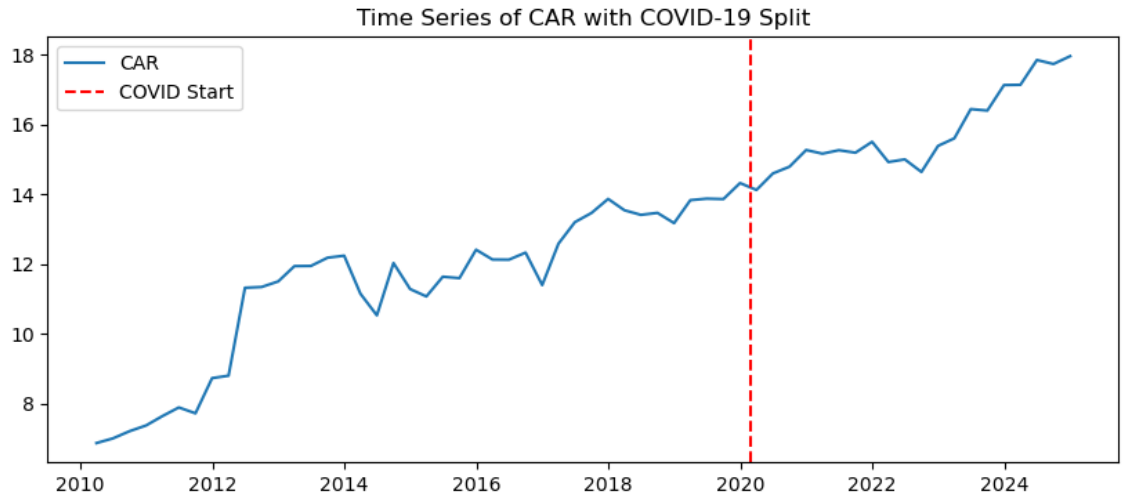


Figure 5 - Time series CAR with covid

The CAR ratio chart from 2010 to 2025 shows a steady increase in the solvency ratio from about 7% to around 18%. This trend has thrown the shock of COVID-19 aside. Pre-pandemic, increased capital buffers were built up from 7% to 14%, asset quality generally improved, and the introduction of Basel III reforms led to a more prudent balance sheet management, with only minor oscillations incident on periods of macroeconomic uncertainty. Anti-cyclical capital preservation, regulatory relief, as well as fiscal and monetary support, paved the way for continuing the upward trend after 2020.

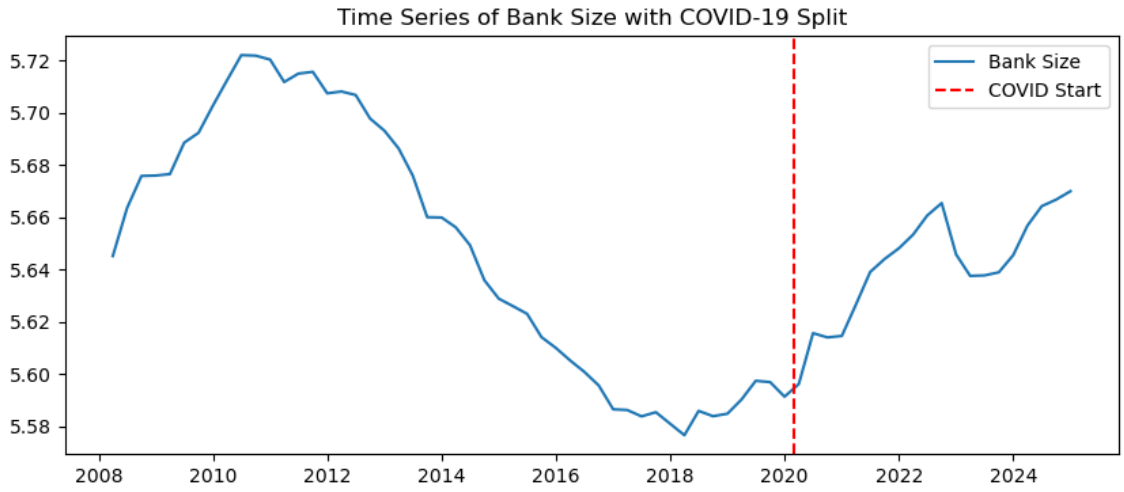


Figure 6 - Time series Bank Size with covid

Bank size modestly expanded between 2008 and 2011-2012, most likely a result of the balance-sheet expansion and consolidation that followed the crisis. From 2012 to the end of 2018, a continuous contraction took place in bank size, reflecting the deleveraging and restructuring, asset sales, and less prominent role of banks in the economy against the background of developments like tightening regulation, weak credit demand in this post-crisis era. Even with and lingering effects of the European sovereign-debt crisis following the initial shock of the COVID-19 pandemic, bank size began to recover steadily from 2020. This may have been due to expansionary monetary and fiscal policies increasing government securities holdings, renewed balance sheet growth as the economy modelled once again.

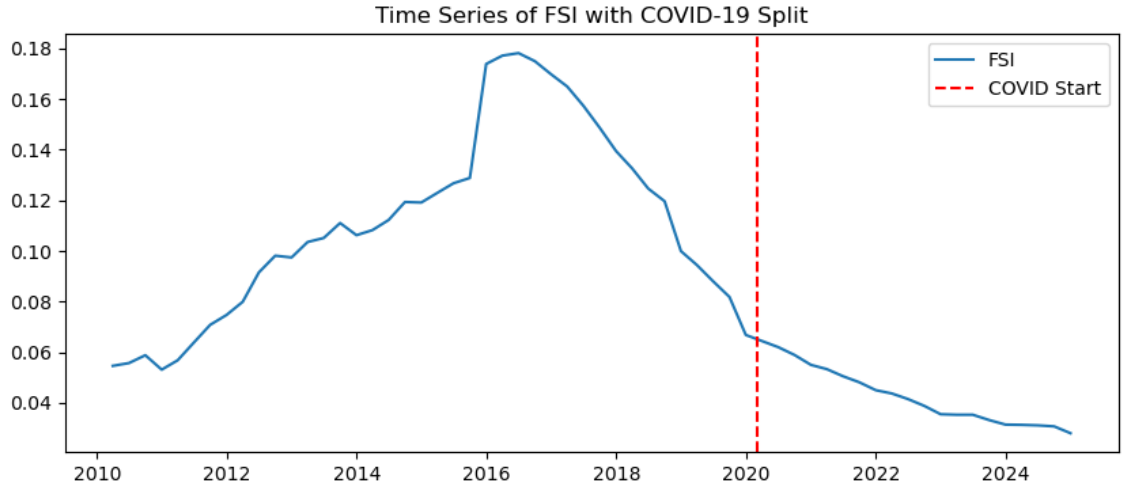


Figure 7 - Time series FSI with covid

Between 2012 and 2016, FSI rose from around 5% to close to 18%, reflecting the lagged effects of the sovereign debt crisis, declining credit quality, and increased regulatory oversight. Between 2016 and 2019, FSI gradually declined to 6%, helping prevent a sharp economic slowdown and aiding balance sheet repair in banks overall. Surprisingly, FSI continued to decline during COVID-19; by 2025, it was revised downward to around 2.5%, likely due to public support measures, regulatory forbearance, and delayed recognition of impairments.

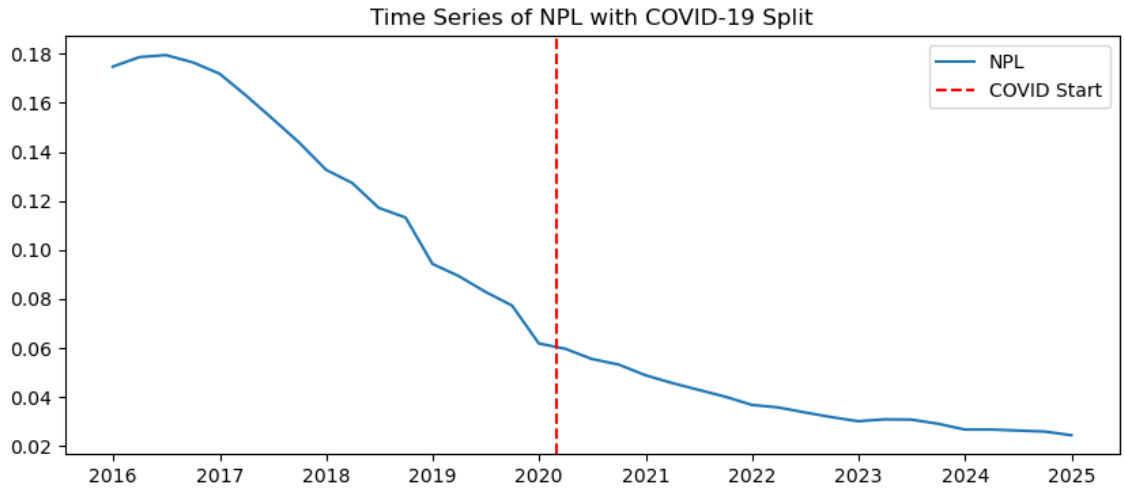


Figure 8 - Time series NPL with covid

In Figure 8, the NPL time series begins in 2015, under the EBA's harmonised definition. In 2017-2016, NPLs remained high, having emerged in a post-crisis period of weak recovery. Thanks to the macroeconomic recovery, dynamic resolution strategies, and regulatory pressure, they fell from over 8% to nearly zero in 2019. Starting in 2012, shortly before reaching historically low levels of around 3.5%, they stabilised at a slightly lower percentage between 2023 and 2025, suggesting the persistence of stricter credit criteria. The absence of a pandemic-driven increase indicates a possible structural break in credit-risk transmission, justifying the use of COVID-specific dummy variables and structural break tests in econometric modelling.

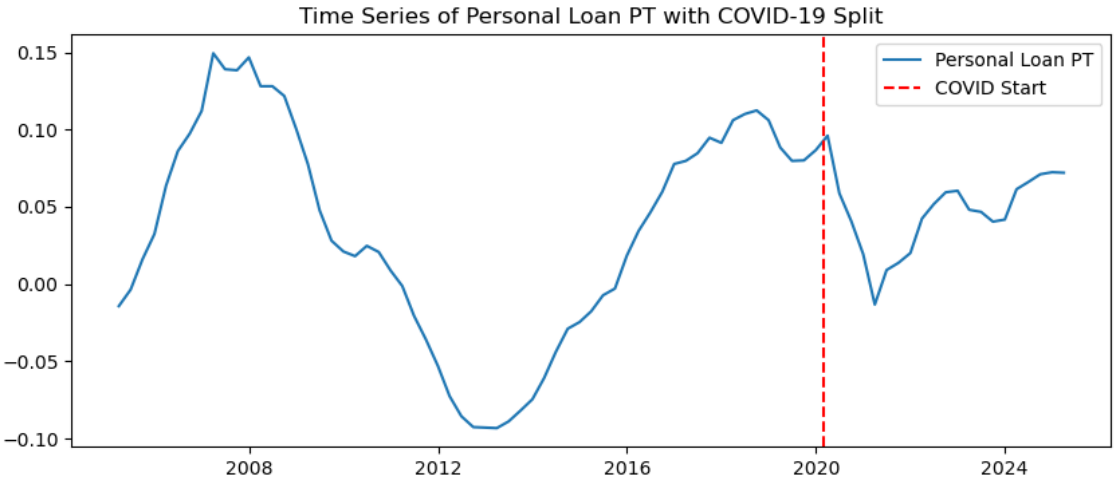


Figure 9 - Time series Personal loan with covid

Figure 9 illustrates that Personal loans increased strongly between 2006 and 2008, reflecting the pre-crisis credit boom, with robust consumer demand and easy access to credit. From 2009 to 2012, growth declined sharply and by 2012 was negative amid the global financial crisis, reflecting tightened credit conditions and household deleveraging. The gradual recovery of loan growth from the end of 2012 to 2017 followed an improving confidence and macroeconomic conditions and received a small but steady increase in the average annual total between 2018 and 2019 due to low interest rates and better job markets. The COVID-19 outbreak in 2020 saw an abrupt halt in this, which can be explained by all the uncertainty and, consequently, restricted access to credit. A partial and volatile recovery during 2021-2022 evolved. By 2023-2024, growth had stabilised and began at moderately positive levels, slightly below pre-pandemic rates after reflecting household caution and disciplined banking.

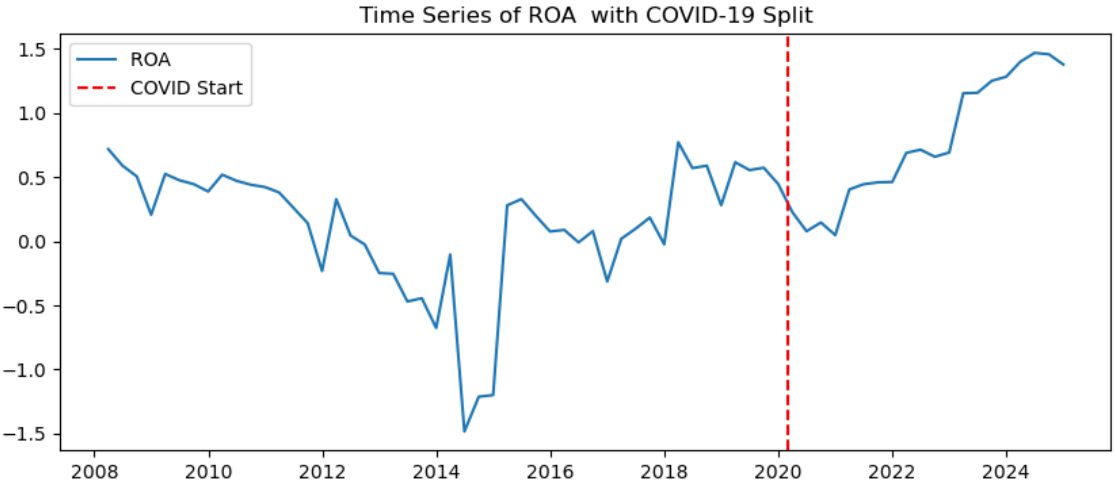


Figure 10 - Time series ROA with covid

Between 2008 and 2009, ROA evidenced a sharp fall, reflecting the global financial crisis's early impacts, losses and margin compression first noticeable in 2006. From 2010 and 2014, ROA went negative, which is partially due to the sovereign debt crisis around 2013–2014, as well as high NPLs and extensive balance sheet restructuring. From 2015 to 2017, shows an improved performance in asset quality that followed a little stability between 2018 and 2019, against a backdrop of low interest rates. The COVID-19 shock in 2020 temporarily reduced ROA: uncertainty and weaker credit demand. Though the partial recovery of economic activity in 2021-2022 brings some slight hope, another year at least will pass before banks can realistically expect to continue growing again based on real economic activity turned back up towards foreign markets once more by revitalisation of production.

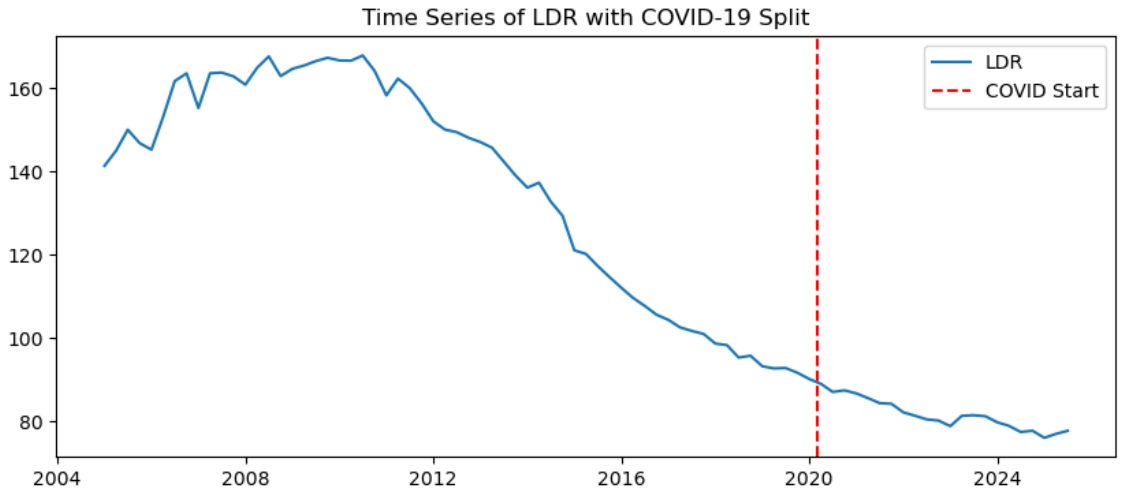


Figure 11 - Times series LDR with covid

From 2004 to 2011, the LDR fluctuated between 140% and 170%. After the financial crisis, the LDR continued to fall inexorably, reaching 89% at the end of 2019, thanks to banks' deleveraging and the strengthening of their deposit bases (Basel III). With the COVID-19 shock between 2020 and 2021, the LDR stabilised around 85-90%, with additional deposits offsetting the increased demand for credit, falling to between 75% and 80% from 2022 to 2025. This highlights caution in lending. The Portuguese banking sector, increasingly solid and less indebted, can absorb the shocks during the pandemic without increasing financial risks.

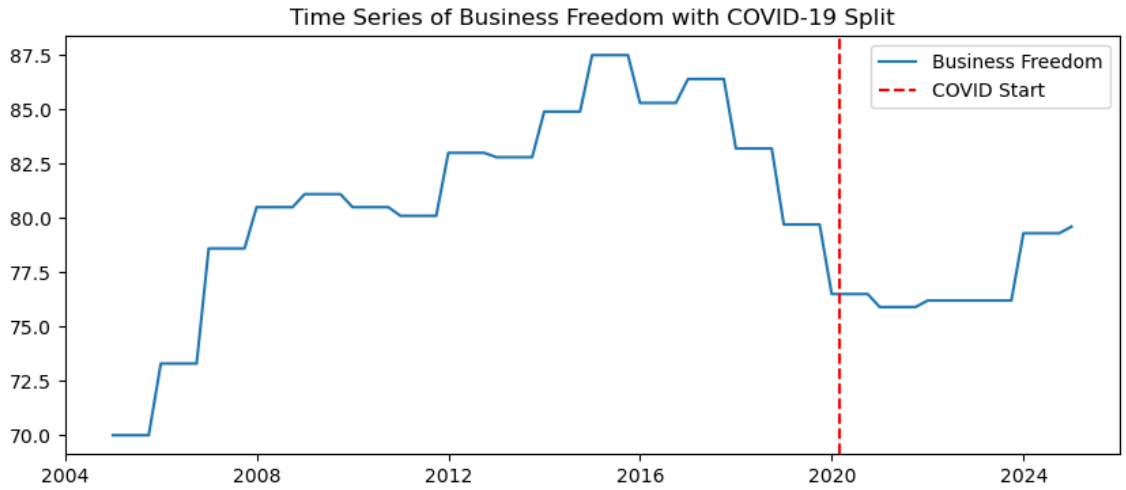


Figure 12 - Time series business freedom with covid

Figure 12 shows that Business Freedom rose steadily from 2004. It reached its peak in 2013–2016 with strong reform prospects and has now dropped slightly in 2017–2019. Nevertheless, 2020 showed an abrupt drop from the COVID-19 shock, followed by further stagnation during 2021 and 2022 before only a partial return in 2023 or 2024. The diminutive un-saw is, in fact, significant for NPL dynamics: a lessening of business freedom, especially against the background of the pandemic, probably lifted SME credit risks; while recovery might improve resistance and help to soften future loan stress.

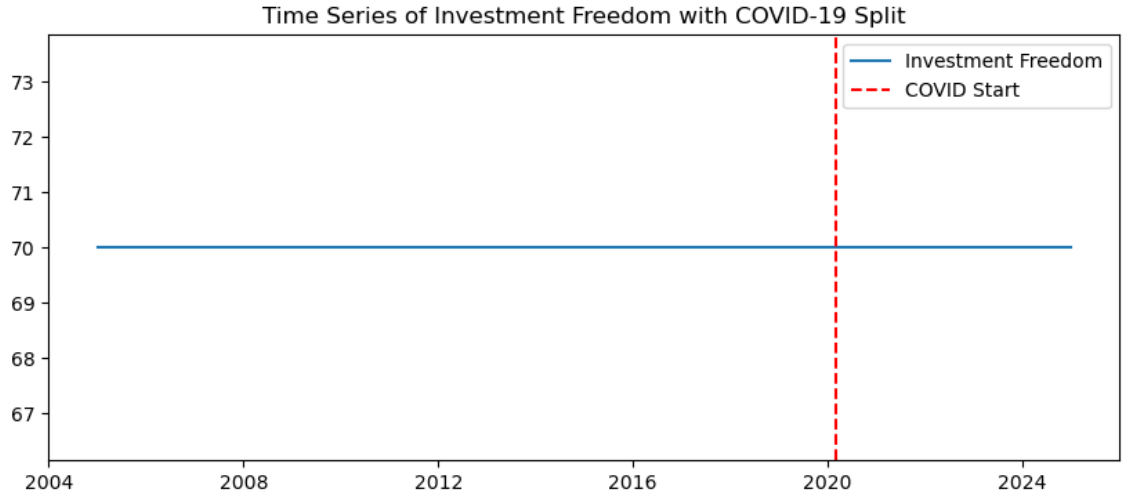


Figure 13 - Times series investment freedom with covid

The investment climate has a score of 70 for the period 2004 to 2024. EU membership, about the free flow of capital, did not, however, mean the complete elimination of restrictions. From the financial debt crisis to the recovery phase, the index

remained stable, highlighting the resilience of established institutions and the absence of concerted reforms to attract investment. This continuity means there were no major shocks to investment regulation, affecting credit dynamics or NPLs; however, the lack of adjustment may have constrained long-term growth.

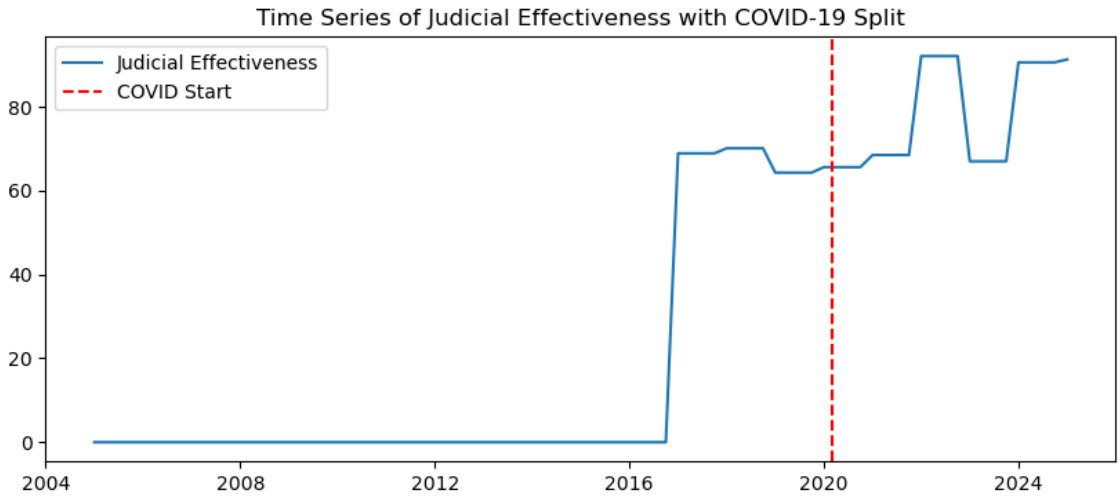


Figure 14 - Time series judicial effectiveness with covid

A structural break can be seen in the judicial effectiveness of Portugal, which was at zero from 2004-2016 - most likely due to omitted data rather than organisational weakness - but then rocketed to 68 in 2017, signifying reform or method changes. This stability continued until 2019, while the COVID-19 period in 2020 resulted in a few changes in scores. In 2021, the indicator became volatile with the trend hitting nearly 90 in 2022, dropping in 2023, and rebounding again in 2024. This reflects systemic pressures, clearing off backlogs or recalibrating indices. As far as NPL analysis is concerned, judicial effectiveness is of some relevance in relation to whether credit is enforced and loans are recovered - the break in 2017 suggests that this should be taken primarily as a post-2017 measure of institutional quality.

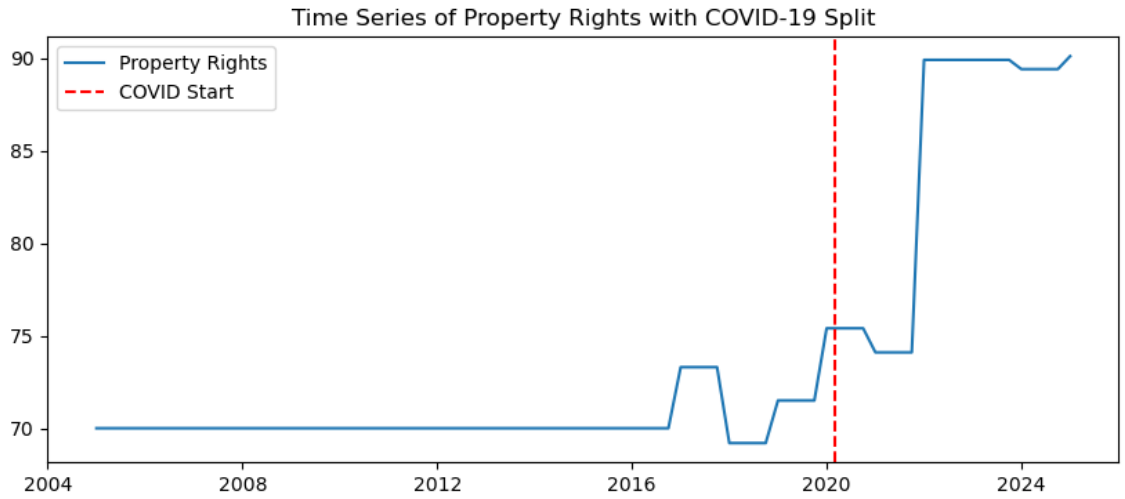


Figure 15 - Time series property rights with covid

In Portugal, property rights remained unchanged at around 70 from 2004 to 2016, marking not a radical institutional reform but a modest one, yet unreformed. In 2017, it entered on a gentle upward curve following modest fluctuation, even during the COVID-19 epidemic. The most significant reform-driven consolidation took place in 2021–2022. By then, the index had rocketed up to 90 and stayed there, indicating stricter enforcement, more precise legal provisions, and improved collateral security. These advances should prove especially useful for NPL dynamics: clearer property rights will foster secured lending, reassure investors, and make it easier to collect loans.

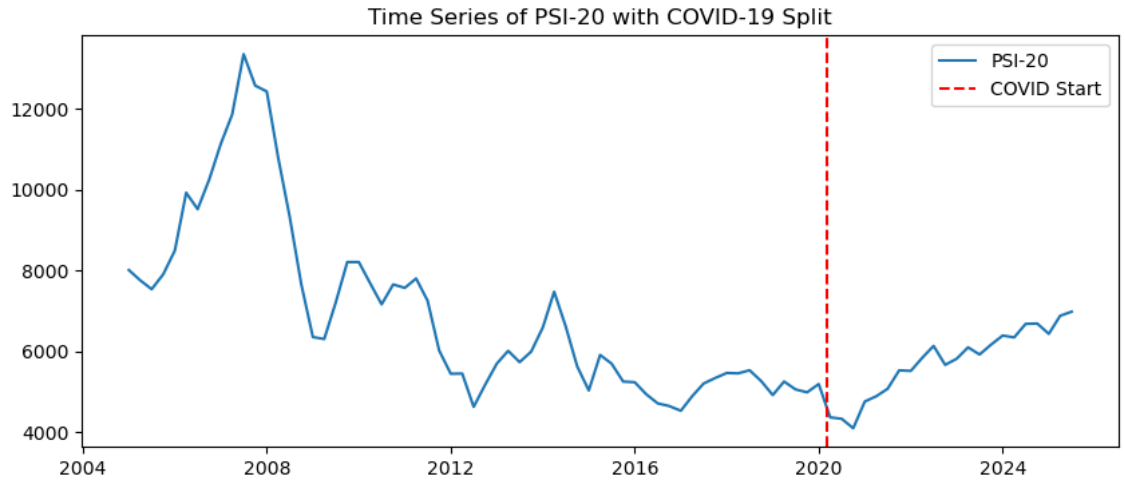


Figure 16 - Time series PSI-20 with covid

Regarding PSI-20, there was a pre-2008 boom followed by the 2008–2009 crash, with its aftermath of sovereign debt-driven volatility (2010–2014) and even a modest recovery (2015–2019); 2020 saw a dip caused by COVID-19, but from there it gradually rebounded post-pandemic to 6,800 in 2024. What these curves reflect, then, is investor sentiment, systemic shocks and economic stress. It is thus a handy proxy for grasping both NPL dynamics and the resilience of credit markets.

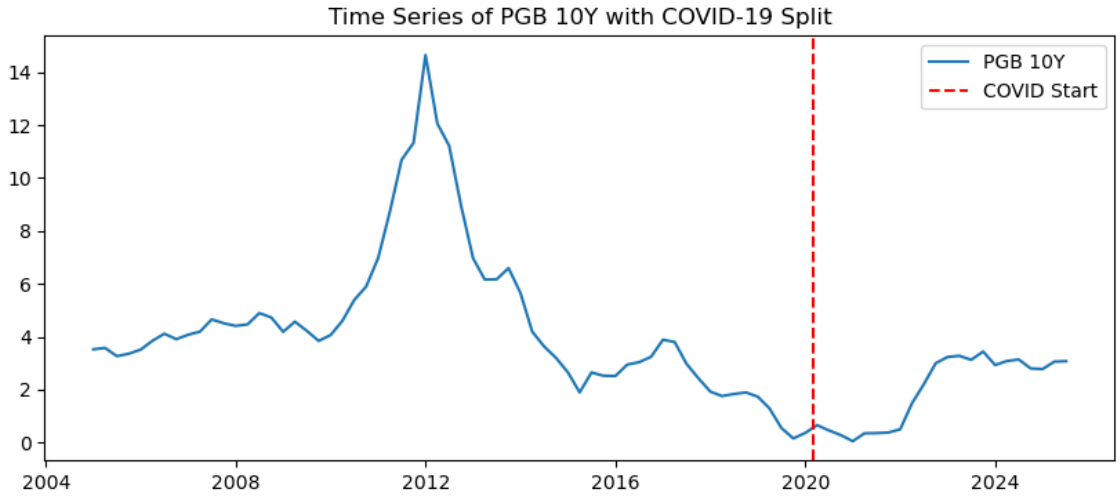


Figure 17 - Time series PGB 10Y with covid

The PGB 10y yield (2004–2024) reflects sovereign risk, market sentiment, and macroeconomic stress. Yields were low and stable (3-5%) pre-2009, rose to ~15% for 2010–2012 during the Eurozone debt crisis, then dropped below 3% post-bailout as fiscal management and ECB buying regained investor confidence. They remained low in 2016-2019, rose slightly for the 2020 COVID shock, but gradually increased to ~3.5% post-2021 global monetary tightening.

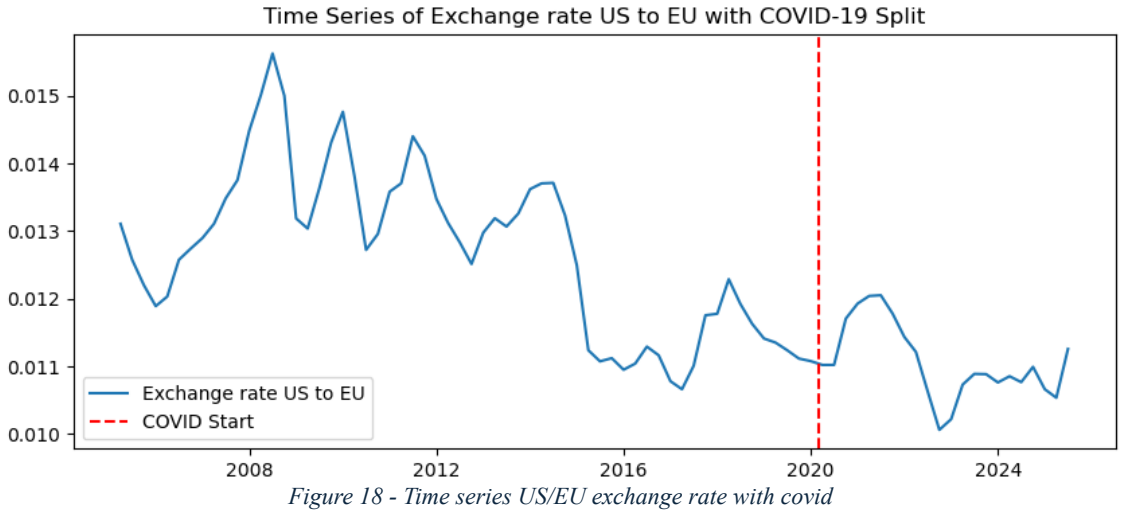


Figure 18 - Time series US/EU exchange rate with covid

USD/EUR during 2004–2008, the USD/EUR exchange rate, the euro rallied both because of America's weaknesses and the firm stance taken by the eurozone. From 2008 to 2010 crisis, investors went from dollars to euros in search of safety, but following a change in US monetary policy conditions stimulated the euro's recovery. From 2011 to 2015, the euro fell as everything turned tough with eurozone debt and US growth became stronger, stabilising moderately in 2016–2019. It rose slightly amid the COVID-19 shock during 2020-2021, but post-2021, the dollar is regaining strength. This he attributes to tighter US monetary conditions and global uncertainty. These movements on import prices in Portugal, eurozone policy transmission, credit conditions and NPL.

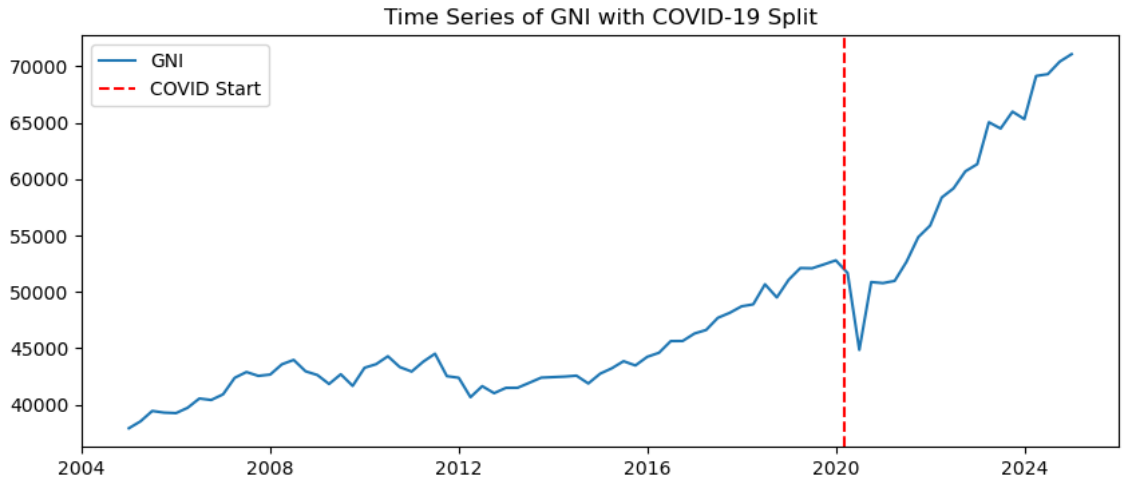


Figure 19 - Time series of GNI with covid

GNI measures the sum of all residents' incomes, or national economic performance and productivity. Between 2004 and 2024, GNI shows long-term growth punctuated by crises: steady pre-2008 expansion, halt during 2009 - 2013 financial and sovereign crises, recovery as of 2019, a sharp COVID-19 downturn in 2020, and strong post-pandemic rebound by 2024. These fluctuations affect household and business debt capacity, as a main indicator for NPL risk and overall economic resilience.

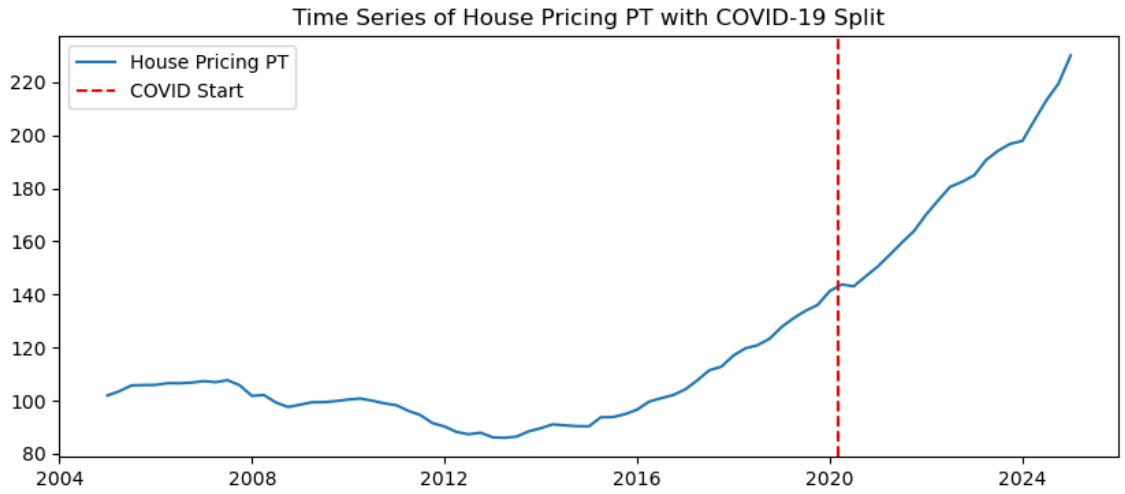


Figure 20 - Time series house pricing with covid

During 2004-2008, house prices climbed moderately as credit conditions held steady and the EU loans rolled in. From 2009 to 2013, the HPI dropped rapidly on the back of financial and sovereign debt crises that cast economies into recession, ushered in austerity and curtailed demand for housing. The Portuguese housing price index then climbed slowly over 2014-2019, it's possibly driven by tourism, foreign investment, urban regeneration and a tighter labour market. In 2020, COVID-19 nudged inflation down a bit with prices stabilising, while 2021 - 2024 saw a strong post-pandemic boom buoyed by low interest rates, supply constraints and even continued foreign demand. This saw wide swings in the value of collateral, at times causing shifts in mortgage default risk and NPL post-COVID growth, then pushing up possible implications of overvaluation concerns because of rapid increases in potential problems.

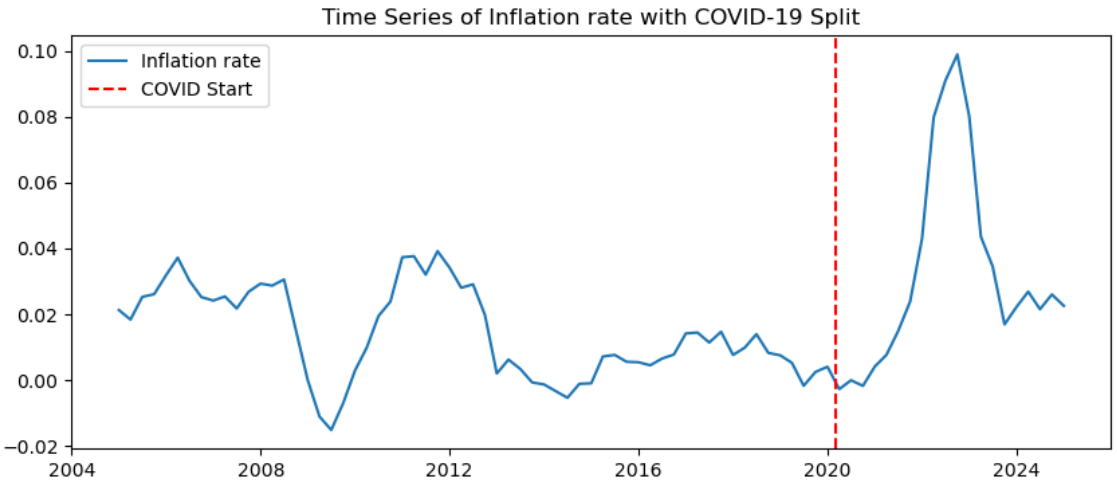


Figure 22 - Time series inflation rate with covid

Between 2004 and 2024, Portugal's inflation rate underwent distinct stages as macroeconomic conditions and policy changed. Mild inflation continued throughout 2004 - 2008, followed by a period of disinflation during the 2009 - 2013 financial and sovereign debt crises. Inflation ran low and stable from 2014 to 2019, but collapsed in 2020 due to COVID-19. A spike then occurred during 2021 - 2022, driven by post-pandemic demand, supply shocks and energy price surge before easing off in 2023 - 2024 with ECB tightening. This affected real debt burden, interest rate conditions and NPL deflation raised repayment pressure, and the 2022 spike briefly jacked up interest rates and solvency risks overall.

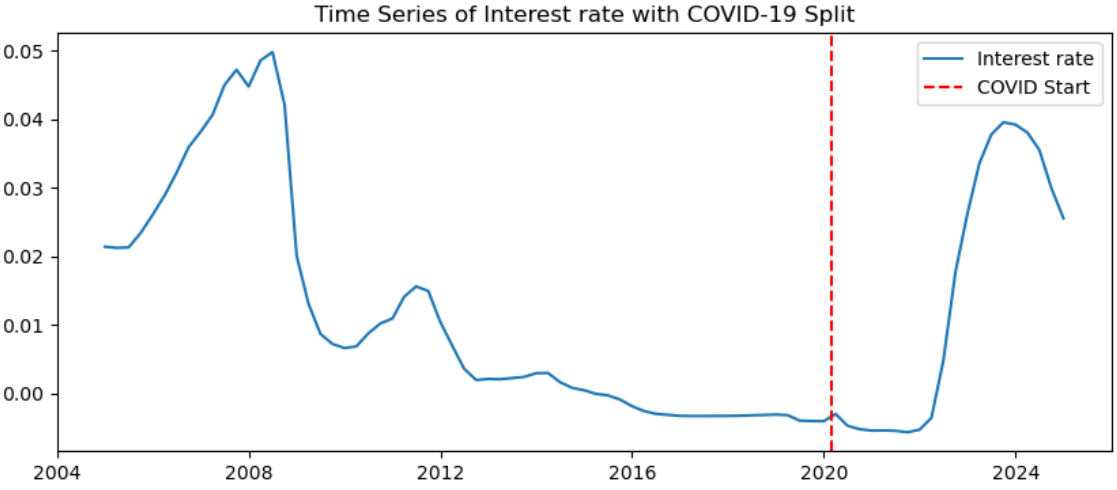


Figure 21 - Time series interest rates with covid

Between 2004 and 2024, Portuguese interest rates, largely aligned with the ECB policy trajectory, underwent distinct phases: pre-crisis tightening (2004 - 2008) with rates rising to contain inflation and credit growth; post-crisis easing (2009 - 2013) as rates fell to stimulate credit and support sovereign financing during the global financial and eurozone crises; a prolonged near-zero period (2014 - 2020) reflecting the zero-lower bound and ultra-loose monetary policies to sustain investment, credit flows, and reduce NPL risk; and a sharp post-COVID tightening (2022 - 2024) with rates climbing to counter record inflation. These cycles affected borrowing costs and credit conditions, with low rates easing repayment and reducing NPLs, while rapid hikes increased payment stress and credit risk for households and SMEs.

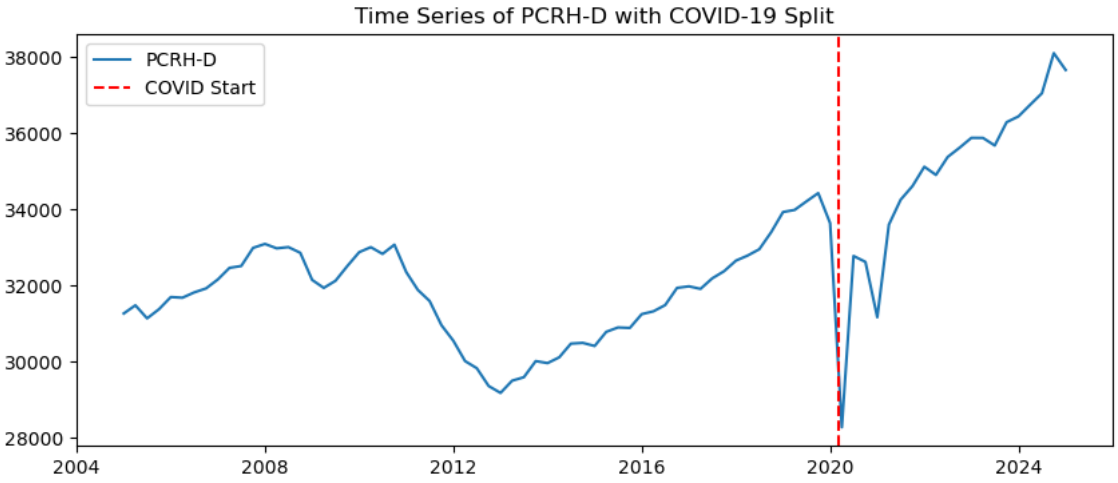


Figure 23 - Time series PCRH-D with covid

Between 2004 and 2024, PPCRH-D followed regular cycles tied to confidence, credit and macroeconomic conditions. Before 2008, consumption increased steadily amid rising incomes and cheap loans. During the 2009 - 2013 crisis period of austerity, it declined sharply. Then, for a period from 2014 to 2019, as wage growth picked up and interest rates remained low, consumption slowly recovered. The 2020 COVID shock led to another collapse in consumption, which was only partially made good by pent-up demands for goods and services. These affected NPL: a fall in consumption might portend rising defaults, while credit-driven booms may cover over fiscal vulnerabilities.

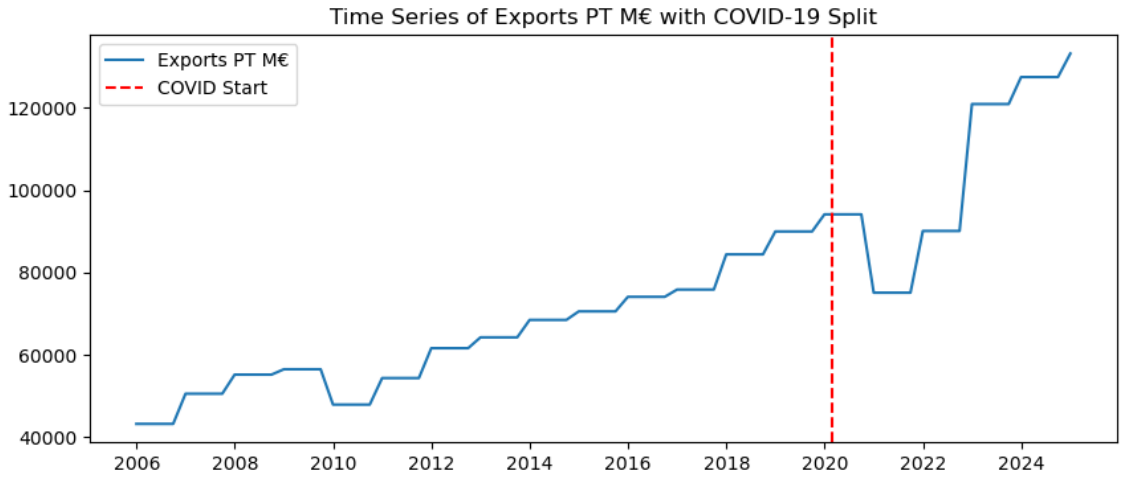


Figure 24 - Time series exports with covid

From 2006 to 2024, the Portuguese export cycles depicted in Figure 24 were formed by demand-driven and internal restructuring. Mild growth in the pre-2009 era was interrupted by the financial crisis. Then, after gains in 2011 - 2014 under austerity, it went steady between 2015 and 2019 through diversification and global integration. The COVID-19 snap in 2020 shattered those numbers as tourism exports dropped dramatically before picking up again after COVID-19 to reach an all-time high by 2024. Yet these export dynamics have a direct impact on the solvency of enterprises and household income, affecting their ability to repay debts or their NPL risks.

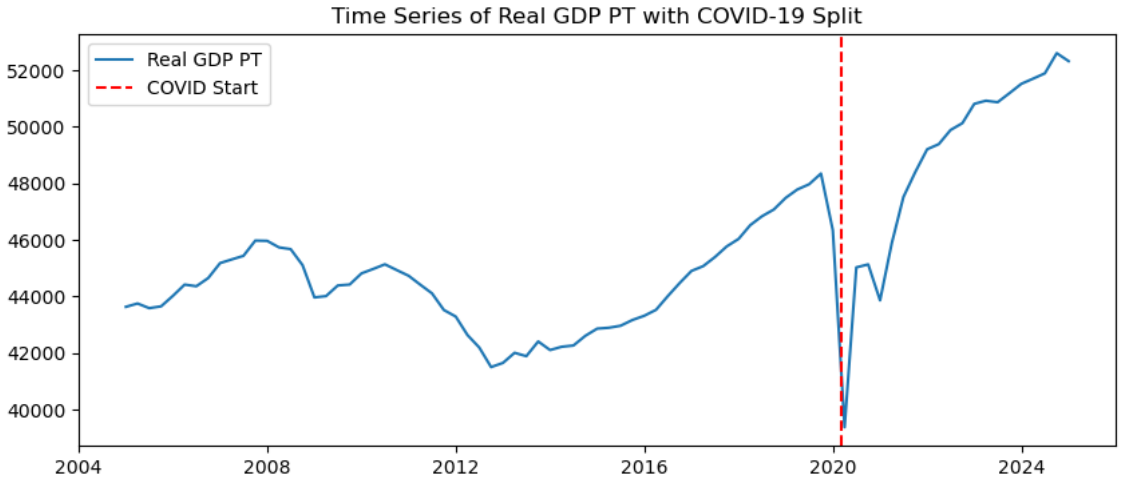


Figure 25 - Time series real GDP with covid

From 2004 to 2024, the real GDP performance of Portugal experienced cycles shaped by crises, structural adjustments, and recovery following the epidemic crisis. Moderate growth between 2004 and 2008 reflected euro area integration: expansion of credit; big reform pushes somewhere along the line made things a little brighter. The global financial shock of 2009 and sovereign debt austerity from 2011 to 2013 led to a deep double-dip recession with mass unemployment, as well as weak domestic demand. Recovery by 2014 - 2019 was continuous but mild economically, driven by exports, capital formation and tourism. In 2020, the year of the COVID-19 pandemic in Portugal, tourism-oriented exports collapsed, leaving a severe downturn in their place, which can at best be described as transient. But after COVID-19 proved a strong V-shaped turnaround from 2021 – 2024, outperforming pre-pandemic levels.

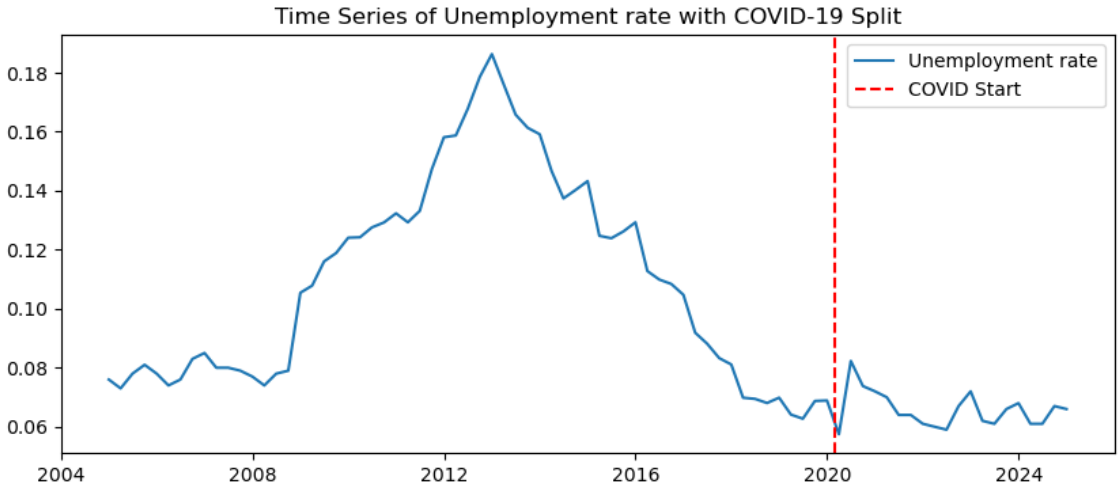


Figure 26 - Time series unemployment rate with covid

From the time 2004 - 2024, it is marked that Portugal's unemployment rate reflects the sensitivity of labour markets to economic cycles and external shocks. Before 2008, unemployment remained low and stable. In 2009 - 2011, however, it surged sharply to above 18%. From this peak, unemployment figures are on the decline until 2013 after a gradual decrease which lasts for five consecutive years; finally, by 2019, they are just below 7%. The 2020 COVID-19 shock caused a minor temporary upturn. Policy support helps to cushion this without letting unemployment increase much, and then rapid post-pandemic recovery brings it down again. By 2024, unemployment will be historically low.

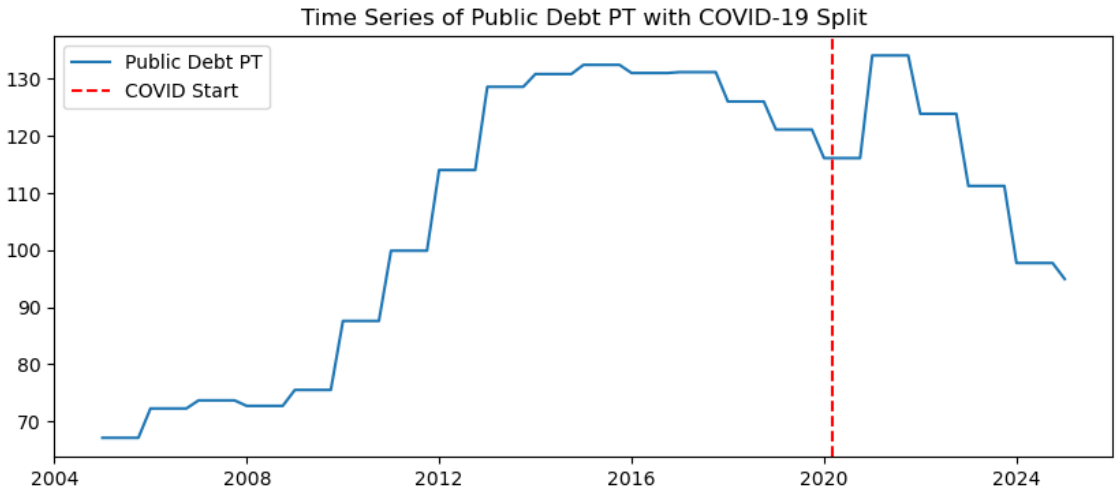


Figure 27 - Time series public debt with covid

Portugal's sovereign and financial crises of 2008 - 2013 greatly increased public debt within the space of just a few short years. In 2019, the fiscal consolidation and economic recovery required the combined contributions of all social sectors. Contrary to expectations produced by this gradual improvement, the shock of COVID-19 in 2020 did not lead to a significant increase in public debt; on the contrary, with support from EU financial aid and targeted fiscal measures to a certain degree, GDP rebounds temporarily. From 2021 to 2024, the debt-to-GDP ratio kept falling as rapid economic recovery continued and fiscal balances improved. This helped ensure the security and resilience of post-pandemic stabilisation. The role of institutional support and macroeconomic resilience in helping to implement this policy is fully illustrated by these data.

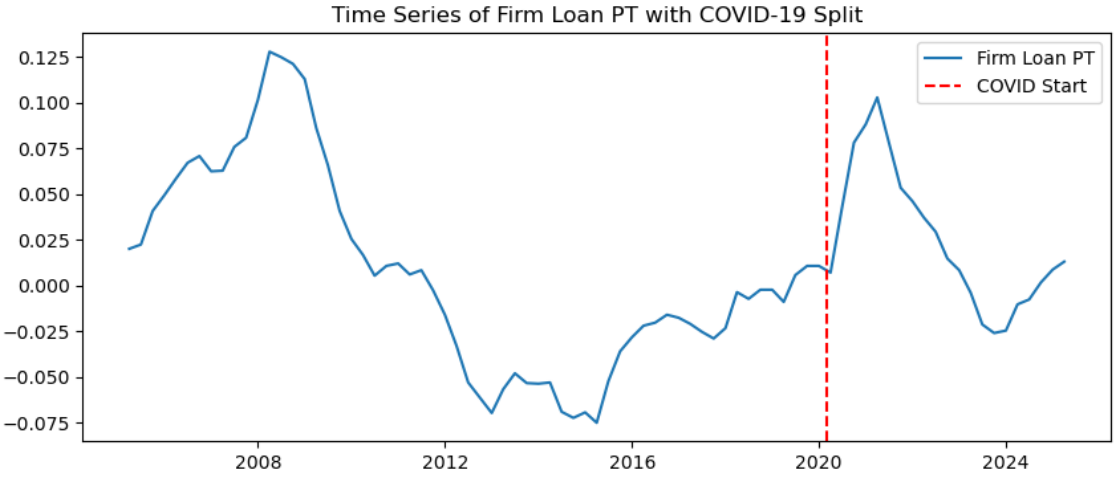


Figure 28 - Time series firm loan with covid

During the pre-crisis credit expansion, between 2006 and 2009, firm loans rose sharply, but then during 2010 - 2012 they all dropped off in an echo of Europe's debt crisis and austerity policies. From 2014 on, however, lending stayed at the lower level that it is today. In 2020, loan volume rose strongly because of the COVID risk governments *"backstopped"* credit and liquidity support but then declined sharply over 2021 - 2022 as both provisions ceased and credit conditions tightened. From 2023 to 2024, growth continued to be weak, with only a small recovery.

B2) Correlation matrix heatmap for all variables

A correlation matrix heatmap can present variable relationships, association direction and degrees of strength by means of colour processing. The correlation matrix integrates both bank-specific indicators and macroeconomic variables, as well as environmental factors such as judicial effectiveness, property rights, and deficit caps of governments. The colour gradient shows how long each pair is a linear relation between two things robustly correlated: red means a high positive correlation, blue means a high negative correlation. Some notable patterns emerge, such as strong links between macroeconomic variables and financial stability indicators, as well as meaningful links between LDR, public debt, and both bank performance and systemic liquidity, emphasising that banks' performance (and systemic liquidity) are in fact systemic as they are brought about by broader economic conditions on which wider society depends.

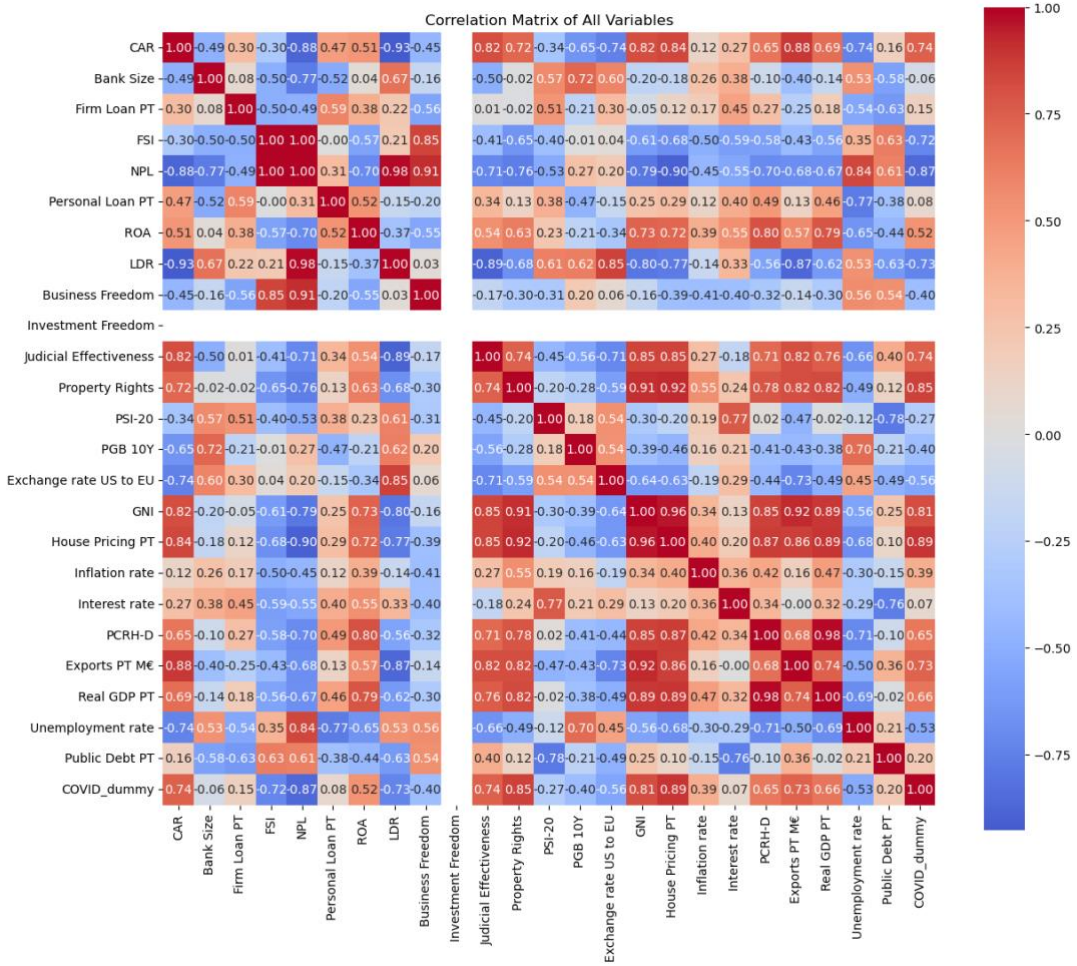


Figure 29 - Correlation matrix of all variables

The CAR has a high positive correlation with GNI (0.82), Judicial Effectiveness (0.88), Exports (0.88), House Prices (0.84) and PCRH-D (0.65), which is indicative of the linkage with high levels of capital buffers and better domestic economic conditions. Negative association with NPL (-0.88), LDR (-0.93), Bank Size (-0.49), and the US/EUR exchange rate (-0.74) as well as Unemployment (-0.74) imply that larger banks maintain relatively weak capital ratios where the capital adequacy ratio (CAR) decreases in the event of increased unemployment or depreciation of the euro. These results suggest that positive macroeconomic conditions (reflected in increasing income, consumption, exports and property prices) favour the stability and strength of banks by diminishing the credit risk and increasing the bank profitability. ROA is positively related to House Prices (0.72), GNI (0.73), PCRH-D (0.80) and Real GDP (0.79) and negatively related to NPL (-0.70) and Unemployment (-0.65), indicating that ROA is higher in a favourable economic situation and lower when the credit risk is higher. Bank Size has a weak sign of association with ROA (0.04), which suggests that larger Portuguese banks aren't necessarily more profitable.

NPLs have strong positive correlations with LDR (0.98), Unemployment (0.84), and Business Freedom (0.91), and strong negative correlations with House Prices (-0.90), GNI (-0.79), PCRH-D (-0.70), and ROA (-0.70), and are also perfectly correlated with FSI, as expected. Such relationships illustrate that relatively high NPLs are highly associated with lower economic conditions, lower household and firm spending, and lower bank profitability, yet are also highly sensitive to the overall credit quality. LDR is highly negatively correlated with CAR (-0.93), Judicial Effectiveness (-0.89), Exports (-0.87), GNI (-0.80), and House Prices (-0.77), and negatively correlated with Property Rights (-0.68) and ROA (-0.37), while it is positively correlated with PGB 10y yields (0.85) and Bank Size (0.67). Thus, larger banks, less economically and institutionally developed conditions, and higher credit risk exposure seem to be associated with higher LDR, implying that riskier lending practices, even though they contribute only moderately to the ROA, can increase vulnerabilities.

Public Debt exhibits both positive and negative correlations with financial and institutional indicators. Positive correlations with FSI (0.63) and NPLs (0.61) suggest that high-leveraged companies often experience more credit risk. It is worth noting that

moderate positive associations with Business Freedom (0.54) and Judicial Effectiveness (0.40) indicate that times of high debt also tend to coincide with better institutional backing. Negative associations with Bank Size (-0.58), LDR (-0.63), Firm Loans (-0.63), the PSI-20 (-0.78), and Interest Rates (-0.76) suggest that increasing debt can come hand in hand with tighter bank balance sheets, lower borrowing and market valuations, and cheaper cost of borrowing. These results suggest aggregate across policy, finance and macro categories, whereby Public Debt can simultaneously be conducive to growth and a source of systemic risk.

Most macroeconomic variables present the expected results. House Prices has significant pure trend positive relationships with PCRH-D (0.87), Real GDP (0.89), CAR (0.84), ROA (0.72), Judicial Effectiveness (0.85), Property Rights (0.92), Exports (0.86) and GNI (0.96), and dose trend negative relationships with NPLs (-0.90), LDR (-0.77), FSI (-0.68) and Unemployment (-0.68), a testament to the pivotal role of the housing market in economic and financial stability. PCRH-D and GDP also present strong positive regression coefficients with institutional indicators and negative linkages with FSI, confirming again the relevance of macroeconomic soundness on financial stability.

Depreciation of the exchange rate is negatively related with CAR (-0.74), House Prices (-0.63), GNI (-0.64), Judicial Effectiveness (-0.71), Exports (-0.73), PCRH-D (-0.44), and positively related with Unemployment (0.45) and tracks the PSI-20 (0.54), PGB 10y yields (0.54) and the LDR (0.85), indicating that domestic conditions and banking system fragility are responsive to currency movements. Interest rates moderate positive relationships are reported with ROA (0.55), Firm Loans (0.45) and PSI-20 (0.77), whereas negative correlations are obtained with Public Debt (-0.76), FSI (-0.59), and NPLs (-0.55), representing how they affect borrowing costs and credit risk.

Institutional indicators, such as Judicial Effectiveness and Property Rights, display strong positive correlations with Solvency, ROA, and macroeconomic performance, and negative correlations with NPLs and FSI, pointing to their stabilising influence in warfare, the banking industry and the wider economy. The dummy of COVID-19 has negative correlations with Unemployment (-0.53), NPLs (-0.87) and FSI (-0.72) and positives correlations with GDP (0.66), PCRH-D (0.65), House Prices (0.89),

Exports (0.73), GNI (0.81) and Property Rights (0.85), so, even with the pandemic shock the Portuguese economy showed a partial resilience due strong institutions and economic activity still in place.

This analysis evidences that the Portuguese banking sector performance is highly influenced by macroeconomic, institutional, financial markets, and fiscal policy factors. CAR and ROA increase with GNI, GDP, consumption, house prices, and exports and NPLs and LDR are enhanced with poor economic circumstances, higher unemployment, and increased credit risk. Institutional drivers, in particular Judicial Effectiveness and Property Rights, are stabilisers to economic fluctuations, and the COVID-19 period highlights the value of institutional resilience in dampening shocks. Public debt is positively associated with growth as well, but at the same time leads to higher NPLs and credit risk, highlighting the double role of the fiscal policy. The depreciation of the exchange rate is consistent with weaker macro-financial conditions, and interest rates are of weak significance for profitability, lending, and markets. The LDR captures the balance between lending expansion and lending stability. Taken as a whole, such results stress the mutual links among macroeconomic, financial, and institutional fundamentals in ensuring the stability of the Portuguese banking sector.

B3) Scatter Plot Analysis of NPL Determinants

This representation allows the visualisation of the relationships between NPLs and their determinants, showing the strength and sign of the associations emerged by the correlation matrix, revealing patterns, non-linearities or outliers that can drive the dynamics of loan quality.

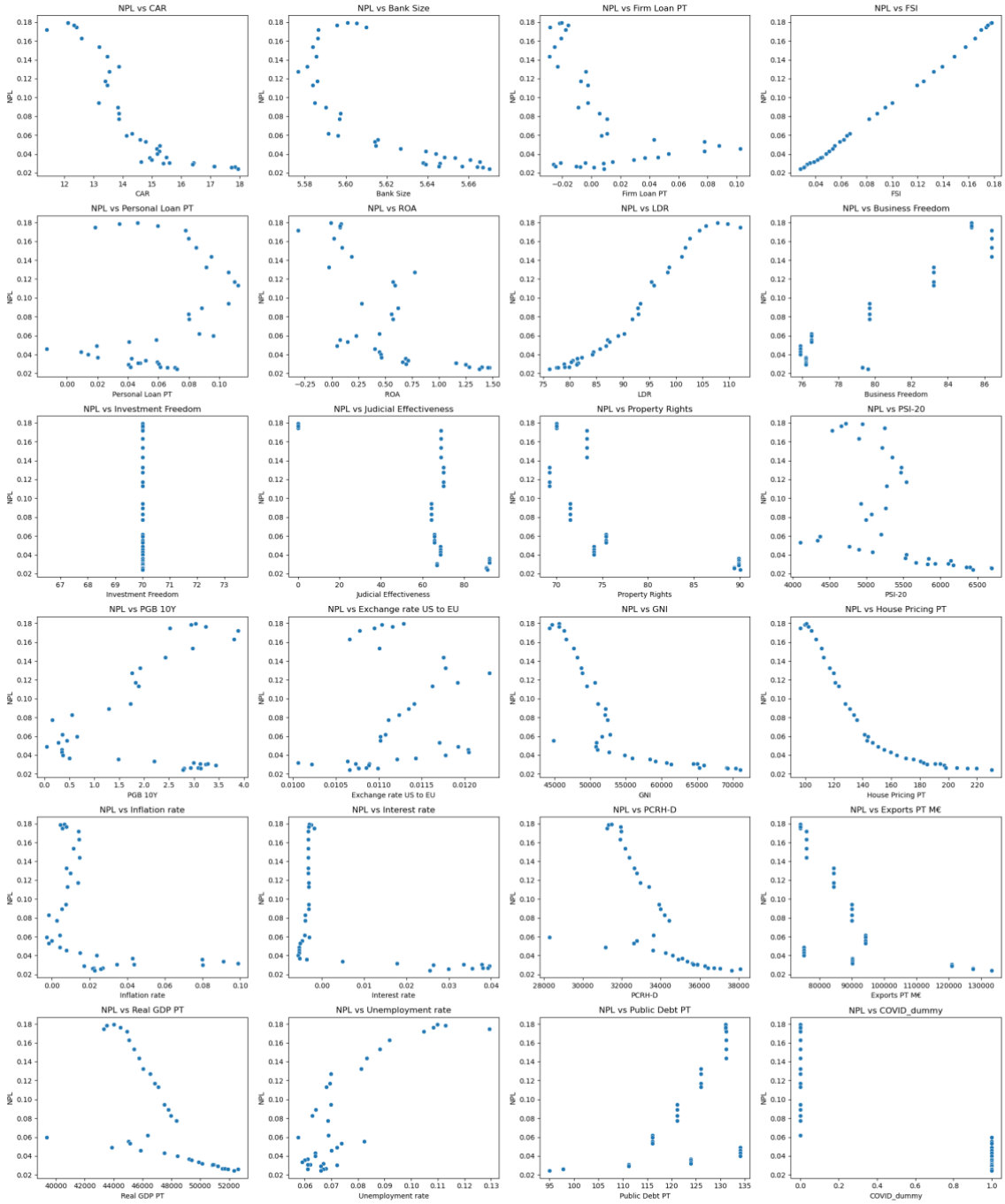


Figure 30 - Scatter plot of NPL determinants

Bank-specific factors reveal distinct patterns. Meanwhile, a negative and significant relationship is becoming apparent between CAR and NPLs, which means that banks with sufficient capital are in a stronger position to control credit risk. On the other hand, the size of the bank has very little explanatory power, which means that the bigger banks do not enjoy lower levels of loan defaults by their customers. There is a positive association between both firm and personal lending and NPLs since a higher credit risk arises from larger exposure, although the LDR depicts an ascending curve, indicating the riskiness derived from more aggressive lending policies. Not surprisingly, credit risk, proxied through FSI, has a very high correlation with NPLs as both capture the credit risk dimension.

Institutional quality related to business in terms of business freedom, investment freedom, judicial effectiveness, and property rights does not show clear linear relationships with NPLs. Instead, they exhibit clustered patterns, indicating that the institutional measures have a second-order association with loan quality, which may reach indirectly through more structural or long-term channels. The market signals sound a more immediate note. The positive PSI-20 stock market performance is associated with both lower NPLs, aligning with the notion of financial stability spillovers. By contrast, the increasing PGB 10y data corresponds to higher NPLs, and the uptrend reflects sovereign and banking sector risk perceptions.

Macroeconomic variables largely behave as expected. An increase in income (output) measured by GNI (real GDP) shows a negative association with NPLs, indicating the significance of economic growth on the ability to repay debt. Housing prices also show a negative, substantial correlation, which reveals an anchoring role of collateral values. On the other hand, unemployment is substantially positively correlated with credit distress, indicating its pivotal role in the causal structure. Interest rates have a positive relationship, suggesting that the higher is debt servicing than the more rollovers are difficult. Public debt also correlates positively with NPLs, associating sovereign weaknesses with the fragility in the banking sector. Unlike PCRH-D, which decreases NPLs from advancing domestic demand, exports are more neutral, with no discernible

pattern of economic response. Inflation has a significantly low relationship with some indication of higher bad credit risk as inflation rises.

Finally, the COVID-19 dummy suggests a structural break, having the NPL tendency to move up in the pandemic. This is an indication of the massive disturbance to economic activity and loan servicing, despite the temporary moratoria in place. In a nutshell, the most powerful modelling drivers for NPLs are unemployment, GDP/GNI, house prices, bond yields, and credit-at-risk variables that directly mirror macro financial soundness and the ability to repay. Bank-specific resilience is positively associated with CAR and prudential lending but minimally related to institutional quality. The period of the pandemic is identified as a key moment characterised by an increase in credit risk, which confirms the sensitivity of NPLs to both business cycle and systemic shocks.

B4) Variables correlated pairs

The correlation analysis is essential for this research, pinpointing those macro-financial and institutional determinants of banking sector development in Portugal. Analysing the pairwise associations facilitate the detection of important connections between solvency, credit risk, returns, and lending activity, highlighting the likely multicollinearity problems to support econometric modelling and undertake the minimum number of variables.

The CAR is strongly negatively correlated with the NPL (-0.88) and LDR (-0.93); this pattern confirms that weak asset quality and aggressive lending eroded the capital buffer. CAR has a positive relation with Judicial Effectiveness (0.82), Property Rights (0.72), GNI (0.82), Exports PT (0.88), House Pricing PT (0.84) and Real GDP PT (0.69), which indicates that strong institutions and macroeconomic performance have stabilising effects. Negative correlations with PGB 10y yield (-0.65), the Exchange Rate US/EUR (-0.74), and the Unemployment Rate (-0.74) make it clear that the vulnerabilities to sovereign, currency, and labour market pressures are captured here, whereas the COVID dummy (0.74) picks up a structural change improvement in solvency post the pandemic.

The NPL ratio is very sensitive to macroeconomic and institutional variable values, and it associated positively to Unemployment Rate (0.84) and Public Debt PT (0.61), negatively to House Pricing PT (-0.90), GNI (-0.79), PCRH-D (-0.70), Exports PT (-0.68), Real GDP PT (-0.67), Judicial Effectiveness (-0.71), and Property Rights (-0.76). NPLs are completely positively correlated with FSI (1.00). Bank Size (-0.77) is negatively related to NPLs, and Business Freedom (0.91) is positively related, which implies that less restrictive regulation environments might increase the cost of credit.

ROA becomes more profitable in an economic boom and is significantly positively associated with PCRH-D (0.80), Real GDP PT (0.79), GNI (0.73), House Pricing PT (0.72) and Property Rights (0.63), and significantly negatively associated with NPL (-0.70) and Unemployment Rate (-0.65).

The LDR seems to be a major driving variable for the vulnerability of the banking sector. It is strongly negatively correlated with CAR (-0.93), Judicial Effectiveness (-

0.89), Property Rights (-0.68), GNI (–0.80), Exports PT (-0.87), House Pricing PT (-0.77), Real GDP PT (-0.62), and COVID dummy (-0.73) and strongly positively correlated with NPL (0.98). Besides, the positive correlation to Bank Size (0.67), PGB 10y (0.62), PSI-20 (0.61) and Exchange Rate US/EUR (0.85), suggests that there are other general market and currency impacts on lending behaviour.

The macroeconomic and institutional dimensions are highly interconnected. GNI and House Pricing PT (0.96), the PCRH-D and Real GDP PT (0.98), and Exports PT (0.92) - at which House Pricing PT (0.86) display the strong relationships between household demand, asset markets, and economic production. Judicial Effectiveness is positively related to House Pricing PT (0.85), GNI (0.85), PCRH-D (0.71), Exports PT (0.82), and Real GDP PT (0.76), whereas Property Rights are highly related to House Pricing PT (0.92), GNI (0.91), PCRH-D (0.78), and Real GDP PT (0.82). Market indicators such as the PGB 10y yield, the Exchange Rate US/EUR, and the Interest Rates continue to be linked to macro-financial stress.

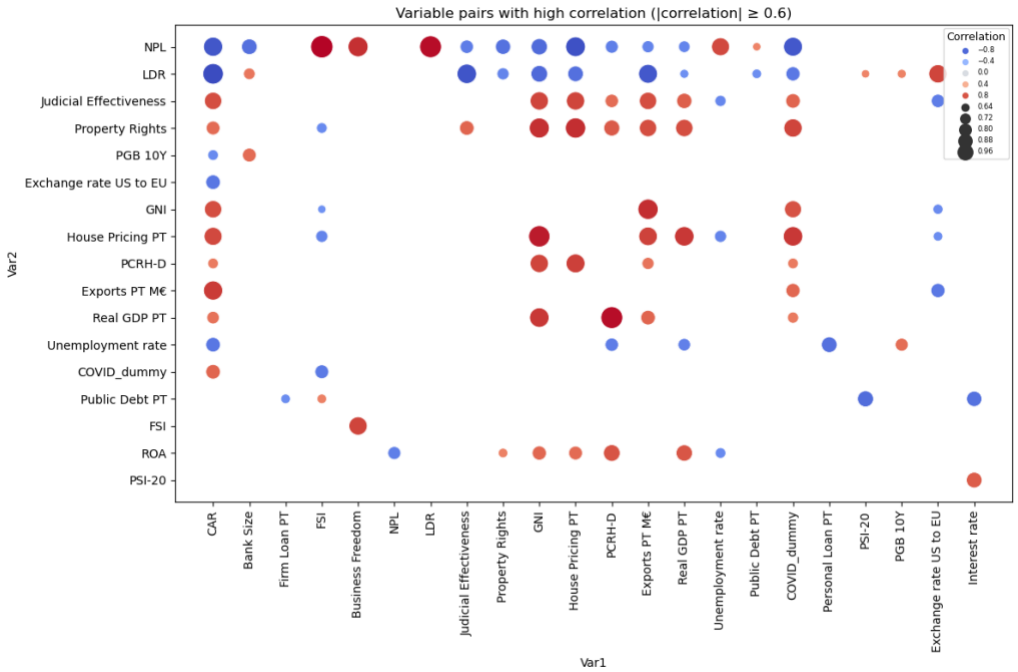


Figure 31 - Variable pairs with high correlation

In addition, considering the figure above, it is also possible to see that the strongest correlations cluster around a few critical pairs: NPL and FSI (1.00), NPL and LDR (0.98), Real GDP PT and PCRH-D (0.98), GNI and House Pricing PT (0.96), and Property Rights and House Pricing PT (0.92). On the risk side, NPL and House Pricing PT (-0.90) and

CAR and LDR (-0.93) represent the most significant negative associations. These findings demonstrate that credit quality, solvency, and profitability are most sensitive to the interplay of household consumption, property markets, and institutional conditions, while excessive lending and macroeconomic stress constitute the main channels of banking sector vulnerability.

Variable highest correlation pairs

```
CAR ↔ NPL : correlation = -0.88
CAR ↔ LDR : correlation = -0.93
CAR ↔ Judicial Effectiveness : correlation = 0.82
CAR ↔ Property Rights : correlation = 0.72
CAR ↔ PGB 10Y : correlation = -0.65
CAR ↔ Exchange rate US to EU : correlation = -0.74
CAR ↔ GNI : correlation = 0.82
CAR ↔ House Pricing PT : correlation = 0.84
CAR ↔ PCRH-D : correlation = 0.65
CAR ↔ Exports PT M€ : correlation = 0.88
CAR ↔ Real GDP PT : correlation = 0.69
CAR ↔ Unemployment rate : correlation = -0.74
CAR ↔ COVID_dummy : correlation = 0.74
Bank Size ↔ NPL : correlation = -0.77
Bank Size ↔ LDR : correlation = 0.67
Bank Size ↔ PGB 10Y : correlation = 0.72
Firm Loan PT ↔ Public Debt PT : correlation = -0.63
FSI ↔ NPL : correlation = 1.00
Business Freedom ↔ FSI : correlation = 0.85
FSI ↔ Property Rights : correlation = -0.65
FSI ↔ GNI : correlation = -0.61
FSI ↔ House Pricing PT : correlation = -0.68
FSI ↔ Public Debt PT : correlation = 0.63
COVID_dummy ↔ FSI : correlation = -0.72
NPL ↔ ROA  : correlation = -0.70
LDR ↔ NPL : correlation = 0.98
Business Freedom ↔ NPL : correlation = 0.91
Judicial Effectiveness ↔ NPL : correlation = -0.71
NPL ↔ Property Rights : correlation = -0.76
GNI ↔ NPL : correlation = -0.79
House Pricing PT ↔ NPL : correlation = -0.90
NPL ↔ PCRH-D : correlation = -0.70
Exports PT M€ ↔ NPL : correlation = -0.68
NPL ↔ Real GDP PT : correlation = -0.67
NPL ↔ Unemployment rate : correlation = 0.84
NPL ↔ Public Debt PT : correlation = 0.61
COVID_dummy ↔ NPL : correlation = -0.87
Personal Loan PT ↔ Unemployment rate : correlation = -0.77
Property Rights ↔ ROA  : correlation = 0.63
GNI ↔ ROA  : correlation = 0.73
House Pricing PT ↔ ROA  : correlation = 0.72
PCRH-D ↔ ROA  : correlation = 0.80
ROA  ↔ Real GDP PT : correlation = 0.79
ROA  ↔ Unemployment rate : correlation = -0.65
```

```
Judicial Effectiveness ↔ LDR : correlation = -0.89
LDR ↔ Property Rights : correlation = -0.68
LDR ↔ PSI-20 : correlation = 0.61
LDR ↔ PGB 10Y : correlation = 0.62
Exchange rate US to EU ↔ LDR : correlation = 0.85
GNI ↔ LDR : correlation = -0.80
House Pricing PT ↔ LDR : correlation = -0.77
Exports PT M€ ↔ LDR : correlation = -0.87
LDR ↔ Real GDP PT : correlation = -0.62
LDR ↔ Public Debt PT : correlation = -0.63
COVID_dummy ↔ LDR : correlation = -0.73
Judicial Effectiveness ↔ Property Rights : correlation = 0.74
Exchange rate US to EU ↔ Judicial Effectiveness : correlation = -0.
71
GNI ↔ Judicial Effectiveness : correlation = 0.85
House Pricing PT ↔ Judicial Effectiveness : correlation = 0.85
Judicial Effectiveness ↔ PCRH-D : correlation = 0.71
Exports PT M€ ↔ Judicial Effectiveness : correlation = 0.82
Judicial Effectiveness ↔ Real GDP PT : correlation = 0.76
Judicial Effectiveness ↔ Unemployment rate : correlation = -0.66
COVID_dummy ↔ Judicial Effectiveness : correlation = 0.74
GNI ↔ Property Rights : correlation = 0.91
House Pricing PT ↔ Property Rights : correlation = 0.92
PCRH-D ↔ Property Rights : correlation = 0.78
Exports PT M€ ↔ Property Rights : correlation = 0.82
Property Rights ↔ Real GDP PT : correlation = 0.82
COVID_dummy ↔ Property Rights : correlation = 0.85
Interest rate ↔ PSI-20 : correlation = 0.77
PSI-20 ↔ Public Debt PT : correlation = -0.78
PGB 10Y ↔ Unemployment rate : correlation = 0.70
Exchange rate US to EU ↔ GNI : correlation = -0.64
Exchange rate US to EU ↔ House Pricing PT : correlation = -0.63
Exchange rate US to EU ↔ Exports PT M€ : correlation = -0.73
GNI ↔ House Pricing PT : correlation = 0.96
GNI ↔ PCRH-D : correlation = 0.85
Exports PT M€ ↔ GNI : correlation = 0.92
GNI ↔ Real GDP PT : correlation = 0.89
COVID_dummy ↔ GNI : correlation = 0.81
House Pricing PT ↔ PCRH-D : correlation = 0.87
Exports PT M€ ↔ House Pricing PT : correlation = 0.86
House Pricing PT ↔ Real GDP PT : correlation = 0.89
House Pricing PT ↔ Unemployment rate : correlation = -0.68
COVID_dummy ↔ House Pricing PT : correlation = 0.89
Interest rate ↔ Public Debt PT : correlation = -0.76
Exports PT M€ ↔ PCRH-D : correlation = 0.68
PCRH-D ↔ Real GDP PT : correlation = 0.98
PCRH-D ↔ Unemployment rate : correlation = -0.71
COVID_dummy ↔ PCRH-D : correlation = 0.65
Exports PT M€ ↔ Real GDP PT : correlation = 0.74
COVID_dummy ↔ Exports PT M€ : correlation = 0.73
Real GDP PT ↔ Unemployment rate : correlation = -0.69
COVID_dummy ↔ Real GDP PT : correlation = 0.66
```

Source: Jupiter Notebook, Python

Appendix C: Data Testing and Modelling

C1) Test the data

The objective of this paper is therefore to explore the determinants of NPLs of the Portuguese banking system, with NPLs as the dependent variable. The research methodology relies on methods similar to Foglia (2022) and Kjosevski, Petkovski, and Naumovska (2019). As the dataset covers the years 2016 to 2024, periods of missing observations were left out. As a result, the total data set contains 36 observations. The selected period provides a homogeneity of variables and a robustness of further statistical analysis.

Before estimating the model, the ADF test was applied to test for the stationarity of the variables in keeping with the methods of Kjosevski et al. (2019), Foglia (2022) and Castro and Santos (2010). The ADF test tests the null hypothesis that a unit root is present in a time series against the alternative that the time series is stationary around a constant level:

- Null hypothesis (H0): The series contains a unit root (non-stationary).
- Alternative hypothesis (H1): The series is stationary.

```
ADF Statistic: -4.892988300610958
p-value: 3.598705593244005e-05
```

Figure 32 - ADF Statistic

The test statistic of the NPL series, , ADF obtained an ADF statistic of - 4.893 with a p-value of 0.0000359. Since the p-value is much smaller than 0.05, the null hypothesis of non-stationarity is turned down, and the NPL series is stationary at levels and is also I(0). This is an important property because it makes it possible for NPL to be incorporated in econometric models such as the ARDL model, which can handle a mix of I(0) and I(1) series. This preliminary step guarantees that, in what follows, the estimation of the model is performed in statistically valid and properly behaved series.

Homogeneity of the variables: linear relationships were examined considering the independent variables and to decide the representative predictors by Pearson's correlation

analysis, following the evaluation of the stationarity, according to the methods used in Cotovio (2023), Kanapickienė et al. (2023) and Gounder and Bajaj (2022).

The correlation matrix showed many pairs of variables with very high correlations among them, which suggests strong linear relationships and possible redundancy. A cautious selection strategy was introduced to reduce multicollinearity and improve interpretability.

According to the Pearson correlation outputs, among the highly correlated variables of macroeconomic indicators, only PCRH-D was selected as the representative variable. Of institutional quality indicators, Judicial Effectiveness was chosen as the item that is most directly related to the credit enforcement environment. The NPL ratio was kept as the dependent variable, removing closely associated ratios such as FSI to prevent the collinearity problem. A COVID dummy was also included to account for structural breaks around the pandemic and to allow us to examine differences in pre- and post-COVID correlations. This strategy is useful in the sense that it can decrease the dimensionality of the original problem and make correlation patterns clear and still consistent in the theoretical sobriety and quality of the data, and generate a strong database for further regression analysis.

VIF estimation was then carried out to formally check for any multicollinearity, as suggested by Foglia (2022), Mixão (2020) and Milenković et al. (2024). Highly correlated variables were then screened, and those that gave rise to inflated VIF values were taken out one at a time in a stepwise manner to maintain stable estimates of coefficients of the regression equations.

	Variable	VIF
0	const	1323.362527
1	NPL	18.250168
2	GNI	23.799227
3	Property Rights	8.504375
4	Unemployment rate	6.061442
5	CAR	12.306951
6	Interest rate	6.931213
7	COVID_dummy	8.060625

Table 5 - Variance Inflation Factor (VIF) output

The last group of variables were chosen by Pearson's correlation analysis, VIF results and empirical reasons. NPL was assigned as the endogenous (dependent) variable because it was a pivotal variable in this research. As far as macro-financial indicators are concerned, GNI proved to be the best measure of economic output and was selected for the representation of economic activity as it takes on the economy as a whole, without overlapping with other highly correlated variables as PCRH-D and Real GDP. Property Rights is the institutional quality variable, which depicts the effectiveness of law and regulation, in particular credit enforcement, and is theoretically correlated with bank credit risk. The Unemployment Rate was chosen to reflect labour market conditions, which may affect loan delinquencies. The CAR and Interest rate are included in the retained indicators to depict the financial soundness of a particular bank and the credit cost, respectively. Lastly, a COVID dummy was added to capture structural breaks and potential NPL behavioural shift due to the pandemic. This selection of variables balances theoretical relevance and empirical strength, reduces the collinearity among them and lays a good base for further regression.

The Elastic Net regression approach was used to further improve the importance of variables and avoid overfitting. Elastic Net is a regularised regression strategy which consists of a combination of Lasso (L1) and Ridge (L2) penalties and aims to induce sparsity as well as the grouping effect of highly correlated predictors (Zou & Hastie, 2005; Friedman, Hastie & Tibshirani, 2010; Friedman, 2020). This method of modelling is particularly suitable for settings (such as the current study) where the number of predictors is high relative to the number of observations.

Furthermore, Elastic Net was used to compromise between the advantages of L1 and L2 regularisation, allowing L1 rule to maintain the most useful features while shrinking coefficients of irrelevant variables to zero. The existing literature on equivalent macro-financial datasets suggests that the use of small values for α (regularisation parameter), usually ranging from 0.0001 to 0.001, allows for maintaining the accuracy of the model whilst taking advantage of the regularisation effect of the Elastic Net (Zou & Hastie, 2005). The choice of an α is important, since too much regularisation yields underfitting and too little keeps the predictive signal while giving more liberty to the inflation of the coefficients (Pedregosa et al., 2011; Zou & Hastie, 2005). Using this

evidence, the regularisation parameter, α, was set at 0.0001, indicative of a greatly undisturbed punishment for the coefficients of the model. A parameter sweep was used to select the combination of penalties, and α = 0.0001 was verified to be the optimal choice on this dataset. Hence, the use of Elastic Net in this analysis has two benefits: it shrinks dimensionality by penalising low contributing predictors and keeps in theoretically important model predictors, improving the interpretability and generalizability of the resulting model.

Intercept: 0.08611038785714285

	Variable	Coefficient
3	FSI	0.045113
6	LDR	0.005340
21	Unemployment rate	0.002445
7	Business Freedom	0.001765
12	PGB 10Y	0.001442
23	COVID_dummy	-0.000612
0	CAR	-0.000600
9	Judicial Effectiveness	-0.000253
16	Inflation rate	-0.000109
14	GNI	0.000037
4	Personal Loan PT	-0.000024
17	Interest rate	-0.000000
22	Public Debt PT	0.000000
2	Firm Loan PT	0.000000
20	Real GDP PT	0.000000
19	Exports PT M€	-0.000000
18	PCRH-D	0.000000
5	ROA	0.000000
13	Exchange rate US to EU	-0.000000
1	Bank Size	0.000000
11	PSI-20	-0.000000
10	Property Rights	0.000000
8	Investment Freedom	0.000000
15	House Pricing PT	0.000000

Table 6 – Elastic Net results

The output of the Elastic Net is used to identify the variables kept in the model (non-zero coefficients), which are the key drivers of NPLs in the Portuguese banking industry. In the Elastic Net model, FSI (expected), LDR and Unemployment rate became significant positive predictors of higher NPLs, thus suggesting that higher FSI ratio, aggressive lending and higher level of unemployment are associated with high NPLs. COVID dummy, CAR, Judicial Effectiveness, and Inflation rate exhibited minor negative effects, which means that regulatory forbearance, strong bank capital, fast judicial systems higher inflation have a minor reduction effect on NPLs. GNI and Personal Loan PT had little influence. On the other hand, variables such as Bank Size, Firm Loan PT, Real GDP, Property Rights, PI-20, exchange rate US to EU, House pricing, Exports,

Interest rate, Investment freedom, ROA, PCRH-D, and Public debt were compressed to zero, which means they had less impact on NPLs and justifying their omission from the analysis. This model selects for the most influential and theoretically promising predictors, leading to a simple and interpretable framework for the analysis of the determinants of NPLs and avoiding overfitting and multicollinearity.

In conclusion, the 2016–2024 dataset was meticulously processed and scrubbed, providing 36 full observations. Stationarity of the NPL variable was verified by the tests for I(0), and the Pearson correlation analysis was also used to identify the representative variables with the lowest redundancy. Multicollinearity was managed by omission of VIF through a sequential stepwise execution in Elastic Net regression; it also allows both variable selection and regularisation to obtain a feasible and stable model.

C2) Vector Autoregression (VAR) Framework

The trend-indicators for NPLs, as well as bank-specific and macroeconomic variables, are used in the VAR framework, which allows for the dynamic relationships between them. This method is very appropriate since it considers all the variables to be endogenous and hence can be used to test the feedback effect of developments in the credit cycle. Previous work, e.g., Marques et al. (2020), underlines the role of VAR in investigating the NPL credit link. For instance, Espinoza and Prasad (2010) find that a one standard deviation increase in the NPL ratio reduces credit growth by 1.5 and 2.2 p.p. in the years two and three ahead, respectively, and Klein (2013), via panel VAR, illustrates that a 1 p.p. elevation in the NPL ratio brings a cumulative decline in the credit-to-GDP ratio. These results provide evidence that NPL shocks have real effects on loan facilitations and the overall macroeconomic environment.

The selection of variables maintained in the Portuguese case study was controlled by the Elastic Net and VIF diagnostics, so it was ensured that the selected variables had significant theoretical relevance and were robust to multicollinearity. The LDR as a proxy for bank liquidity, the Unemployment rate as a macroeconomic driver of credit risk, the CAR for bank solvency, Judicial Effectiveness to control for institutional quality, the Interest rate as the main macro-financial variable and the PGB 10y yield and the inflation rate were the final variables included in the above model specification. This selective

choice prevented redundancy, so that, for example, the FSI has been removed (because collinear with NPL), preserving, on the other hand, relevant macro-financial and institutional dimensions. The COVID-19 dummy was not added to the VAR specification as it is a structural break rather than a dynamic endogenous variable for autoregressive modelling.

```
Colunas constantes (serão removidas): []
  Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Sun, 07, Sep, 2025
Time:                     14:59:21
--------------------------------------------------------------------
No. of Equations:         7.00000    BIC:                   -30.1615
Nobs:                     35.0000    HQIC:                  -31.7910
Log likelihood:           279.736    FPE:                7.00240e-15
AIC:                     -32.6501    Det(Omega_mle):     1.65746e-15
--------------------------------------------------------------------
Results for equation Unemployment rate
=================================================================================================
                                coefficient       std. error           t-stat            prob
-------------------------------------------------------------------------------------------------
const                             0.032295         0.073093            0.442           0.659
L1.Unemployment rate              0.369209         0.166179            2.222           0.026
L1.CAR                            0.001137         0.002344            0.485           0.628
L1.Interest rate                 -0.271110         0.272349           -0.995           0.320
L1.Judicial Effectiveness        -0.000217         0.000087           -2.501           0.012
L1.LDR                            0.000052         0.000553            0.093           0.926
L1.PGB 10Y                        0.004113         0.003179            1.294           0.196
L1.Inflation rate                -0.016036         0.082035           -0.195           0.845
=================================================================================================

Results for equation CAR
=================================================================================================
                                coefficient       std. error           t-stat            prob
-------------------------------------------------------------------------------------------------
const                            14.450666         4.207434            3.435           0.001
L1.Unemployment rate             15.751985         9.565673            1.647           0.100
L1.CAR                            0.406596         0.134935            3.013           0.003
L1.Interest rate                 24.072965        15.677099            1.536           0.125
L1.Judicial Effectiveness         0.012191         0.004999            2.439           0.015
L1.LDR                           -0.083936         0.031843           -2.636           0.008
L1.PGB 10Y                        0.011534         0.183012            0.063           0.950
L1.Inflation rate               -11.403575         4.722144           -2.415           0.016
=================================================================================================
```

Results for equation Interest rate

	coefficient	std. error	t-stat	prob
const	0.022450	0.016469	1.363	0.173
L1.Unemployment rate	-0.013270	0.037442	-0.354	0.723
L1.CAR	-0.001551	0.000528	-2.936	0.003
L1.Interest rate	1.091876	0.061364	17.793	0.000
L1.Judicial Effectiveness	-0.000010	0.000020	-0.499	0.618
L1.LDR	0.000020	0.000125	0.162	0.872
L1.PGB 10Y	-0.001531	0.000716	-2.137	0.033
L1.Inflation rate	0.145260	0.018484	7.859	0.000

Results for equation Judicial Effectiveness

	coefficient	std. error	t-stat	prob
const	120.788465	146.307472	0.826	0.409
L1.Unemployment rate	-288.353562	332.632565	-0.867	0.386
L1.CAR	4.904031	4.692162	1.045	0.296
L1.Interest rate	-1223.613179	545.148633	-2.245	0.025
L1.Judicial Effectiveness	0.406101	0.173820	2.336	0.019
L1.LDR	-1.591055	1.107299	-1.437	0.151
L1.PGB 10Y	14.243323	6.363997	2.238	0.025
L1.Inflation rate	-272.864459	164.205789	-1.662	0.097

Results for equation LDR

	coefficient	std. error	t-stat	prob
const	20.460182	12.604341	1.623	0.105
L1.Unemployment rate	9.033551	28.656188	0.315	0.753
L1.CAR	-0.404516	0.404228	-1.001	0.317
L1.Interest rate	-24.086386	46.964378	-0.513	0.608
L1.Judicial Effectiveness	-0.009169	0.014975	-0.612	0.540
L1.LDR	0.821291	0.095393	8.610	0.000
L1.PGB 10Y	0.418492	0.548256	0.763	0.445
L1.Inflation rate	-8.591148	14.146275	-0.607	0.544

Results for equation PGB 10Y

	coefficient	std. error	t-stat	prob
const	-6.383126	3.909432	-1.633	0.103
L1.Unemployment rate	10.585469	8.888161	1.191	0.234
L1.CAR	0.179906	0.125378	1.435	0.151
L1.Interest rate	9.818983	14.566731	0.674	0.500
L1.Judicial Effectiveness	-0.003957	0.004645	-0.852	0.394
L1.LDR	0.038777	0.029588	1.311	0.190
L1.PGB 10Y	0.656387	0.170050	3.860	0.000
L1.Inflation rate	17.358841	4.387687	3.956	0.000

Results for equation Inflation rate

	coefficient	std. error	t-stat	prob
const	-0.157783	0.124895	-1.263	0.206
L1.Unemployment rate	-0.066638	0.283951	-0.235	0.814
L1.CAR	0.009999	0.004005	2.496	0.013
L1.Interest rate	-1.072928	0.465365	-2.306	0.021
L1.Judicial Effectiveness	-0.000061	0.000148	-0.413	0.679
L1.LDR	0.000219	0.000945	0.232	0.816
L1.PGB 10Y	0.005108	0.005433	0.940	0.347
L1.Inflation rate	0.934666	0.140174	6.668	0.000

Correlation matrix of residuals

	Unemployment rate	CAR	Interest rate	Judicial Effectiveness	LDR	PGB 10Y	Inflation rate
Unemployment rate	1.000000	0.198195	-0.170000	-0.196966	-0.407987	0.061114	0.201091
CAR	0.198195	1.000000	-0.173191	-0.353962	-0.330116	-0.480115	-0.069557
Interest rate	-0.170000	-0.173191	1.000000	0.168566	0.171170	-0.033101	0.019626
Judicial Effectiveness	-0.196966	-0.353962	0.168566	1.000000	-0.096351	0.391024	0.336070
LDR	-0.407987	-0.330116	0.171170	-0.096351	1.000000	-0.279012	-0.416980
PGB 10Y	0.061114	-0.480115	-0.033101	0.391024	-0.279012	1.000000	0.506080
Inflation rate	0.201091	-0.069557	0.019626	0.336070	-0.416980	0.506080	1.000000

Table 7 - Vector Autoregression (VAR) results

The estimation results of VAR validate the adequacy of this specification. The model achieved a strong overall fit, with an AIC of (- 32.65), BIC of (- 30.16), and a log-likelihood of ~279.74. Several meaningful dynamics emerged. For the unemployment equation, persistence existed as lag unemployment was still significant and Judicial Effectiveness negatively affected (p-value = 0.012), which implies that higher institutional quality mitigates the fragility of the labour market. For the solvency equation, the lagged CAR was highly significant (p-value = 0.003), but solvency was also negatively determined by LDR (p-value = 0.008) and inflation (p = 0.016), and positively by institutional quality. For the interest rate equation, the partial autocorrelation ($p < 0.001$) was complemented by significant negative effects from the CAR (p = 0.003) and the PGB 10y yield (p = 0.033) and a strong positive effect of inflation ($p < 0.001$). Judicial Effectiveness displayed the interest rate shock persistence (p = 0.019) but negatively affected (p = 0.025) the same shocks. As anticipated, LDR exhibited strong persistence ($p < 0.001$), and the PGB 10y yield equation indicated significant autoregressive effects ($p < 0.001$) and strong sensitivity to inflation ($p < 0.001$). Finally, lagged CAR had a positive impact on inflation (p = 0.013), and interest rates were impacted negatively by inflation (p = 0.021), which further emphasised the feedback between macroeconomic and financial indicators.

The residual correlation matrix also indicates that the interdependencies are well explained and there is no evidence of excessive unexplained co-movements. In sum, the VAR analysis not only delivers a robust statistical specification but also is consistent with theoretical expectations in that the role of unemployment, capitalisation, liquidity and institutional quality in driving NPL dynamics in Portugal is underscored.

C3) Autoregressive Distributed Lag (ARDL)

Keeping the banking sector stable and efficient means that, becomes important to understand the reasons underlying NPLs. Studies such as Kjosevski et al. (2019) and Foglia (2022), to this end, have highlighted that the ARDL framework stands out for its small-sample, unmeasured with absolute consistency when this characteristic is inherent in the data, as it allows a clear distinction between short- and long-run effects. Before estimating the ARDL model, it is essential to determine the maximum number of lags to

include. The optimal lag length is selected using regression specification tests and information criteria such as the AIC and BIC, ensuring both statistical validity and parsimony. This structure allows the model to capture immediate responses to macroeconomic and bank-specific shocks, such as variations in GDP or LDR ratios, as well as their persistent effects over time. The ARDL approach can accommodate variables integrated of different orders such as I(0) or I(1), the cases that are relevant to the determinants of NPLs in our study, are suitable for deriving a dynamic ECM and also permit explicit resolution into short - long term problems (Banerjee et al., 1993; Pesaran & Shin, 1995; Pattichis, 1999). It is by combining a robust lag selection and this dual point of view that the ARDL methodology yields sound inference as to the determinants of NPLs in Portugal's banking system: it is economical to implement and makes sense in terms of policy. Selection of independent variables for the ARDL model was guided by Pearson correlation, VIF and VAR results, as well as theoretical relevance, its own empirical basis and the extent to which they influenced predicted output.

Variable	Importance for the model
Unemployment rate	Strong macroeconomic driver; positively affects NPLs; significant lag effects
CAR	Key bank-specific determinant; higher CAR reduces NPL risk
Judicial Effectiveness	Institutional factor: efficient enforcement reduces default risk
LDR	Reflects bank liquidity and credit exposure; positive effect on NPLs
Inflation rate	Weak short-term macro influence; included for robustness and completeness
Real GDP PT	Captures economic growth effects; reduces NPL risk; core macro determinant
Interest rate	Macro-financial relevance; captures cost of borrowing and policy impact
Bank Size	Captures scale-related effects on NPLs
COVIDdummy	Captures structural shocks from the pandemic; essential for 2020–2024 period

Table 8 - ARDL variable selection

C4) ARDL Estimation Results – Elastic Net, ARDL-based

After the estimation of the ARDL model, an Elastic Net regression was conducted with the independent variables selected to determine the relative elasticities and further establish the priority determinants influencing NPLs. Application of the Elastic Net in this setting offers a compromise between the LASSO and Ridge regression steps; it simultaneously applies variable selection and mitigates any possible multicollinearity in the highly correlated bank-specific and macroeconomic variables (Zou & Hastie, 2005;

Tibshirani, 1996). This method is particularly suitable when dealing with small samples with many correlated predictors, as in the present study.

Variable	Elasticity
NPL.L1	0.853500
LDR.L0	0.069933
LDR.L1	-0.020397
Real GDP PT.L0	2.216040
Bank Size.L0	24.844644
COVID_dummy.L0	0.629213
COVID_dummy.L1	-0.325843

Table 9 – Elastic Net ARDL-based results

Most importantly, the Elastic Net estimation using the AIC-selected ARDL model indicates the dominance of bank-level variables and NPL persistence. Lagged NPL (NPL. L1: 0.8535) is significantly positive, suggesting that historical credit defaults are significantly good predictors of the current NPLs, which is an empirical result consistent with the existence of a significant serial pattern of credit risk in the literature (Beck et al., 2013). Bank Size (24.8446 at Lag 0) affects more immediately than others and suggests that a larger bank is more exposed to credit risk since its lending is higher and its portfolio is more complex. LDR (0.0699 at Lag 0; -0.0204 at Lag 1) has a moderate instantaneous positive impact, then a small negative correction, consistent with short-term balance-sheet adjustments. Real GDP PT (2.2160 at Lag 0) has a contemporaneous effect and presents context-related combinations of economic growth and NPL determinants. The COVID dummy (0.6292 at Lag 0; -0.3258 at Lag 1) captures its temporary shock, highlighted by the immediate rise in NPLs followed by a partial correction, which reflects the influence of policy measures such as loan moratoria and capital support. Overall, these results suggest that lagged NPL and bank-level factors largely drive outcomes, whereas macro shocks and the pandemic take place mainly in the short run or as a transitional state. By integrating ARDL-based lag selection with Elastic Net estimation, this method confers physical meaning while reducing overfitting and remarkably recovers the main determinants of NPLs (Tibshirani, 1996; Zou & Hastie, 2005).

C5) ARDL Estimation Results – OLS simple regression, ARDL-based

After Elastic Net, it has been estimated that the OLS regression is a more desirable alternative for investigating the dynamics of NPL in the Portuguese banking sector.

OLS Regression Results

Dep. Variable:	NPL	R-squared:	0.998
Model:	OLS	Adj. R-squared:	0.998
Method:	Least Squares	F-statistic:	2292.
Date:	Sun, 07 Sep 2025	Prob (F-statistic):	9.47e-36
Time:	15:07:47	Log-Likelihood:	166.34
No. Observations:	35	AIC:	-316.7
Df Residuals:	27	BIC:	-304.2
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.8552	0.256	-3.338	0.002	-1.381	-0.330
NPL_lag1	0.8535	0.064	13.416	0.000	0.723	0.984
LDR_lag0	0.0024	0.000	4.967	0.000	0.001	0.003
LDR_lag1	-0.0007	0.001	-1.375	0.180	-0.002	0.000
Real GDP PT_lag0	1.131e-06	5.01e-07	2.256	0.032	1.02e-07	2.16e-06
Bank Size_lag0	0.1172	0.047	2.470	0.020	0.020	0.215
COVID_dummy_lag0	0.0168	0.005	3.218	0.003	0.006	0.028
COVID_dummy_lag1	-0.0087	0.004	-2.463	0.020	-0.016	-0.001

Omnibus:	3.943	Durbin-Watson:	2.297
Prob(Omnibus):	0.139	Jarque-Bera (JB):	2.608
Skew:	-0.610	Prob(JB):	0.271
Kurtosis:	3.545	Cond. No.	3.11e+07

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 3.11e+07. This might indicate that there are strong multicollinearity or other numerical problems.

Table 10 - OLS simple regression, ARDL-based results

ARDL (1,1,0,0,1) specification exhibits a high R-squared of 0.998, suggesting that almost all the variation in NPLs is accounted for. The overall significance of the model is supported by an adjusted R-squared = 0.998 and F-statistic = 2292.0 ($p < 0.001$). Information criteria AIC (-316.7) and BIC (-304.2) support a compromise between model fit and parsimony. Important variables are highly statistically significant and have economic meaning. Lagged NPL (NPL. L1: 0.8535, $p < 0.001$) reflects the strong persistence of credit risk, which is consistent with previous studies that find autocorrelation in credit defaults (Beck et al., 2013). The LDR (LDR. L0: 0.0024, $p < 0.001$) is positively related to NPLs, while Real GDP PT (L0: 1.131e-06, $p = 0.032$) consists of an economic growth stabilising effect. Bank Size (L0: 0.1172, $p = 0.020$) suggests that larger banks are subject to a little more NPLs, and the COVID dummy reflects pandemic concerns, with an immediate shift effect (L0: 0.0168, $p = 0.003$) and

its partial reversal in the following period (L1: -0.0087, p = 0.020). It suggests that, in this macroeconomic shock, especially, the NPL persistence and bank-specific variables play an important role, and they are dominant, whereas the macroeconomic shock and the pandemic sentiment shocks have small and temporary effects.

Residual diagnostics support the credibility of the model, as residuals are approximately normal (Jarque-Bera = 2.608; p = 0.271) and there is low evidence of autocorrelation (Durbin-Watson = 2.297). The condition number is high (3.11e+07) and suggests that there may be strong multicollinearity among the model features, but model stability is preserved. The OLS lag structure aligns with the ARDL framework and is consistent in capturing both short-run and long-run dynamics, facilitating interpretation and enabling coherent analysis of how macroeconomic and bank-specific determinants affect NPLs's time (Tibshirani, 1996; Zou, & Hastie, 2005; Beck et al, 2013).

The first stage of the Elastic Net estimate on the ARDL relational model with the AIC-selected lag-length pinpoints the most influential NPLs determinants through the persistence of credit risk with the lagged NPLs and the importance of the bank level characteristics, in particular Bank Size and LDR. By selecting variables and controlling for correlation between them, Elastic Net can reduce the risk of overfitting and multicollinearity, which is crucial when working with a small sample size and several lags (Tibshirani, 1996; Zou & Hastie, 2005). It was found that COVID-induced shocks have short-term effects that are partly reversed in the following periods. Extending these findings, the OLS regression, specified by a similar ARDL specification, furnishes very interpretable and statistically significant coefficients that support Elastic Net results. Elastic Net improves on model sparsity and benefits variable significance; at the same time, OLS provides for precise magnitudes and more direct economic interpretation. Combined, the two methodologies provide supporting evidence, increasing confidence in the main drivers of NPLs in the Portuguese banking system and laying a strong empirical base for policy implications.

C6) ARDL Estimation Results – Dynamic ECM estimation ARDL-based

After estimating the ARDL model and further Elastic Net and OLS regression analysis, the ECM has been calculated to decouple the short-run from long-run behaviour of NPLs in the Portuguese banking system. Especially from the viewpoint of both dynamic and policy implications, the ECM is crucial as it reflects conditions for which a differential away from the long-run equilibrium is adjusted over time (Banerjee et al., 1993; Pesaran & Shin, 1995; Kjosevski et al., 2019).

OLS Regression Results

Dep. Variable:	d_NPL	R-squared:	0.401
Model:	OLS	Adj. R-squared:	0.294
Method:	Least Squares	F-statistic:	3.754
Date:	Sun, 07 Sep 2025	Prob (F-statistic):	0.0100
Time:	15:08:07	Log-Likelihood:	145.00
No. Observations:	34	AIC:	-278.0
Df Residuals:	28	BIC:	-268.8
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.0029	0.001	-3.236	0.003	-0.005	-0.001
d_LDR	0.0024	0.001	3.798	0.001	0.001	0.004
d_Real GDP PT	7.356e-07	6.71e-07	1.096	0.283	-6.4e-07	2.11e-06
d_Bank Size	0.0831	0.104	0.797	0.432	-0.130	0.297
d_COVID_dummy	0.0065	0.007	1.001	0.325	-0.007	0.020
ECT_lag1	-0.5514	0.310	-1.781	0.086	-1.186	0.083

Omnibus:	1.431	Durbin-Watson:	0.800
Prob(Omnibus):	0.489	Jarque-Bera (JB):	1.302
Skew:	-0.445	Prob(JB):	0.521
Kurtosis:	2.646	Cond. No.	8.17e+05

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 8.17e+05. This might indicate that there are strong multicollinearity or other numerical problems.

Table 11 - Dynamic ECM estimation ARDL-based results

LDR turn out to be the only statistically significant related factor in the narrative, reflecting that bank liquidity has an influencing effect on credit risk, with the short-run relationship requiring a percentage of 40% of NPLs. The other short-run variables, ΔReal GDP, ΔBank Size, and ΔCOVID dummy, remain relatively insignificant and small, suggesting that short-term macroeconomic and health crisis effects are relatively insignificant. The ECT_lag1 has a highly negative coefficient, which means that adjustments from the long-run equilibrium are gradually adjusted at each period. The ECT is only weakly significant at the 10% level ($p = 0.086$), but its value implies inherently significant long-run adjustment talks of convergence are supported; indeed, the

persistence of NPLs is revealed in the ARDL and Elastic Net analysis. Diagnostic tests indicate a reasonable set of residuals having a slight skew and kurtosis and a bit of autocorrelation (Durbin-Watson = 0.800); multicollinearity is present, as indicated by the high condition number.

The comparison of the ARDL, Elastic Net, OLS and ECM methods reveals the revealed that each methodology has its own advantages. The complete (1,1,0,0,1) ARDL incorporates the short- and long-run lagged dynamics in addition to macroeconomic, bank-specific, and COVID-specific effects; however, its large size results in a minor degree of overfitting and non-significant lags. In the limited ARDL (1,0,0,0,1), more parsimoniously only includes core drivers NPL persistence, LDR, bank size, GDP and COVID shocks with similar goodness of fit - albeit with less complexity. Elastic Net, via coefficients ARDL, reduces multicollinearity and presents the most important predictors, which were previously found (the dominance of lagged NPLs and bank level), but does not give direct significance testing. OLS estimation provides very intuitive and statistically valid coefficients as expected in the ARDL framework; however, the very high R-squared and high condition number suggest multicollinearity. Lastly, the ECM captures the short-run dynamics where ΔLDR is the only short-term driver and there is an intermediate increase of lender controls towards equilibrium through the ECT in line with stubborn NPLs. These methods are combined to provide consistent answers on the short-term and long-term factors of NPLs, where the short-term and long-term dynamic relationship is provided by ARDL and ECM, Elastic Net is responsible for variable selection, while OLS is used to check the robustness of the selected variables.

OLS Regression Results

Dep. Variable:	d_NPL	R-squared:	0.343
Model:	OLS	Adj. R-squared:	0.256
Method:	Least Squares	F-statistic:	3.755
Date:	Sun, 07 Sep 2025	Prob (F-statistic):	0.0136
Time:	15:08:27	Log-Likelihood:	147.43
No. Observations:	35	AIC:	-284.9
Df Residuals:	30	BIC:	-277.1
Df Model:	4		
Covariance Type:	HC1		

	coef	std err	z	P>|z|	[0.025	0.975]
const	-0.0029	0.001	-4.478	0.000	-0.004	-0.002
d_LDR	0.0227	0.007	3.397	0.001	0.010	0.036
d_Real GDP PT	0.0041	0.002	1.963	0.050	7.24e-06	0.008
d_COVID_dummy	0.0112	0.006	2.021	0.043	0.000	0.022
ECT_lag1	-0.1861	0.093	-2.002	0.045	-0.368	-0.004

Omnibus:	4.768	Durbin-Watson:	0.916
Prob(Omnibus):	0.092	Jarque-Bera (JB):	3.281
Skew:	-0.625	Prob(JB):	0.194
Kurtosis:	3.828	Cond. No.	153.

Notes:
[1] Standard Errors are heteroscedasticity robust (HC1)

Table 12 - First differenced ECM and lagged ECT

The first differenced ECM with core explanatory variables and the lagged ECT is a parsimonious and intuitively appealing representation for NPL dynamics in the Portuguese bank sector. The model accounts for about 34% of the variation in the short run and is significant ($F = 3.755$, $p < 0.014$). Short-run dynamics are driven by a dominant effect of liquidity (ΔLDR) and transitory effects of COVID, and a marginal one of GDP. The negative and significant ECT value implies that there is a significant long-run relationship and that the system corrects for deviations by ~18.61% per period (Engle & Granger, 1987; Banerjee et al., 1998).

The diagnoses indicate stability and low multicollinearity (condition number = 153) and nearly normal residuals. This ECM estimation trades off parsimony, significance and economic interpretation, emphasising the substantive importance of bank liquidity, short-run shocks and gradual long-run adjustment, laying solid foundations for policy and risk considerations.

Appendix D: Covid dummy variable extended treatment

The first, a binary specification for the dummy set to 0 before the pandemic, 1 during the pandemic (1st March 2020 – 30th September 2022) and back to 0 after the pandemic. This follows the technique of Plikas et al. (2024), who regard COVID-19 as a temporary shock that generated a greater macroeconomic impact in post-pandemic periods, whilst Suradi, Siregar, and Sartono (2021) treated the pandemic as a transient variable that reflects an immediate but non-permanent effect on NPLs.

The second, phase-based representation splits the life of the economy into five stages following an approach similar to Cotovio (2023): Pre-COVID (from January 2008 to February 2020), COVID-2020 (from March to December 2020), COVID-2021 (from January to December 2021), COVID-2022 (from January to September 2022), and Post-COVID (from October 2022 to December 2024). Such a structure enables a more detailed examination of uniformity in the pandemic effects on NPL or its variation across stages and greater insight into credit risk dynamics in the NPLs of the Portuguese banking sector.

D1) Covid dummy variable: Binary Specification

The COVID dummy variable, which was added as a binary value to account for the structural break in 2020–2022, is negatively signed and reflects the temporary effect of emergency measures on credit risk. Based on these findings, the extended version of the study uses variance in the binary COVID dummy to analyse the dynamic transmission of NPLs to pandemic-induced shocks.

The variables used for the next step VAR analysis, FSI, LDR, unemployment rate, Real GDP PT, interest rate, CAR, and COVID dummy, were chosen by considering both statistical and economic criteria, similar to the method used in the initial version of the COVID dummy.

	Variable	VIF
0	const	1152.658465
1	NPL	14.101570
2	ROA	10.919783
3	Real GDP PT	5.801665
4	Unemployment rate	9.644202
5	Judicial Effectiveness	3.322122
6	Interest rate	7.344269
7	COVID_dummy	4.581330

Table 13 - VIF results - Binary Specification

Intercept: 0.08611038785714285

	Variable	Coefficient
3	FSI	0.045773
6	LDR	0.005346
21	Unemployment rate	0.002347
7	Business Freedom	0.001680
12	PGB 10Y	0.001184
0	CAR	-0.000895
14	GNI	0.000323
23	COVID_dummy	-0.000221
9	Judicial Effectiveness	-0.000199
16	Inflation rate	-0.000132
22	Public Debt PT	-0.000000
2	Firm Loan PT	-0.000000
20	Real GDP PT	0.000000
19	Exports PT M€	-0.000000
18	PCRH-D	0.000000
17	Interest rate	-0.000000
5	ROA	0.000000
4	Personal Loan PT	-0.000000
13	Exchange rate US to EU	-0.000000
1	Bank Size	0.000000
11	PSI-20	-0.000000
10	Property Rights	0.000000
8	Investment Freedom	0.000000
15	House Pricing PT	-0.000000

Table 14 – Elastic Net results - Binary Specification

The issue of high multicollinearity was tackled by using VIF analysis, such that each variable brings in separate explanatory power, and the Elastic Net results showed the most important NPLs drivers focusing on systemic credit risk, bank leverage and macroeconomic stress. FSI and LDR articles internal bank risk, systemic credit exposure, and CAR relate to bank solvency buffer, unemployment, and Real GDP consist of cyclical macroeconomic conditions, interest rate is used for monetary policy transmission, and COVID dummy explains temporal structural shocks, from the economic point of view.

This setup enables the VAR to trace the dynamic relationship between macroeconomic conditions, banking system health, and credit risk over time, while explicitly taking into account structural breaks and policy support measures adopted in the pandemic. The VAR model focused on a selected number of variables: FSI, LDR,

unemployment rate, Real GDP PT, rate of interest, CAR, and COVID dummy, where the COVID dummy is also addressed by the binary specification to control the structural break over 2020–2022.

```
  Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Sun, 07, Sep, 2025
Time:                     19:31:40
----------------------------------------------------------------------
No. of Equations:         7.00000    BIC:                   -3.43133
Nobs:                     34.0000    HQIC:                  -5.08798
Log likelihood:          -180.637    FPE:                  0.00278798
AIC:                     -5.94533    Det(Omega_mle):     0.000635180
----------------------------------------------------------------------
Results for equation FSI
================================================================================================
                           coefficient       std. error           t-stat            prob
------------------------------------------------------------------------------------------------
const                        -0.014951         0.127379           -0.117           0.907
L1.FSI                       -0.071348         0.175699           -0.406           0.685
L1.LDR                       -0.363469         0.147776           -2.460           0.014
L1.Unemployment rate         -0.250119         0.170312           -1.469           0.142
L1.Real GDP PT                0.061557         0.159413            0.386           0.699
L1.Interest rate              0.877285         0.193110            4.543           0.000
L1.CAR                       -0.009643         0.133241           -0.072           0.942
L1.COVID_dummy                0.640106         0.167862            3.813           0.000
================================================================================================

Results for equation LDR
================================================================================================
                           coefficient       std. error           t-stat            prob
------------------------------------------------------------------------------------------------
const                         0.037373         0.151906            0.246           0.806
L1.FSI                       -0.212934         0.209530           -1.016           0.310
L1.LDR                       -0.211132         0.176231           -1.198           0.231
L1.Unemployment rate          0.023875         0.203105            0.118           0.906
L1.Real GDP PT                0.201448         0.190108            1.060           0.289
L1.Interest rate              0.421635         0.230294            1.831           0.067
L1.CAR                        0.422190         0.158896            2.657           0.008
L1.COVID_dummy                0.294804         0.200185            1.473           0.141
================================================================================================
```

Results for equation Unemployment rate

	coefficient	std. error	t-stat	prob
const	0.010691	0.135463	0.079	0.937
L1.FSI	0.327006	0.186849	1.750	0.080
L1.LDR	0.111960	0.157154	0.712	0.476
L1.Unemployment rate	0.106535	0.181120	0.588	0.556
L1.Real GDP PT	-0.656189	0.169530	-3.871	0.000
L1.Interest rate	-0.006895	0.205365	-0.034	0.973
L1.CAR	-0.046519	0.141696	-0.328	0.743
L1.COVID_dummy	0.167597	0.178515	0.939	0.348

Results for equation Real GDP PT

	coefficient	std. error	t-stat	prob
const	0.002946	0.160799	0.018	0.985
L1.FSI	0.462557	0.221798	2.085	0.037
L1.LDR	-0.231282	0.186548	-1.240	0.215
L1.Unemployment rate	-0.154592	0.214997	-0.719	0.472
L1.Real GDP PT	-0.218968	0.201238	-1.088	0.277
L1.Interest rate	-0.011679	0.243776	-0.048	0.962
L1.CAR	-0.135458	0.168199	-0.805	0.421
L1.COVID_dummy	0.192952	0.211905	0.911	0.363

Results for equation Interest rate

	coefficient	std. error	t-stat	prob
const	0.048199	0.037802	1.275	0.202
L1.FSI	-0.041365	0.052143	-0.793	0.428
L1.LDR	0.030923	0.043856	0.705	0.481
L1.Unemployment rate	0.001234	0.050544	0.024	0.981
L1.Real GDP PT	0.015560	0.047309	0.329	0.742
L1.Interest rate	1.034795	0.057310	18.056	0.000
L1.CAR	-0.014671	0.039542	-0.371	0.711
L1.COVID_dummy	0.093971	0.049817	1.886	0.059

Results for equation CAR

	coefficient	std. error	t-stat	prob
const	0.015083	0.172067	0.088	0.930
L1.FSI	-0.219588	0.237339	-0.925	0.355
L1.LDR	0.199148	0.199620	0.998	0.318
L1.Unemployment rate	-0.072067	0.230062	-0.313	0.754
L1.Real GDP PT	0.039625	0.215339	0.184	0.854
L1.Interest rate	0.367131	0.260858	1.407	0.159
L1.CAR	-0.363775	0.179985	-2.021	0.043
L1.COVID_dummy	0.017341	0.226753	0.076	0.939

Results for equation COVID_dummy

	coefficient	std. error	t-stat	prob
const	-0.001411	0.089378	-0.016	0.987
L1.FSI	-0.108728	0.123283	-0.882	0.378
L1.LDR	0.007540	0.103690	0.073	0.942
L1.Unemployment rate	0.002788	0.119503	0.023	0.981
L1.Real GDP PT	-0.081088	0.111855	-0.725	0.468
L1.Interest rate	-0.104104	0.135499	-0.768	0.442
L1.CAR	0.132311	0.093491	1.415	0.157
L1.COVID_dummy	0.893869	0.117784	7.589	0.000

Correlation matrix of residuals

	FSI	LDR	Unemployment rate	Real GDP PT	Interest rate	CAR	COVID_dummy
FSI	1.000000	0.340501	-0.390555	-0.227117	-0.001139	0.048151	0.408666
LDR	0.340501	1.000000	-0.077858	0.052454	0.388263	0.075899	-0.004045
Unemployment rate	-0.390555	-0.077858	1.000000	0.408823	-0.057003	-0.138234	-0.341932
Real GDP PT	-0.227117	0.052454	0.408823	1.000000	-0.168881	0.145557	-0.569831
Interest rate	-0.001139	0.388263	-0.057003	-0.168881	1.000000	-0.118135	-0.172636
CAR	0.048151	0.075899	-0.138234	0.145557	-0.118135	1.000000	-0.229988
COVID_dummy	0.408666	-0.004045	-0.341932	-0.569831	-0.172636	-0.229988	1.000000

Correlation matrix of residuals:

	FSI	LDR	Unemployment rate	Real GDP PT	\
FSI	1.000000	0.340501	-0.390555	-0.227117	
LDR	0.340501	1.000000	-0.077858	0.052454	
Unemployment rate	-0.390555	-0.077858	1.000000	0.408823	
Real GDP PT	-0.227117	0.052454	0.408823	1.000000	
Interest rate	-0.001139	0.388263	-0.057003	-0.168881	
CAR	0.048151	0.075899	-0.138234	0.145557	
COVID_dummy	0.408666	-0.004045	-0.341932	-0.569831	

	Interest rate	CAR	COVID_dummy
FSI	-0.001139	0.048151	0.408666
LDR	0.388263	0.075899	-0.004045
Unemployment rate	-0.057003	-0.138234	-0.341932
Real GDP PT	-0.168881	0.145557	-0.569831
Interest rate	1.000000	-0.118135	-0.172636
CAR	-0.118135	1.000000	-0.229988
COVID_dummy	-0.172636	-0.229988	1.000000

Table 15 - Vector Autoregression (VAR) results - Binary Specification

With the binary COVID dummy, FSI retention is significantly influenced by lagged interest rate (0.877, t = 4.543, $p < 0.001$), lagged COVID dummy (0.640, t = 3.813, $p < 0.001$), and lagged LDR (-0.363, t = -2.460, p = 0.014), capturing the relation of macro-financial situation, bank leverage, and pandemic induced crisis. LDR is positively influenced by lagged CAR (0.422, t = 2.657, p = 0.008), and unemployment rate reacts negatively with lagged Real GDP PT (-0.656, t = -3.871, $p < 0.001$), tying macroeconomic growth to labour market performances. Lagged FSI (0.463, t = 2.085, p = 0.037) affects positively Real GDP PT - signalling feedback from the stress occurring in the banking sector to the overall economy. Interest rate is still highly autoregressive (1.035, t = 18.056, $p < 0.001$) and has a borderline effect from a lagged COVID dummy (0.094, t = 1.886, p = 0.059). CAR is significantly influenced by its own lag (-0.364, t = -2.021, p = 0.043), and the binary dummy for COVID shows strong persistence (0.894, t = 7.589, $p < 0.001$). Interestingly, residual correlations exhibit moderate relationships, especially FSI and COVID dummy (0.409) and Real GDP PT and COVID dummy (-0.570). The binary coding of the COVID dummy allows the VAR to account for structural shocks and their transmission mechanisms through bank-specific and macroeconomic variables, providing a more accurate picture of the NPL determinants than the former model.

Based on the earlier variable obtained results conducted by the VIF, Elastic Net, and VAR methods, it was employed the ARDL model was employed to investigate the

short- and long-run factors of NPLs. The AIC-selected specification is ARDL (2,0,0,0,0,0,0), while the BIC-optimal model is more parsimonious, ARDL (1,0,0,0,0,0,0). In all cases, lagged NPL is strongly persistent, and with NPL.L1 coefficients of 0.613 (AIC, p = 0.011) and 1.023 (BIC, p < 0.001), highlighting the persistence of credit risk and how existing non-performing loans in the previous months will affect current levels.

```
Dependent var in columns: True
Independent vars in columns: True
Optimal ARDL order using AIC: (2, 0, 3, 0)
Summary of ARDL model selected by AIC:
                              ARDL Model Results
==============================================================================
Dep. Variable:                      NPL   No. Observations:                   36
Model:             ARDL(2, 0, 0, 0, 0, 0, 0)   Log Likelihood                 152.972
Method:                  Conditional MLE   S.D. of innovations              0.003
Date:                   Sun, 07 Sep 2025   AIC                           -285.944
Time:                           19:32:25   BIC                           -270.681
Sample:                      09-30-2016   HQIC                          -280.739
                           - 12-31-2024
==============================================================================
                                coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const                         -0.0525      0.027     -1.919      0.066      -0.109       0.004
NPL.L1                         0.6135      0.223      2.746      0.011       0.153       1.074
NPL.L2                         0.3926      0.212      1.848      0.076      -0.045       0.830
CAR.L0                         0.0005      0.001      0.491      0.628      -0.001       0.002
Unemployment rate.L0           0.1048      0.110      0.956      0.348      -0.121       0.331
Real GDP PT.L0             7.006e-07   4.32e-07      1.623      0.117   -1.89e-07    1.59e-06
Interest rate.L0               0.2129      0.078      2.722      0.012       0.052       0.374
Judicial Effectiveness.L0 -8.654e-05   5.07e-05     -1.706      0.100      -0.000    1.79e-05
COVID_dummy.L0                 0.0085      0.003      2.906      0.008       0.002       0.014
==============================================================================
Optimal ARDL order using BIC: (1, 0, 0)
Summary of ARDL model selected by BIC:
                              ARDL Model Results
==============================================================================
Dep. Variable:                      NPL   No. Observations:                   36
Model:             ARDL(1, 0, 0, 0, 0, 0, 0)   Log Likelihood                 154.797
Method:                  Conditional MLE   S.D. of innovations              0.003
Date:                   Sun, 07 Sep 2025   AIC                           -291.595
Time:                           19:32:33   BIC                           -277.597
Sample:                      06-30-2016   HQIC                          -286.763
                           - 12-31-2024
==============================================================================
                                coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const                         -0.0276      0.023     -1.174      0.251      -0.076       0.021
NPL.L1                         1.0232      0.048     21.367      0.000       0.925       1.121
CAR.L0                         0.0004      0.001      0.458      0.651      -0.002       0.002
Unemployment rate.L0           0.0079      0.099      0.080      0.937      -0.195       0.211
Real GDP PT.L0             3.556e-07   3.84e-07      0.927      0.362   -4.32e-07    1.14e-06
Interest rate.L0               0.2034      0.082      2.490      0.019       0.036       0.371
Judicial Effectiveness.L0   -9.2e-05    4.6e-05     -2.002      0.055      -0.000     2.3e-06
COVID_dummy.L0                 0.0074      0.003      2.562      0.016       0.001       0.013
==============================================================================
```

Table 16 - ARDL model results - Binary Specification

The interest rate factor is the key driver with coefficients of 0.2129 (AIC, p = 0.012) and 0.203 (BIC, p = 0.019), respectively, suggesting that the risk of credit increases with higher interest rates, most likely due to higher debt servicing for borrowers. Likewise, the binary COVID dummy is positive and significant in both specifications

(AIC: 0.0085, p = 0.008; BIC: 0.0074, p = 0.016), suggesting the pandemic causes a structural shock to the credit quality and thereby triggers temporary strains for bank portfolios.

The impact of Judicial Effectiveness is negative in both the models (AIC: -8.65e-05, p = 0.100; BIC: -9.2e-05, p = 0.055). This means stronger institutions and legal setup may somewhat reduce the credit risk, albeit at marginally significant levels. On the contrary, the remaining factors, as CAR, Unemployment rate, and Real GDP PT, demonstrate no statistically significant effects in this binary specification, thus leading to a limited short-run impact of these independent variables on NPLs within this framework.

As a whole, the ARDL analysis emphasises the primacy of interest rates and pandemic-related shocks in determining NPL dynamics and confirms the permanence of credit risk while revealing the possible dampening impact of institutional quality.

The ARDL framework allows us to provide additional insights about the determinants of NPLs, both under the binary COVID dummy specification using Elastic Net and OLS estimations.

Variable	Elasticity
NPL.L1	1.023200
CAR.L0	0.000960
Unemployment rate.L0	0.011060
Real GDP PT.L0	1.564640
Interest rate.L0	0.827431
Judicial Effectiveness.L0	-0.001380
COVID_dummy.L0	0.014800

Table 17 – Elastic Net ARDL-based - Binary Specification

The lagged NPL dominates in the Elastic Net ARDL-based estimates, confirming that credit risk is very persistent (elasticity of 1.023). The interest rate elasticities are (0.827), further suggesting the importance of interest rate in NPL dynamics, affirming the theoretical proposition that higher borrowing cost serves to raise the chance of NPL occurring. Intuitively, the elasticities of each variable indicate how macroeconomic variables (such as GDP PT) have a moderate effect (or at least a potentially nonlinear one) on credit risk. The COVID dummy remains positively significant (~0.015), and the sign of this effect is an indication that pandemic shocks are still exerting downward

pressure on NPL, while the elastic effect of Judicial Effectiveness is found to be modestly negative (-0.001), implying that better institutional quality can slightly mitigate NPL.

OLS Regression Results

Dep. Variable:	NPL	R-squared:	0.997
Model:	OLS	Adj. R-squared:	0.996
Method:	Least Squares	F-statistic:	1119.
Date:	Sun, 07 Sep 2025	Prob (F-statistic):	5.62e-31
Time:	19:38:24	Log-Likelihood:	156.81
No. Observations:	35	AIC:	-295.6
Df Residuals:	26	BIC:	-281.6
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.7050	0.381	-1.849	0.076	-1.489	0.079
NPL_lag1	1.0244	0.046	22.232	0.000	0.930	1.119
CAR_lag0	-0.0003	0.001	-0.270	0.789	-0.002	0.002
Unemployment rate_lag0	-0.0620	0.103	-0.603	0.552	-0.273	0.149
Real GDP PT_lag0	-4.58e-07	5.88e-07	-0.779	0.443	-1.67e-06	7.5e-07
Interest rate_lag0	0.1359	0.087	1.557	0.131	-0.043	0.315
Judicial Effectiveness_lag0	-7.933e-05	4.48e-05	-1.771	0.088	-0.000	1.27e-05
Bank Size_lag0	0.1304	0.073	1.780	0.087	-0.020	0.281
COVID_dummy_lag0	0.0020	0.004	0.497	0.623	-0.006	0.011

Omnibus:	18.935	Durbin-Watson:	2.940
Prob(Omnibus):	0.000	Jarque-Bera (JB):	31.104
Skew:	-1.323	Prob(JB):	1.76e-07
Kurtosis:	6.784	Cond. No.	3.49e+07

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 3.49e+07. This might indicate that there are strong multicollinearity or other numerical problems.

Table 18 - OLS regression results ARDL-based – Binary Specification

The OLS estimate further strengthens the importance of NPL persistence, when NPL_lag1 is found to have a coefficient of 1.024 ($t = 22.232$, $p < 0.001$). None of the other covariates reaches significance levels due to the strong autocorrelation in the dependent variable. Importantly, the COVID dummy coefficient in OLS (0.002, $p = 0.623$) is lower and insignificant, implying that after controlling for the lagged NPL, the immediate effect of the COVID shock on credit risk disappears in the short term. The OLS ARDL-based model is quite explanatory ($R^2 = 0.997$), but the high condition number indicates some possible multicollinearity, mainly between lagged NPL and the other explanatory variables, as would be expected in the persistence effects.

In sum, the empirical evidence suggests that the persistence of NPLs emerges as a determinant of short-run dynamics, thus suggesting that the role of interest rates and macroeconomic activity remains relevant. The Elastic Net results persist in identifying the hidden origins of the structural shock and institutional determinants, while the OLS

approach reinforces the dominant influence of historical credit risk, underlining the path-dependency of the bank portfolios in the case of the Portuguese. These results reinforce the correlations that exist between macroeconomic factors and banks and the possible effects towards NPLs, but also how the instant effects take the second role when the remaining is the persistence of past NPLs.

Subsequently, the ECM is estimated under the binary COVID dummy specification to understand the short-run dynamics of the NPLs. The model explains a moderate amount of the variability of NPLs ($R^2 = 0.417$); however, much of its short-run variation is not explained (adj. $R^2 = 0.189$), probably since the credit risk factor is slowly evolving, and multicollinearity among regressors would be present, as shown by the high condition number.

OLS Regression Results

Dep. Variable:	d_NPL	R-squared:	0.417
Model:	OLS	Adj. R-squared:	0.189
Method:	Least Squares	F-statistic:	1.831
Date:	Sun, 07 Sep 2025	Prob (F-statistic):	0.117
Time:	19:39:45	Log-Likelihood:	140.76
No. Observations:	33	AIC:	-261.5
Df Residuals:	23	BIC:	-246.6
Df Model:	9		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.0012	0.001	-0.859	0.399	-0.004	0.002
d_CAR	-0.0012	0.002	-0.593	0.559	-0.005	0.003
d_Unemployment rate	0.0068	0.137	0.049	0.961	-0.277	0.291
d_Interest rate	0.1143	0.224	0.510	0.615	-0.349	0.578
d_Real GDP PT	8.316e-07	6.34e-07	1.312	0.202	-4.8e-07	2.14e-06
d_Judicial Effectiveness	-5.439e-05	6.15e-05	-0.884	0.386	-0.000	7.29e-05
d_COVID_dummy	0.0045	0.004	1.040	0.309	-0.004	0.014
ECT_lag1	-0.0666	0.076	-0.877	0.390	-0.224	0.091
d_NPL_lag1	0.1742	0.205	0.848	0.405	-0.251	0.599
d_NPL_lag2	0.5541	0.198	2.803	0.010	0.145	0.963

Omnibus:	19.917	Durbin-Watson:	2.153
Prob(Omnibus):	0.000	Jarque-Bera (JB):	28.462
Skew:	-1.562	Prob(JB):	6.60e-07
Kurtosis:	6.308	Cond. No.	5.98e+05

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 5.98e+05. This might indicate that there are strong multicollinearity or other numerical problems.

Table 19 - Dynamic ECM estimation ARDL-based results - Binary Specification

The second lag of ΔNPL coefficient of 0.554 (t = 2.803, p = 0.010) appears to be the most important short-term determinant, pointing out the importance of the recent NPL dynamics. In contrast, other differenced explanatory variables, such as ΔCAR,

ΔUnemployment rate, ΔInterest rate, ΔReal GDP PT, ΔJudicial Effectiveness and ΔCOVID dummy, do not turn out to be statistically significant, implying that short-run macroeconomic or institutional shocks do not have much direct effect on NPL changes after controlling for lagged NPL effects.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  d_NPL   R-squared:                       0.417
Model:                            OLS   Adj. R-squared:                  0.189
Method:                 Least Squares   F-statistic:                     1.831
Date:                Sun, 07 Sep 2025   Prob (F-statistic):              0.117
Time:                        19:41:44   Log-Likelihood:                140.76
No. Observations:                  33   AIC:                            -261.5
Df Residuals:                      23   BIC:                            -246.6
Df Model:                           9
Covariance Type:            nonrobust
================================================================================================
                                   coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------------------------
const                          -0.0012      0.001     -0.859      0.399      -0.004       0.002
d_CAR                          -0.0012      0.002     -0.593      0.559      -0.005       0.003
d_Unemployment rate             0.0068      0.137      0.049      0.961      -0.277       0.291
d_Interest rate                 0.1143      0.224      0.510      0.615      -0.349       0.578
d_Real GDP PT                8.316e-07   6.34e-07      1.312      0.202    -4.8e-07    2.14e-06
d_Judicial Effectiveness    -5.439e-05   6.15e-05     -0.884      0.386      -0.000    7.29e-05
d_COVID_dummy                   0.0045      0.004      1.040      0.309      -0.004       0.014
d_NPL_lag1                      0.1742      0.205      0.848      0.405      -0.251       0.599
d_NPL_lag2                      0.5541      0.198      2.803      0.010       0.145       0.963
ECT_lag1                       -0.0666      0.076     -0.877      0.390      -0.224       0.091
==============================================================================
Omnibus:                       19.917   Durbin-Watson:                   2.153
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               28.462
Skew:                          -1.562   Prob(JB):                     6.60e-07
Kurtosis:                       6.308   Cond. No.                     5.98e+05
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 5.98e+05. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Table 20 - First differenced ECM and lagged ECT - Binary Specification

The ECT_lag1 is negative (~ -0.0666) and not significant (t = -0.877, p = 0.390), meaning that deviations from the long-run equilibrium occur, if at all, rather sluggishly in the short run. This is consistent with the persistence detected in the previous ARDL and VAR estimates and relates to the overwhelming weight of past NPL stock in the short-run dynamics.

The ECM results under the binary COVID dummy specification suggest that NPLs dynamics are mostly driven by their own past level, with pandemic shocks and macroeconomic contemporaneous dynamics having limited immediate impacts. This reflects the lagged nature of credit risk in the Portuguese banking system and the importance of taking lags of NPLs into account in short-run policy evaluations. These

results support those of Plikas, J. H. et al. (2024), who also find broadly low overall importance of external shocks.

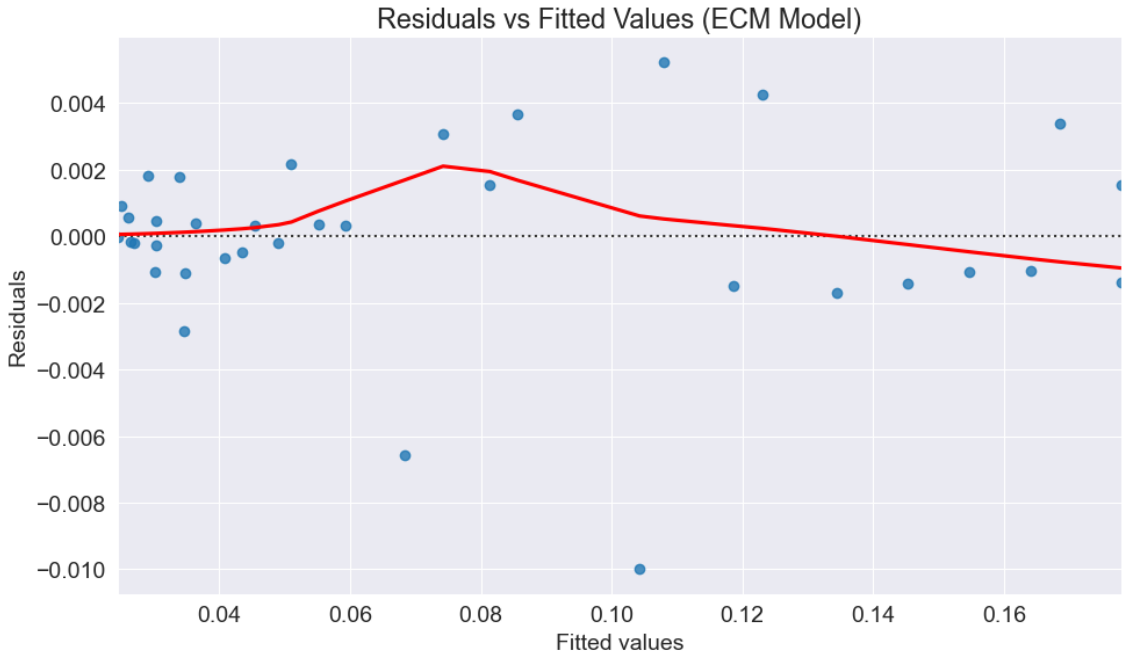


Figure 33 - Residuals Vs Fitted values - Binary Specification

The residuals against the fitted values scatter plot for the ECM model indicates that the residuals are predominantly distributed close to zero, which implies that the model captures a lot of the short-run dynamics of NPLs. However, the slight curve does indicate some kind of systematic bias, with under-prediction at mid-range fitted values to the over-prediction at the high end, and so mild non-linearity. The scatter of residuals is more spread out at larger fitted values; this may suggest heteroskedasticity, and a few outliers may be due to extreme shocks such as the one brought by the pandemic. All in all, the ECM works reasonably well.

D2) Covid dummy variable: Phase-by-Phase Specification

In the phase-specific specification, the COVID dummy is structured to capture discrete time phases of the pandemic and recovery. The dummy is equal to 1 during the following periods: COVID-2020 (1 March 2020 – 31 December 2020), COVID-2021 (1 January 2021 – 31 December 2021), COVID-2022 (1 January 2022 – 30 September 2022) and Post-COVID (1 October 2022 – 31 December 2024) while 0 otherwise. This temporal segmentation allows the model to account for the heterogeneous impact of the pandemic and the subsequent recovery of NPLs in the form of immediate, medium-term, and residual effects. By timing individual phases, the approach offers a more accurate description of structural shocks and policy responses over time, allowing for fine-tuning of credit risk behaviour during the crisis and in the post-crisis period.

Through correlation, in the phase-based specification it is possible to observe that the highlighted banking and macro-financial variables are largely intertwined, with CAR negatively correlated with NPL (-0.88) and LDR (-0.93) while they are positively correlated with Judicial Effectiveness (0.82), GNI (0.82), and House Pricing PT (0.84) signalling a multi-dimensional relationship between capital adequacy, institutional quality and macro-financial aspects. In addition, NPLs are substantially positively correlated with LDR (0.98) and Business Freedom (0.91); negatively correlated with Judicial Effectiveness (- 0.71) and Property Rights (-0.76), demonstrating that both bank-specific and systemic or institutional factors have an important role in affecting credit quality.

	Variable	VIF
0	const	3967.649234
1	PSI-20	5.729407
2	Real GDP PT	8.407705
3	Judicial Effectiveness	2.871774
4	LDR	8.781936
5	Inflation rate	2.224578
6	Unemployment rate	4.105687
7	COVID_dummy	3.979750

Table 21 - VIF results - Phase-by-phase Specification

Intercept: 0.08611038785714285

	Variable	Coefficient
3	FSI	0.045773
6	LDR	0.005346
21	Unemployment rate	0.002347
7	Business Freedom	0.001680
12	PGB 10Y	0.001184
0	CAR	-0.000895
14	GNI	0.000323
23	COVID_dummy	-0.000221
9	Judicial Effectiveness	-0.000199
16	Inflation rate	-0.000132
22	Public Debt PT	-0.000000
2	Firm Loan PT	-0.000000
20	Real GDP PT	0.000000
19	Exports PT M€	-0.000000
18	PCRH-D	0.000000
17	Interest rate	-0.000000
5	ROA	0.000000
4	Personal Loan PT	-0.000000
13	Exchange rate US to EU	-0.000000
1	Bank Size	0.000000
11	PSI-20	-0.000000
10	Property Rights	0.000000
8	Investment Freedom	0.000000
15	House Pricing PT	-0.000000

Table 22 – Elastic Net results - Phase-by-phase Specification

Multicollinearity continues to be a concern as evidenced by high VIFs for LDR (8.78), Real GDP PT (8.41) and the COVID dummy (3.98); however, the Elastic Net results highlight the strongest drivers of NPLs: FSI, LDR and Unemployment rate, alongside little effect from CAR and COVID dummy, highlighting the relative inactivity of NPLs to contemporaneous shocks. In this specification, phase-based dummy enables examining temporal enhancements in pandemic impacts, deciding that the structural breaks in the years of COVID-19 have a differential effect on credit risk, and post-pandemic, the behaviour of NPL has long-term management to historical levels and systemic financial stability measures.

In this phase-based specification, the VAR model was also estimated, incorporating the following endogenous variables: NPL, LDR, Unemployment rate, Judicial Effectiveness, Inflation rate, PGB 10Y and Business Freedom. The choices of variables combine bank-specific risk indicators (NPL, LDR), macro-financial conditions (Unemployment, Inflation, PGB 10Y), and institutional quality measures (Judicial Effectiveness, Business Freedom) to ensure that it presents the internal and external

determinants of banking sector stability in both correlating pairs, VIF, Elastic Net and Pearson correlations. The phase-based COVID dummy is included in the first block to take into account that the model must capture short-term shocks and lagged responses and to correct for structural shifts occurring during the pandemic phases.

```
Constant columns (will be removed): []
  Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Sun, 07, Sep, 2025
Time:                     21:16:30
--------------------------------------------------------------------
No. of Equations:         7.00000    BIC:                   -26.6625
Nobs:                     35.0000    HQIC:                  -28.4957
Log likelihood:           230.947    FPE:                1.74048e-13
AIC:                     -29.4621    Det(Omega_mle):     3.50733e-14
--------------------------------------------------------------------
Results for equation NPL
=============================================================================================
                              coefficient       std. error           t-stat            prob
---------------------------------------------------------------------------------------------
const                            0.047632         0.100432            0.474           0.635
COVID_dummy                      0.002723         0.002381            1.144           0.253
L1.NPL                           0.995358         0.161762            6.153           0.000
L1.LDR                          -0.000360         0.000755           -0.478           0.633
L1.Unemployment rate             0.101242         0.097394            1.040           0.299
L1.Judicial Effectiveness       -0.000080         0.000054           -1.463           0.144
L1.Inflation rate               -0.000996         0.035324           -0.028           0.977
L1.PGB 10Y                       0.001405         0.001057            1.330           0.184
L1.Business Freedom             -0.000313         0.000786           -0.398           0.691
=============================================================================================

Results for equation LDR
=============================================================================================
                              coefficient       std. error           t-stat            prob
---------------------------------------------------------------------------------------------
const                           83.093042        29.114077            2.854           0.004
COVID_dummy                     -1.085926         0.690271           -1.573           0.116
L1.NPL                         121.883669        46.892892            2.599           0.009
L1.LDR                           0.434461         0.218724            1.986           0.047
L1.Unemployment rate            17.003443        28.233571            0.602           0.547
L1.Judicial Effectiveness        0.003844         0.015782            0.244           0.808
L1.Inflation rate               -7.566809        10.239934           -0.739           0.460
L1.PGB 10Y                      -0.287264         0.306327           -0.938           0.348
L1.Business Freedom             -0.542883         0.227747           -2.384           0.017
=============================================================================================
```

Results for equation Unemployment rate

	coefficient	std. error	t-stat	prob
const	-0.160363	0.175262	-0.915	0.360
COVID_dummy	0.008712	0.004155	2.097	0.036
L1.NPL	-0.223826	0.282288	-0.793	0.428
L1.LDR	0.001085	0.001317	0.824	0.410
L1.Unemployment rate	0.286225	0.169962	1.684	0.092
L1.Judicial Effectiveness	-0.000246	0.000095	-2.588	0.010
L1.Inflation rate	0.031984	0.061643	0.519	0.604
L1.PGB 10Y	0.003299	0.001844	1.789	0.074
L1.Business Freedom	0.001734	0.001371	1.264	0.206

Results for equation Judicial Effectiveness

	coefficient	std. error	t-stat	prob
const	640.468826	350.858085	1.825	0.068
COVID_dummy	11.255962	8.318559	1.353	0.176
L1.NPL	1193.924589	565.113250	2.113	0.035
L1.LDR	-6.776776	2.635874	-2.571	0.010
L1.Unemployment rate	-421.583112	340.246982	-1.239	0.215
L1.Judicial Effectiveness	0.376799	0.190190	1.981	0.048
L1.Inflation rate	-228.669255	123.402970	-1.853	0.064
L1.PGB 10Y	0.340400	3.691595	0.092	0.927
L1.Business Freedom	-0.617568	2.744607	-0.225	0.822

Results for equation Inflation rate

	coefficient	std. error	t-stat	prob
const	-0.088298	0.315222	-0.280	0.779
COVID_dummy	0.020137	0.007474	2.694	0.007
L1.NPL	0.155885	0.507715	0.307	0.759
L1.LDR	-0.001525	0.002368	-0.644	0.519
L1.Unemployment rate	-0.222155	0.305688	-0.727	0.467
L1.Judicial Effectiveness	-0.000127	0.000171	-0.743	0.458
L1.Inflation rate	0.925542	0.110869	8.348	0.000
L1.PGB 10Y	-0.001749	0.003317	-0.527	0.598
L1.Business Freedom	0.002977	0.002466	1.207	0.227

Results for equation PGB 10Y

	coefficient	std. error	t-stat	prob
const	-4.014996	9.931402	-0.404	0.686
COVID_dummy	0.344040	0.235465	1.461	0.144
L1.NPL	-0.933187	15.996117	-0.058	0.953
L1.LDR	-0.039260	0.074611	-0.526	0.599
L1.Unemployment rate	1.694918	9.631044	0.176	0.860
L1.Judicial Effectiveness	-0.012391	0.005384	-2.302	0.021
L1.Inflation rate	11.884686	3.493049	3.402	0.001
L1.PGB 10Y	0.823416	0.104494	7.880	0.000
L1.Business Freedom	0.104166	0.077689	1.341	0.180

Results for equation Business Freedom

	coefficient	std. error	t-stat	prob
const	82.032788	27.813283	2.949	0.003
COVID_dummy	0.416023	0.659430	0.631	0.528
L1.NPL	103.749940	44.797755	2.316	0.021
L1.LDR	-0.396391	0.208951	-1.897	0.058
L1.Unemployment rate	-2.451341	26.972117	-0.091	0.928
L1.Judicial Effectiveness	0.001871	0.015077	0.124	0.901
L1.Inflation rate	-28.160660	9.782421	-2.879	0.004
L1.PGB 10Y	0.763730	0.292641	2.610	0.009
L1.Business Freedom	0.300986	0.217571	1.383	0.167

Correlation matrix of residuals

	NPL	LDR	Unemployment rate	Judicial Effectiveness	Inflation rate	PGB 10Y	Business Freedom
NPL	1.000000	0.313621	−0.122728	−0.217662	0.084614	0.062902	0.631331
LDR	0.313621	1.000000	−0.303831	−0.196077	−0.415042	−0.334328	−0.081684
Unemployment rate	−0.122728	−0.303831	1.000000	−0.212651	0.104613	−0.076631	−0.049026
Judicial Effectiveness	−0.217662	−0.196077	−0.212651	1.000000	0.321039	0.220689	0.265327
Inflation rate	0.084614	−0.415042	0.104613	0.321039	1.000000	0.371534	0.412117
PGB 10Y	0.062902	−0.334328	−0.076631	0.220689	0.371534	1.000000	0.078915
Business Freedom	0.631331	−0.081684	−0.049026	0.265327	0.412117	0.078915	1.000000

Table 23 - Vector Autoregression (VAR) results - Phase-by-phase Specification

The persistence of NPLs is very evident from the VAR results: the own lagged NPL coefficient is statistically significant (L1. NPL = 0.995, t = 6.15, p < 0.001), however, contemporary macro-financial and institutional time-varying variables such as the COVID dummy have small simultaneous impacts on NPL. LDR is controlled by the previous value of NPL (L1. NPL = 121.88, t = 2.60, p = 0.009), as well as its lag (L1. LDR = 0.434, t = 1.99, p = 0.047), and Business Freedom had significantly negative impact on LDR (t = −2.38, p = 0.017), indicating that the flexibility in regulation can change bank's liquidity management. A negative and significant effect is found in the unemployment rate with the immediate impact of the pandemic (COVID dummy = − 0.0087, t = − 2.10, p = 0.036) and judicial quality (L1. Judicial Effectiveness = -0.00025, t = -2.59, p = 0.010), further supporting that there is an association between institutional setting and labour market resilience. As past NPLs' Judicial Effectiveness reacts to increases in NPL (1193.92, t = 2.11, p = 0.035) and LDR stress (L1. LDR = -6.78, t = -2.57, p = 0.010), indicating institutional responses to banking industry pressures. There is a high autocorrelation of inflation and PGB 10Y (L1. Inflation = 0.93), (L1. PGB 10Y=0.82), but COVID shocks are temporarily raising inflation. Business Freedom seems to be sensitive to economic strain (L1. NPL = 103.75, t = 2.32, p = 0.021), inflation, and the signal from the bond market, regulatory responses to macro-financial contexts.

In comparison with VAR model on the original data of the data frame with CAR, Interest rate serve as the main banking risk indicator instead of Business Freedom and NPL, the phase model contains the same views as the VAR model in modelling the persistence of banking and macro-financial variables, but the latter more easily showcase that the specific effects of pandemic to unemployment and inflation. The residual correlations are moderate, and the model is quite robust for short-term dynamics given

the small number of observations (N=35), validating the assumptions that NPLs are mostly inertia-driven and that institutional and macro-financial shocks take time to generate their full impact.

Following the variable selection from VIF, Elastic Net and VAR, an ARDL model was used to analyse the short-run and long-run determinants of NPLs in Portugal. The dependent variable is the NPL ratio, independent ones are important bank-specifi, macroeconomic, institutional and market variables and an exogenous COVID dummy to account for the shock from the pandemic. Variable inclusion was cautiously informed by multicollinearity problems, predictive relevance from VIF, Elastic Net, and VAR evidence of dynamic linkages across assets, excluding near-perfectly collinear variables such as FSI.

```
Dependent var in columns: True
Independent vars in columns: True

Optimal ARDL order using AIC: (2, 3, 2, 1)
Summary of ARDL model selected by AIC:
                           ARDL Model Results
==============================================================================
Dep. Variable:                    NPL   No. Observations:                   36
Model:               ARDL(2, 3, 2, 1)   Log Likelihood                165.029
Method:               Conditional MLE   S.D. of innovations             0.002
Date:                Sun, 07 Sep 2025   AIC                         -304.058
Time:                        21:16:31   BIC                         -284.215
Sample:                  12-31-2016   HQIC                        -297.291
                       - 12-31-2024
================================================================================================
                                   coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------------------------
const                           -0.0250      0.042     -0.602      0.553      -0.111       0.061
NPL.L1                           0.4602      0.168      2.745      0.012       0.112       0.808
NPL.L2                           0.4207      0.160      2.627      0.015       0.089       0.753
Unemployment rate.L0             0.1131      0.094      1.201      0.243      -0.082       0.308
Unemployment rate.L1            -0.0285      0.086     -0.331      0.744      -0.207       0.150
Unemployment rate.L2             0.1812      0.085      2.132      0.044       0.005       0.357
Unemployment rate.L3             0.1348      0.080      1.682      0.107      -0.031       0.301
CAR.L0                          -0.0018      0.001     -1.689      0.105      -0.004       0.000
CAR.L1                          -0.0007      0.001     -0.554      0.585      -0.003       0.002
CAR.L2                           0.0034      0.001      2.343      0.029       0.000       0.006
LDR.L0                           0.0020      0.001      3.002      0.007       0.001       0.003
LDR.L1                          -0.0021      0.001     -3.825      0.001      -0.003      -0.001
================================================================================================

Optimal ARDL order using BIC: (1, 1)
Summary of ARDL model selected by BIC:
                           ARDL Model Results
==============================================================================
Dep. Variable:                    NPL   No. Observations:                   36
Model:                  ARDL(1, 1)   Log Likelihood                147.130
Method:               Conditional MLE   S.D. of innovations             0.004
Date:                Sun, 07 Sep 2025   AIC                         -284.260
Time:                        21:16:32   BIC                         -276.483
Sample:                  06-30-2016   HQIC                        -281.575
                       - 12-31-2024
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0035      0.031      0.112      0.912      -0.060       0.067
NPL.L1         0.9762      0.072     13.547      0.000       0.829       1.123
LDR.L0         0.0013      0.001      1.817      0.079      -0.000       0.003
LDR.L1        -0.0013      0.001     -1.875      0.070      -0.003       0.000
==============================================================================
```

Table 24 - ARDL Model results - Phase-by-phase Specification

The ARDL estimation delivered two specifications of the model. The AIC-optimal ARDL (2,3,2,1) has 2 NPL lags, 3 Unemployment rate lags, 2 CAR lags and 1 LDR lags, to effectively account for the short- and medium-term dynamics. Significance of the coefficients of NPL lags are positive and statistically significant (NPL. L1 = 0.460, p = 0.012; NPL. L2 = 0.421, p = 0.015), indicating high persistence in credit risk. LDR has a direct beneficial impact on NPL (LDR. L0 = 0.002, p = 0.007) as well as a reverse lagged effect (LDR. L1 = -0.002, p = 0.001), which would indicate risk as well as corrections. Unemployment and CAR contribute with lagged dynamics, whereas all other short-term coefficients are largely irrelevant. On the other hand, the BIC-optimal ARDL

(1,1) is less discrete, comprising NPL.L1 (0.976; $p < 0.001$) and LDR lags, further indicating the strong past influence of NPL on the current level, leading to a strong, simpler form, robust specification for the small sample.

Taken together, the ARDL findings indicate persistence of NPLs, short-term effects of bank leverage, and lag effects of macroeconomic and bank health variables, providing supplementary evidence to the phase-dependent VAR analysis and supporting both credit risk management and policy evaluation.

Next, the estimation of Elastic Net was performed by with the ARDL framework in this extended study with the addition of a COVID dummy as well as a phase-based treatment.

Variable	Elasticity
NPL.L1	0.46020
NPL.L2	0.42070
Unemployment rate.L0	0.15834
Unemployment rate.L1	-0.03990
Unemployment rate.L2	0.25368
Unemployment rate.L3	0.18872
CAR.L0	-0.00540
CAR.L1	-0.00210
CAR.L2	0.01020
LDR.L0	0.03600
LDR.L1	-0.03780

Table 25 – Elastic Net ARDL-based - Phase-by-phase Specification

Based on the Elastic Net ARDL's results, credit risk is largely influenced by lagged NPL, with NPL. L1 and NPL. L2 elasticities of 0.460 and 0.421, respectively, thus supporting high persistence in NPLs. Of the macroeconomic variables, Unemployment stands out with a mixed picture, overall being significantly positive with elasticities varying between around -0.040 (L1) and 0.254 (L2) and 0.189 (L3), suggesting that labour market stress impacts on NPLs with a lag. As for bank-specific variables, it presents more complex results: LDR generates positive contemporaneous impact (L0 = 0.036) and negative simultaneous impact (L1 = -0.038), which seems to imply that more leveraging finances leads to an increase in credit risk first, but over time, it can make banks adjust their behaviour to reduce this risk. So does CAR have a little contemporaneous negative effect (L0 = -0.005) and a positive lagged effect (L2 = 0.010) describing the adjustment of bank capital adequacy to and credit risk. Overall, these

findings emphasise the high degree of persistence of NPLs, the short-run sensitivity to a change in bank leverage, and the long-run but significant effects of macroeconomic as well as bank health indicators on NPLs.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    NPL   R-squared:                       0.998
Model:                            OLS   Adj. R-squared:                  0.997
Method:                 Least Squares   F-statistic:                     1074.
Date:                Sun, 07 Sep 2025   Prob (F-statistic):           2.71e-26
Time:                        21:16:32   Log-Likelihood:                 159.68
No. Observations:                  33   AIC:                            -295.4
Df Residuals:                      21   BIC:                            -277.4
Df Model:                          11
Covariance Type:            nonrobust
================================================================================
                               coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
const                    -0.0250      0.042     -0.602      0.554      -0.112       0.061
NPL_L1                    0.4602      0.168      2.745      0.012       0.112       0.809
NPL_L2                    0.4207      0.160      2.627      0.016       0.088       0.754
Unemployment rate_L0      0.1131      0.094      1.201      0.243      -0.083       0.309
Unemployment rate_L1     -0.0285      0.086     -0.331      0.744      -0.207       0.150
Unemployment rate_L2      0.1812      0.085      2.132      0.045       0.004       0.358
Unemployment rate_L3      0.1348      0.080      1.682      0.107      -0.032       0.301
CAR_L0                   -0.0018      0.001     -1.689      0.106      -0.004       0.000
CAR_L1                   -0.0007      0.001     -0.554      0.585      -0.003       0.002
CAR_L2                    0.0034      0.001      2.343      0.029       0.000       0.006
LDR_L0                    0.0020      0.001      3.002      0.007       0.001       0.003
LDR_L1                   -0.0021      0.001     -3.825      0.001      -0.003      -0.001
==============================================================================
Omnibus:                        6.163   Durbin-Watson:                   2.122
Prob(Omnibus):                  0.046   Jarque-Bera (JB):                4.873
Skew:                          -0.918   Prob(JB):                       0.0875
Kurtosis:                       3.413   Cond. No.                     7.22e+04
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 7.22e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Table 26 - OLS regression model ARDL-based - Phase-by-phase Specification

The OLS estimates confirm the existence of significant persistence in the NPL series, as the coefficients of NPL_L1 and NPL_L2 (0.4602 and 0.4207, respectively) are significant at the 5% level, thus demonstrating that past non-performing loans have a strong effect on present credit risk. Of the macro variables, only Unemployment_L2 is significant (0.1812, p = 0.045) and is indicative of the lagged response to labour market distress. Bank-specific indicators exhibit short-term dynamics: LDR_L0 is positive (0.0020, p = 0.007) and LDR_L1 is negative (-0.0021, p = 0.001), signalling corrective behaviour, while CAR effects are overall arbitrary. The COVID dummy is small and insignificant, especially when NPL persistence is considered. The model explains 99.8% of the variance ($R^2 = 0.998$). The condition number is 7.22e+04, suggesting that there might be multicollinearity in the model. Overall, the findings also reflect a high

(macro)inertia of NPLs and a combination of macro-effects with bank-behaviour adjustments in the short-run.

The results indicate that NPL persistence remains the major force moulding credit risk, and lagged coefficients are still significantly high, as evident from Elastic Net ARDL and OLS estimation. Influence from the macroeconomic indicators, particularly the Unemployment rate, is also transmitted with a lag, while bank-specific variables such as LDR and CAR affect short-run dynamics on an instant and lag basis. With the other NPL variables included in the model, the COVID dummy hardly has any direct effect.

OLS Regression Results

Dep. Variable:	d_NPL	R-squared:	0.676
Model:	OLS	Adj. R-squared:	0.471
Method:	Least Squares	F-statistic:	3.303
Date:	Sun, 07 Sep 2025	Prob (F-statistic):	0.00990
Time:	22:39:46	Log-Likelihood:	145.39
No. Observations:	32	AIC:	-264.8
Df Residuals:	19	BIC:	-245.7
Df Model:	12		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.0008	0.001	-0.538	0.597	-0.004	0.002
d_Unemployment rate	0.0309	0.053	0.578	0.570	-0.081	0.143
d_CAR	-0.0007	0.001	-0.849	0.406	-0.003	0.001
d_LDR	0.0009	0.000	2.149	0.045	2.32e-05	0.002
d_NPL_lag1	0.2210	0.194	1.141	0.268	-0.184	0.626
d_NPL_lag2	0.4098	0.168	2.446	0.024	0.059	0.760
d_Unemployment rate_lag0	0.0309	0.053	0.578	0.570	-0.081	0.143
d_Unemployment rate_lag1	-0.0196	0.124	-0.159	0.876	-0.278	0.239
d_Unemployment rate_lag2	0.1761	0.129	1.370	0.187	-0.093	0.445
d_Unemployment rate_lag3	0.1810	0.117	1.553	0.137	-0.063	0.425
d_CAR_lag0	-0.0007	0.001	-0.849	0.406	-0.003	0.001
d_CAR_lag1	-0.0009	0.002	-0.493	0.628	-0.004	0.003
d_CAR_lag2	0.0030	0.002	1.717	0.102	-0.001	0.007
d_LDR_lag0	0.0009	0.000	2.149	0.045	2.32e-05	0.002
d_LDR_lag1	-0.0015	0.001	-1.716	0.102	-0.003	0.000
ECT_lag1	-0.0427	0.096	-0.445	0.662	-0.244	0.158

Omnibus:	11.154	Durbin-Watson:	2.184
Prob(Omnibus):	0.004	Jarque-Bera (JB):	10.578
Skew:	-1.093	Prob(JB):	0.00505
Kurtosis:	4.777	Cond. No.	1.18e+17

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The smallest eigenvalue is 1.07e-32. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.

Table 27 - Dynamic ECM estimation ARDL-based results - Phase-by-phase Specification

After that, the ECM was estimated with phase-based specification to capture the short-run dynamics of NPLs. The model accounts for a moderate amount of the variation in first differences of NPLs ($R^2 = 0.676$), with the adjusted R-squared (0.471) indicating that a large fraction of short-term fluctuations still go unexplained, plausibly related to

the persistence of credit risk as well as the complexity of interactions between the regressors. The condition number is very high (~1.18e+17), indicating strong multicollinearity or other numerical problems in the design matrix, and diagnostic statistics (Durbin-Watson = 2.184; normality check for the residuals, including the Omnibus and Jarque-Bera tests) show relatively small departures from normality, a common feature that happens in financial time series.

The second lag of ΔNPL remains the most relevant short-run driver, with a coefficient of 0.410 (t = 2.446, p = 0.024), highlighting the persistence of NPL changes over time. Additionally, the contemporaneous change in LDR shows a positive and significant effect (0.0009, t = 2.149, p = 0.045), indicating that short-term increases in bank leverage immediately raise credit risk. In contrast, the error correction term (ECT_lag1) is negative as expected (-0.043, p = 0.662) but not statistically significant, suggesting that adjustment toward the long-run equilibrium is weak in this sample. Other short-run drivers, including ΔUnemployment rate, ΔCAR, and lagged differences of these variables, are generally insignificant, reinforcing the dominant role of past NPL dynamics and bank-level leverage in shaping short-term NPL fluctuations.

OLS Regression Results

Dep. Variable:	d_NPL	R-squared:	0.490
Model:	OLS	Adj. R-squared:	0.437
Method:	Least Squares	F-statistic:	9.285
Date:	Sun, 07 Sep 2025	Prob (F-statistic):	0.000184
Time:	21:16:32	Log-Likelihood:	142.95
No. Observations:	33	AIC:	-277.9
Df Residuals:	29	BIC:	-271.9
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.0009	0.001	-0.999	0.326	-0.003	0.001
d_LDR	0.0019	0.001	3.130	0.004	0.001	0.003
d_NPL_lag2	0.4294	0.136	3.149	0.004	0.150	0.708
ECT_lag1	-0.1143	0.070	-1.623	0.115	-0.258	0.030

Omnibus:	11.040	Durbin-Watson:	1.627
Prob(Omnibus):	0.004	Jarque-Bera (JB):	10.890
Skew:	-1.024	Prob(JB):	0.00432
Kurtosis:	4.929	Cond. No.	359.

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Table 28 - First differenced ECM and lagged ECT - Phase-by-phase Specification

The percentage adjustment parameters (ECT_lag1) are found to have a negative value (~ -0.114), which is statistically insignificant (t = -1.623, p = 0.115), indicating that the short-run partial adjustment to the long-run equilibrium is only partial. It is consistent with the persistence found in the ARDL and Elastic Net analyses and confirms the primary importance of past NPL levels in shaping the short-run dynamic. In the short term, factors influencing ΔNPL are mainly the second lag of ΔNPL (0.429, t = 3.149, p = 0.004) and simultaneous LDR (0.0019, t = 3.130, p = 0.004), suggesting that both past loan quality and short-term bank leverage are dominant determinants of changes in NPLs.

Overall, the simplified phase-based ECM results emphasise that NPL dynamics are largely driven by their past, while possible contemporaneous macroeconomic or institutional modifications have relatively limited short-run effects. This contribution highlights the persistent presence of credit risk in the Portuguese banking system and the relevance of lagged NPLs in short-run policy considerations.

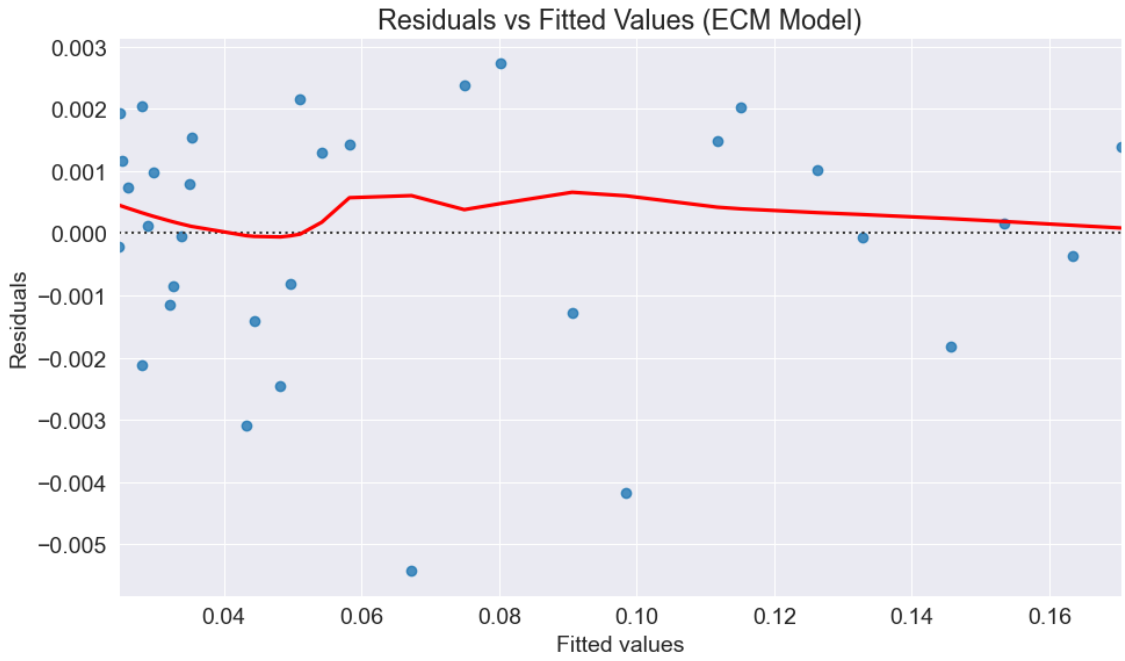


Figure 34 - Residuals Vs Fitted values - Phase-by-phase Specification

Notice that the residuals in the residual versus fitted values plot for the ECM model are generally centred at zero, indicating that the model explains a good amount of the short-run dispersion in NPLs. The global scatter implies no extreme bias, while the red smoothing curve does indicate slight non-linear deviations in the sense of a curve

below smaller fitted values. This indicates a slight upward bias in the predictions in the low end, and when the line flattens out at higher fitted values, there are fewer systematic errors in that range. The spread of residuals is also kind of wider at lower fitted values, which is an indication of mild heteroscedasticity, but it is not so evident. There are a few apparent outliers, which might be interpreted as severe shocks, as with the COVID-19 pandemic. In general, the ECM is quite well, neutral, but further improvements by using robust standard errors, extending explanatory variables, or other non-linear dynamics can lead to better robustness and predictive accuracy.

The alternative COVID dummy specifications offer further insights into NPL dynamics during the pandemic. The binary specification emphasises the role of interest rates and structural shocks, but once past NPLs are added to the equation, it loses some explanatory power, which is a confirmation of the prevailing presence of persistence. The phase-based one, which considers different periods of the pandemic and accounts for unemployment and liquidity, provides a more detailed picture; however, it is sensitive to multicollinearity and one-step error correction.

Overall, the original specification is a reliable and complete framework that maintains a balanced explanatory power with simplicity. While phase-based models give a sense of temporal dynamics, a binary COVID specification signals structural shocks. In general, the evidence confirms that NPL persistence and bank-specific factors drive Portuguese credit risk; macroeconomic and pandemic shocks are secondary to this, usually only temporary phenomena.