Essays on Systemic Risk and Banking

Resumo

Esta tese de doutorado é composta por três ensaios independentes sobre risco sistêmico e economia bancária. O primeiro artigo busca mensurar risco bancário utilizando o modelo estrutural de Merton (1974) e a métrica Z-Score para avaliar o impacto da regulação de capital na probabilidade de falência (PD) bancária. Os resultados confirmam a importância de medidas regulatórias como o acordo de Basileia III e destacam a necessidade de requerimentos de capital otimizados para fortalecer a estabilidade financeira. O segundo artigo estima diferentes medidas para entender a contribuição de cada banco para o risco sistêmico do mercado brasileiro e propõe um modelo de corrida bancária que considera a PD idiossincrática dos bancos e um processo de risco sistêmico no qual quebras adicionais ocorrem através de diferentes canais de contágio. Por meio de nossa abordagem, estimamos a distribuição de perdas, a PD do fundo garantidor de depósitos e um nível otimizado de requerimento de capital para o sistema bancário brasileiro. Por fim, o terceiro artigo propõe o procedimento de Estimativa Bootstrap com Seleção de Variáveis (BEVS) para estimar os determinantes da PD no sistema bancário brasileiro. Neste método, combinamos técnicas como regressão Lasso, Loess e bagging, mostrando que esta abordagem integrada gera resultados superiores se comparado aos obtidos pelo desempenho individual de cada técnica. Nossos resultados indicam que o BEVS não apenas refina a estimativa da PD, mas também oferece uma visão mais abrangente do impacto dos fatores macroeconômicos durante o período avaliado.

Palavras-chave: Risco Sistêmico; Corrida Bancária; Estabilidade Financeira; Fundo Garantidor de Depósitos; Probabilidade de Falência; Instituições Financeiras; Requerimento de Capital.

Abstract

This doctoral dissertation consists of three self-contained essays on systemic risk and banking. The first paper examines bank risk using the structural model of Merton (1974) and Z-Score to evaluate the impact of bank capital regulation on banks' probability of default (PD). The results confirm the importance of regulatory measures such as Basel III and highlight the need for balanced capital requirements to enhance financial stability. The second paper estimates different measures to understand how much systemic risk each bank brings to the Brazilian market and proposes a bank run model that accounts for idiosyncratic PD of banks and a systemic risk process in which additional defaults occur through different channels of contagion. Through our approach, we estimate the loss distribution, the PD of the deposit insurance agency, and an optimized capital adequacy ratio for the Brazilian banking system. Lastly, the third paper proposes the Bootstrap Estimator with Variable Selection (BEVS) procedure to estimate the determinants of the PD in the Brazilian banking system. In this method, we combine techniques such as Lasso regression, Loess smoothing, and bagging, showing that this integrated approach yields improved results compared to those obtained through their individual performance. Our findings indicate that BEVS not only refines the estimate of PD but also offers a comprehensive view of the impact of macroeconomic factors over the study period.

Keywords: Systemic Risk; Bank Runs; Financial Stability; Deposit Insurance Agency; Probability of Default; Financial Institutions; Capital Requirement.

List of Figures

List of Tables

List of Codes

Contents

General Introduction

The recent Global Financial Crisis (GFC) has emphasized the importance of policies that improve the overall stability of the financial system. This crisis served as one of the clearest illustrations in history of systemic risk, in which banks and credit play a particularly important role (De Bandt and Hartmann, 2019). Although the topic has generated extensive literature and intense financial interest, many macroeconomists and policymakers recognize a significant gap in understanding the channels of system-wide risk and the contribution of systemic risk by individual financial institutions (FIs)[1] to the broader economy (Christiano et al., 2018; Fidrmuc and Lind, 2020). Furthermore, the critical nature of this issue has generated a growing consensus among policymakers to adopt a macroprudential approach to regulation and supervision, as it is considered essential to ensure a resilient global financial system (Hannoun, 2010; Borio, 2011; Galati and Moessner, 2013).

In broader terms, systemic risk refers to the risk of a financial crisis or market failure that affects the stability of the financial system and has widespread effects on the economy as a whole. This type of risk is of particular concern in financial systems due to the interconnectedness of financial institutions, which can amplify the effects of individual failures and propagate shocks throughout the system as a negative domino effect (Brunnermeier and Oehmke, 2013; Adrian and Brunnermeier, 2016).

Understanding how financial institutions affect systemic risk, whether through their idiosyncratic characteristics or its connections with the rest of the economy, is fundamental for effective action by central banks and policymakers in time of crisis. Furthermore, identifying the factors that influence the probability of default (PD) of banks is important

[1] In this thesis, the terms "bank" and "financial institution" are used interchangeably, even though banks can be viewed as a subset of the financial services sector. In this narrower sense, banks are financial institutions that accept deposits into various savings and demand deposit accounts, a service that non-banking financial institutions (such as investment banks, leasing companies, insurance companies, investment funds, finance firms, and others) cannot offer. For more information, see Hagendorff (2019).

for assessing and measuring the impact of regulatory policies on these metrics. Although a substantial body of literature derives PD estimates based on stock market data, such as prices, bonds, derivatives, and credit default swaps (Altman, 1968; Merton, 1974; Lehar, 2005; Davydov et al., 2021; da Rosa München, 2022), these methods are primarily applicable to listed banks. This creates significant difficulties for economies with fewer publicly traded banks, as is often the case in several emerging markets. In that sense, one of the pressing challenges in empirical banking literature is to create robust measures based on publicly balance sheet data, which can be used for both listed and non-listed banks (Souza et al., 2015; Guerra et al., 2016; Silva et al., 2016).

In Brazil, as in several other economies, non-listed banks constitute the majority of the financial system and play an important role in supporting both retail and commercial activities (BCB, 2022a). As of September 2022, only 24 (13.2%) of the Brazilian member institutions covered by the private deposit insurance agency (DIA), *Fundo Garantidor de Créditos* (FGC), are listed on the stock exchange. Considering that these institutions account for 80% of total assets, one could argue that concentrating the analysis on these listed banks would offer a reasonable representation in terms of total assets. On the other hand, such a focus overlooks the broader spectrum of institutions and their interconnections within the system. However, it is clear that the absence of market data for non-listed banks amplifies the complexity of risk assessment and, consequently, the evaluation of regulatory impacts on the banking system.

An extension of this challenge lies not just in the numerical distinction between non-listed and listed institutions, but also in the different regulatory guidelines set forth by the Central Bank of Brazil (BCB) based on the Basel III Accord (BCBS, 2011). Larger FIs, which are typically the listed ones, are subject to a more stringent set of regulations due to their substantial exposure size. In contrast, smaller banks, often non-listed, face a different set of regulatory standards. Although there exists a large literature exploring Basel III's impact on banks' balance sheets and its transmission to the real economy, especially since the GFC, the topic continues to be a subject of active discussion and research. Despite the broad consensus on the benefits of Basel III in enhancing financial stability, there are unresolved questions about the effects on capital cost, banking lending, and banking concentration, specifically regarding the trade-offs faced by regulators when establishing regulatory frameworks.

This context sets the stage for the research and the questions addressed in this thesis, which consists of three self-contained essays on systemic risk and banking. The database of all Brazilian FIs and the method to estimate the probability of default are the common factors for the three essays. Specifically, in Chapter 1 we examine bank risk using the structural model of Merton (1974) and Z-Score, employing public balance sheet data to evaluate the impact of bank capital regulation on banks' probability of default. To achieve this, we utilized the two-way fixed effects model and a difference-in-difference approach, analyzing data from the Brazilian banking market from December 2000 to September 2022. The results confirm the practical importance of regulatory measures such as Basel III and highlight the need for balanced capital requirements to enhance financial stability.

In Chapter 2, we estimate different measures to understand how much systemic risk each bank brings to the Brazilian market and proposed a bank run model that accounts for idiosyncratic probability of default of banks and a systemic risk process in which additional defaults occur through different channels of contagion. Our approach follows a similar theoretical framework in which portfolio risk is calculated in banking organizations and in which banking losses are estimated in deposit insurance schemes. Through the application of our model to a reduced sample of 24 listed banks and a full sample of 182 covered banks for September 2022, we estimate the loss distribution, the PD of the deposit insurance agency, and the optimized capital adequacy ratio of the Brazilian banking system.

Lastly, in Chapter 3, we propose the Bootstrap Estimator with Variable Selection (BEVS) procedure to estimate the determinants of the probability of default in the Brazilian banking system over the period from December 2007 to September 2022. In this method, we combine techniques such as Lasso regression, Loess smoothing, and bagging, showing that this integrated approach yields improved results compared to those obtained through their individual performance. Our findings indicate that BEVS not only refines the estimate of PD but also offers a comprehensive view of the impact of macroeconomic factors over the study period.

Chapter 1

Impact of Capital Regulation on Banks' Probability of Default

1.1 Introduction

Financial instability and crises are a recurrent, though infrequent, phenomenon in history. Since the Global Financial Crisis, researchers and policymakers have dedicated a great effort to understand the causes of such fragility and proposed strategies to mitigate future downturns (De Bandt and Hartmann, 2019). One strategy that has gained significant attention requires financial institutions to hold substantially more capital, which would lead to a lower probability of default given the greater capacity to absorb losses (BCBS, 2011; Berger and Bouwman, 2013; Thakor, 2014; Van Der Weide and Zhang, 2019; BCBS, 2021).

Identifying the factors that influence the PD of banks is essential to properly evaluate and measure the regulatory impact on these metrics. Although a substantial body of literature derives PD estimates based on stock market data, such as prices, bonds, derivatives, and credit default swaps (Altman, 1968; Merton, 1974; Lehar, 2005; Kato and Hagendorff, 2010; Chiaramonte and Casu, 2013; Davydov et al., 2021; da Rosa München, 2022), these methods are primarily applicable to listed banks. This creates significant difficulties for economies with fewer publicly traded banks, as is often the case in several emerging markets. In that sense, one of the pressing challenges in empirical banking literature is to create robust measures based on balance sheet data, which can be used for both listed and non-listed banks (Souza et al., 2015; Guerra et al., 2016; Silva et al., 2016).

In Brazil, as in several other economies, non-listed banks constitute the majority of the financial system and play an important role in supporting both retail and commercial activities (BCB, 2022a). As of September 2022, only 24 (13.2%) of the Brazilian member institutions covered by the private deposit insurance agency, *Fundo Garantidor de Créditos*, are listed on the stock exchange. Considering that these institutions account for 80% of total assets, one could argue that concentrating the analysis on these listed banks would offer a reasonable representation in terms of total assets. On the other hand, such a focus overlooks the broader spectrum of institutions and their interconnections within the system. However, it is clear that the absence of market data for non-listed banks amplifies the complexity of risk assessment and, consequently, the evaluation of regulatory impacts on the banking system.

An extension of this challenge lies not just in the numerical distinction between non-listed and listed institutions, but also in the different regulatory guidelines set forth by the Central Bank of Brazil based on the Basel III Accord (BCBS, 2011). Larger FIs, which are typically the listed ones, are subject to a more stringent set of regulations due to their substantial exposure size. In contrast, smaller banks, often non-listed, face a different set of regulatory standards. This differentiation is outlined in the CMN resolution no. 4.553/2017, segmenting banks from S1 to S5 based on their total exposures and defining their respective rules[1]. This approach mirrors the principles emphasized by Tarullo (2019), where the post-2008 US banking regulations, such as the Dodd-Frank Act, formulated guidelines that corresponded to the size, risk profile, and systemic importance of banks.

While there are numerous studies assessing regulatory impacts on the Brazilian banking system, to the best of our knowledge, there is a lack of research encompassing all financial institutions using public data. To address this gap, we employed public balance sheet data to estimate bank risk using the structural model of Merton (1974) and the Z-Score (Souza et al., 2016). Furthermore, we evaluated the effect of bank capital regulation on a bank's probability of default. To achieve this, we utilized the two-way fixed effects and

[1] In accordance with Resolution CMN n.º 4.553/2017, the Central Bank of Brazil segments financial institutions and other licensed entities into five categories: S1 includes universal banks, commercial banks, investment banks, foreign exchange banks, and federal savings banks with a size equal to or exceeding 10% of the GDP or those engaging in relevant international activity; S2 encompasses the same bank categories as S1 but with a size less than 10% and equal to or exceeding 1% of the GDP, and other institutions with a size equal to or exceeding 1% of the GDP; S3 is for institutions with a size below 1% but equal to or exceeding 0.1% of the GDP; S4 consists of institutions with a size less than 0.1% of the GDP; and S5 includes institutions below 0.1% of the GDP that utilize a simplified optional methodology for determining minimum equity requirements, excluding the banks listed in S1 and S2.

the difference-in-difference model, analyzing the Brazilian banking market from December 2000 to September 2022.

The structure of this work consists of five sections. Following this introduction, Section 1.2 outlines the theoretical framework of all our analyzes to estimate the impact of capital regulation on banks' probability of default. Section 1.3 presents the data used in our models. Section 1.4 offers our results and a discussion, while Section 1.5 concludes with final remarks.

1.2 Theoretical Framework

This section presents the theoretical framework for estimating the probability of default and Z-Score for each financial institution, including the application of an econometric model used for inference.

1.2.1 Structural Model for Probability of Default

In this work, we utilize the structural model of Merton (1974) to estimate the idiosyncratic PD for each financial institution[2], which models credit risk using the Contingent Claim Analysis[3] (Souza et al., 2015, 2016; Guerra et al., 2016; Coccorese and Santucci, 2019; da Rosa München, 2022). It is a structural model because it establishes a relationship between the debt and the value of the bank. The intuition of this approach is to consider the bank's assets as the underlying asset of a European call option, with a strike price equal to its obligations and a time to maturity T. Thus, if the bank defaults, equity holders receive nothing because the bank does not have enough resources to repay its

[2] There are works that utilize approaches such as CAMELS (Capital adequacy, Asset quality, Management, Earnings, Liquidity, Sensitivity to market risk) to estimate the PD using balance sheet variables (Valahzaghard and Bahrami, 2013; Calabrese and Giudici, 2015; Rosa and Gartner, 2018; Parrado-Martínez et al., 2019). Although this technique allows the utilization of more granular and specific bank variables, it performs better when a large number of observable bank defaults are available for robust logit model estimation, which is not the case for the Brazilian economy. Even with strategies to increase the default variable, such as considering interventions by supervisors, capital below the minimum required, or mergers motivated by financial difficulties (Vazquez and Federico, 2015), criteria that have an inherent subjective aspect, we may still not reach a sufficient number of defaults in certain economies to estimate a robust model.

[3] Contingent Claim Analysis (CCA) is a generalization of the option pricing theory presented in Black and Scholes (1973) to analyze the corporate capital structure. A contingent claim is an asset whose future payoff depends on the outcome of an uncertain event. Thus, CCA analyzes how the value of the contingent claim changes as the value of the firm fluctuates over time. For a detailed examination, see Jobst and Gray (2013).

obligations. Otherwise, if it does not default, the equity holders receive the difference between the values of assets and liabilities.

Although Merton (1974)'s model establishes that a default occurs when the bank's assets (granted loans) fall below its obligations (deposits received), in reality, operations can continue with negative equity[4]. This is due to contract breach or liquidity scarcity problems when the bank must sell assets or renegotiate debt (Guerra et al., 2016). To address these concerns, the literature introduces a threshold called distress barrier (DB) as a trigger for default, defined as a proportion of the debt's face value. Specifically, the bank defaults if its asset value drops below its DB. This threshold is calculated using accounting data based on the KMV[5] model, as given by equation 1.1 (Crosbie and Bohn, 2003).

$$DB_{it} = STD_{it} + \alpha_{it} LTD_{it} \tag{1.1}$$

In which STD_{it} and LTD_{it} stands for the sort-term (maturity $\leq$ 1 year) and long-term (maturity greater than 1 year) liabilities, respectively. The parameter $0 \leq \alpha_{it} \leq 1$ proxies the share of a bank's long-term liabilities subject to early redemption under stress. Due to the unavailability of time to maturity data of total liabilities for the Brazilian case, we follow Souza et al. (2016) and assume that they are predominantly short-term debts ($STD_{it} = 0.7$) with a significant long-term debt share ($LTD_{it} = 0.3$). In general, the literature suggests $\alpha_{it} = 0.5$ if $LTD_{it}/STD_{it} < 1.5$, resulting in $DB_{it} = 0.85TL_{it}$ (Crosbie and Bohn, 2003; Souza et al., 2015; Guerra et al., 2016; Coccorese and Santucci, 2019).

Using these definitions on the Black and Scholes (1973)'s model, the option's payoff for the equity holder at time T is given by 1.2.

$$E_{it} = max(A_{it}\, \mathcal{N}(d_{1it}) - DB_{it}\, e^{-r_t T} \mathcal{N}(d_{2it}), 0) \tag{1.2}$$

Where A_{it} is the value of assets, r_t is the risk-free interest rate, $\mathcal{N}(.)$ is the cumulative normal distribution function,

[4]11 banks operated with negative equity in the Brazilian market between March 2000 and September 2022 (BCB, 2023a).

[5]Originally an acronym for its founders Kealhofer, McQuown, and Vasicek, KMV is known for its development of a credit risk model that estimates a firm's probability of default. KMV was acquired by Moody's Corporation in 2002, expanding Moody's Analytics credit risk management product offerings.

$$d_{1it} = \frac{ln(\frac{A_{it}}{DB_{it}}) + (r_t + \frac{\sigma^2_{Ait}}{2})T}{\sigma_{Ait}\sqrt{T}} \quad \text{and}$$

$$d_{2it} = d_{1it} - \sigma_{Ait}\sqrt{T} = \frac{ln(\frac{A_{it}}{DB_{it}}) + (r_t - \frac{\sigma^2_{Ait}}{2})T}{\sigma_{Ait}\sqrt{T}},$$

in which σ_{Ait} denotes the volatility of assets.

The time to maturity T is assumed to be one-year and is in line with i) the usual assets' classification into short-term and long-term liabilities that are required by the model, ii) the expected time for banks to adapt to capital increases, and iii) the stress test exercises conducted by the regulatory authority (BCBS, 2010, 2011; O'Keefe and Ufier, 2017). Following this approach, we were able to produce robust estimates of the probability of default for each financial institution in the study.

To calculate each PD_{it}, two important assumptions are made. The first is that the bank's asset values are log-normally distributed (Crouhy et al., 2000; Lehar, 2005; Guerra et al., 2016; Souza et al., 2016; da Rosa München, 2022). The second is that investors are risk-neutral, that is, the demanded rate of return is the risk-free rate of return r_t, which is lower than that required by risk-averse investors. This assumption results in conservative (higher) PD_{it} estimates. Thus, the idiosyncratic PD_{it} of a FI in a time horizon T, computed in $t = 0$, is given by 1.3.

$$\begin{aligned} PD_{it} &= P(DB_{it} \geq A_{it}) \\ &= P(\ln DB_{it} \geq \ln A_{it}) \\ &= \mathcal{N}(-d_{2it}) \\ &= \mathcal{N}\left[-\frac{ln(\frac{A_{it}}{DB_{it}}) + (r_t - \frac{\sigma^2_{Ait}}{2})T}{\sigma_{Ait}\sqrt{T}} \right] \end{aligned} \tag{1.3}$$

Note that the PD_{it} is the area under the default barrier, i.e., a fraction of total liabilities. Also, note that the negative of d_{2it} can also be used to compute the distance to distress for a risk-neutral environment, which is the distance between the bank's asset value and the distress barrier in $t = 0$, measured in assets value' standard deviations.

1.2.2 Z-Score Model

In addition to employing the Merton (1974) model to estimate the probability of default for each financial institution, we utilized the Z-Score as a robustness check. This score, an adaptation of the indicator created by Altman (1968), is widely recognized as a measure of bank stability and risk within the banking literature (Beck et al., 2013; Chiaramonte et al., 2015; Rosa and Gartner, 2018; World Bank, 2020; Nguyen, 2021; Stewart and Chowdhury, 2021). It explicitly compares buffers (capitalization and returns) with risk (volatility of returns) to measure a bank's solvency risk. The specification, proposed by Boyd and Runkle (1993), is given by 1.4.

$$Z\text{-}Score_{it} = \frac{EA_{it} + \mu(ROA_{it})}{\sigma(ROA_{it})} \tag{1.4}$$

Where EA_{it} represents the bank's equity-to-assets ratio and ROA_{it} denotes the historical return on assets. In scenarios with uncertainty, one could consider EA_{it} as deterministic and model ROA as a stochastic variable with finite mean $\mu(ROA_{it})$ and variance $\sigma(ROA_{it})^{2}$[6]. The Z-Score quantifies the number of standard deviations by which returns must fall from the mean in order to net the bank equity. A higher Z-Score indicates increased bank stability and a lower likelihood of insolvency risk.

For a specific case where ROA follows a normal distribution with mean $\mu(ROA_{it})$ and standard deviation $\sigma(ROA)$, the probability of insolvency $P(ROA \leq -EA)$, can be determined as shows in Equation 1.5.

$$P(ROA \leq -EA) = \mathscr{N}\left(\frac{-EA - \mu(ROA)}{\sigma(ROA)}\right) = 1 - \mathscr{N}\left(\frac{EA + \mu(ROA)}{\sigma(ROA)}\right) \tag{1.5}$$

In which $\mathscr{N}(.)$ is the cumulative normal distribution function, and the second equality holds because the normal distribution is symmetric (Mare et al., 2017).

While the Z-Score offers valuable insights, it presents certain limitations as a financial stability measure. For instance, it may not detect short-term fluctuations in bank risk, given that its variance is calculated using extended historical data. Moreover, since Z-

[6]Some works utilize ROA_{it} point-in-time and a three-year rolling standard deviation to increase the sensitivity and fluctuation of the Z-Score (Beck et al., 2013; Lepetit and Strobel, 2013; Chiaramonte et al., 2015; Mare et al., 2017).

Scores rely solely on accounting data, they are constrained by the quality of the underlying accounting and auditing framework. If financial institutions manage to smooth out the reported data, the Z-Score might offer an overly favorable assessment of their stability. Nonetheless, an advantage of the Z-Score is its applicability to institutions lacking more sophisticated market-based data (Lepetit and Strobel, 2013; Mare et al., 2017).

1.2.3 Empirical Strategy

The empirical strategy to estimate the impact of bank capital regulation on FI's PD employs the two-way fixed effects model. This framework allows controlling the intrinsic characteristics of each FI and the influence of macroeconomic factors on the probability of default over time, assuming linear additive effects for this case (Petersen, 2008; de Chaisemartin and D'Haultfœuille, 2020; Callaway and Sant'Anna, 2021; Baltagi, 2021; Imai and Kim, 2021; Davydov et al., 2021; Baker et al., 2022; Mateev et al., 2022; Karolyi et al., 2023). For each estimated model, both PD and Z-Score are used as dependent variables[7].

Consider a balanced panel data set of N financial institutions and T time periods. Let $\boldsymbol{X}_{it}$ and Y_{it} represent the vector of covariates and the observed outcome variable for an FI i at time t, respectively. Then, the two-way fixed effects model is given by 1.6.

$$Y_{it} = \boldsymbol{\beta}\boldsymbol{X}_{it} + \alpha_i + \gamma_t + \varepsilon_{it} \tag{1.6}$$

Where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_N)$ is a vector of unknown model coefficients, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)$ is the unobserved individual specific effects, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_T)$ is the unobserved time specific effects and ε_{it} is the error term, for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Conditioning the unobserved effects serves to control for endogeneity as the individual and time effects capture unobserved heterogeneity that can be related to the covariates. In the statistical analysis, robust standard errors are used, and they are double-clustered to provide more accurate standard error estimates[8].

[7] The purpose of analyzing the impact on Z-Score is to provide a robustness check for the use of Merton (1974)'s structural model for the probability of default. Accordingly, while we will present results for both the Z-Score and PD, our discussion will focus primarily on PD.

[8] The double-clustering approach is particularly beneficial in panel data settings where variables of interest might be correlated both across time (autocorrelation) and cross-sectional units (heteroskedasticity). Such correlations can violate the Ordinary Least Squares (OLS) assumptions of independence and homoscedasticity. Taking into account these dual correlations, the method aims to mitigate possible violations and improve the accuracy of standard error estimates, ensuring robustness against both

The first econometric specification, given by 1.7, considers the impact of bank regulation on FI's PD, while also controlling for lagged capital adequacy ratio to capture temporal dynamics inherent in banking decisions and reduce potential issues related to reverse causality (Berger and Bouwman, 2013).

$$PD_{it} = \beta_1 CAR_{it} + \beta_2 CAR_{it-1} + \alpha_i + \gamma_t + \varepsilon_{it} \tag{1.7}$$

Where $PD_{it} \in [0, 1]$ is defined in 1.3 and CAR_{it} is the capital adequacy ratio (tier I and II) of each FI_{it}.

Because there is significant literature that considers the impact of profitability and performance on FI's PD in addition to the capital requirement (Jayadev, 2013; Giordana and Schumacher, 2017), model 1.8 specifies the equation that discusses this case as a robustness check.

$$PD_{it} = \beta_1 Performance_{it} + \beta_2 CAR_{it} + \beta_3 CAR_{it-1} + \alpha_i + \gamma_t + \varepsilon_{it} \tag{1.8}$$

In which $Performance_{it}$ is captured by $Spread_{it}$, ROE_{it} (return on equity) and ROA_{it} (return on assets). ROE_{it} and ROA_{it} are annualized[9] and calculated by the net income divided by equity and net income divided by total assets, respectively, of each FI_{it}. The $Spread_{it}$ is defined as the difference between interest income and interest expense taking into account loan loss provisions[10], which is given by 1.9.

$$Spread_{it} = \frac{Net\ Interest\ Income_{it} - Net\ Loan\ Loss\ Provisions_{it}}{Gross\ Interest\ Income_{it}} \tag{1.9}$$

In order to assess the impact of bank regulation on each FI's PD and considering the importance of Basel III as a comprehensive set or reform measures for this purpose (BCBS,

cross-sectional and time-series dependence (Sun et al., 2018; Abadie et al., 2023).

[9]The annualization is done from the sum of net income over 1 year divided by the average of equity (ROE) and assets (ROA) in the period.

[10]It is well known that there is a vast literature that discusses the components of spread for banks. Although some works for the Brazilian banking sector use the difference between the rate of return obtained in credit operations and the cost funding in the construction of the ex-post spread (Dantas et al., 2012; Fiche et al., 2017; Magalhães-Timotioa et al., 2018), our approach considers all sources of interest income in addition to credit operations (such as lease, security, financial derivative, foreign exchange and mandatory reserve income). This definition is equivalent to the fraction of net interest income over gross interest income. However, because net interest income takes into account net loan loss provisions on the balance sheet of financial institutions in Brazil, the appropriate definition of spread, which is used to account for the gain from the collection of interest represented as a margin, should disregard these accounting data

2011), we used a difference-in-differences approach with the two-way fixed effect model (Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021; Baker et al., 2022) in equation 1.10 to estimate the impact of Basel III Accord in each Brazilian FI's PD classified in Segment 1.

$$PD_{it} = \delta^{DD} D_{it} + \beta CAR_{it-1} + \alpha_i + \gamma_t + \varepsilon_{it} \tag{1.10}$$

Where D_{it} denotes the interaction term between the indicator variable for the treated unit, D_i (bank Segment 1, in this case), and the indicator variable for observations in periods t, $Post_t$ (Basel III since December 2013, in this case). CAR_{it-1} is the lagged capital adequacy ratio used to account for the pre-existing levels and trends in CAR prior to the implementation of Basel III. Note that the presence of both α_i and γ_t subsumes the main effects of D_i and $Post_t$.

To examine the potential conflict that regulators often face, where the pursuit of a safer banking system could lead to higher banking concentration and an increase in capital costs (Alexandre et al., 2022), we test the specification given by equation 1.11. This model assesses the impact of banking concentration on the FI's PD, taking into account variations in capital requirement.

$$PD_{it} = \beta_1 Concentration_{it} + \beta_2 CAR_{it} + \beta_3 CAR_{it-1} + \alpha_i + \gamma_t + \varepsilon_{it} \tag{1.11}$$

In which $Concentration_{it}$ is measured by asset concentration as in Rosa and Gartner (2018), which is calculated through the proportion of $bank_{it}$ assets in relation to the sum of assets of the banking system in t, and CAR_{it} is intended to capture the influence of regulatory capital on PD_{it}.

Lastly, recognizing that an increase in CAR can lead to strategic adaptations by banks (Shim, 2013; Uluc and Wieladek, 2018; Gropp et al., 2018; Pariès et al., 2022; Alexandre et al., 2022), our last model focuses on the mechanisms through which CAR affects loan supply and demand. Therefore, to investigate the relationship between an increase in the CAR and its effects on the loan market, particularly during periods of economic uncertainty, we employed the specification given by equation 1.12.

$$Loan_{it} = \beta_1 PD_{it} + \beta_2 Concentration_{it} + \beta_3 CAR_{it} + \beta_4 CAR_{it-1} + \alpha_i + \gamma_t + \varepsilon_{it} \tag{1.12}$$

Where $Loan_{it}$ represents loan operations by risk level for bank i at time t. The inclusion of both current and lagged CAR allows us to capture both the immediate and delayed effects of changes in capital adequacy on lending behavior.

1.3 Data

We utilized quarterly data from December 2000 to September 2022 for 244 Brazilian financial institutions, yielding an unbalanced panel data with 9,653 observations[11]. All balance sheet data employed in this study are publicly provided by the Central Bank of Brazil (BCB, 2023a).

The data set considers financial conglomerates and independent institutions until December 2014, and the prudential conglomerates and independent institutions before March 2015[12] with the business model category of $b1$, $b2$, $b4$, and $n1$[13], provided there are at least six valid observations in the studied period. The final data set represents 99.82% of total assets and 99.75% of total credit of covered member institutions in September 2022, with an average of 98.94% and 99.07% throughout the period, respectively. For the interest rate, we used public data provided by B3, the Brazilian financial market infrastructure company (B3, 2023).

To estimate the probability of default on a one-year horizon for each FI using the Merton (1974)'s structural model, we applied the following variables: adjusted total

[11]Although information is available from 2000:I-2000:III in the database, we used these first three quarters to calculate assets volatility since bank capital information is available from December 2000.

[12]Note that until December 2013, the Central Bank of Brazil registered only the institution type of financial conglomerates and independent institution. Starting before March 2014, the perspective of prudential conglomerate and independent institution was included. However, capital information from bank's DLO (Statement of Operating Limits) was published only in the prudential conglomerate and independent institution perspective before March 2015. The difference between the two filters lies in the latter's inclusion of institutions other than those belonging to the financial conglomerate, such as: (i) consortium administrators, (ii) payment institutions, (iii) companies that perform acquisition of credit operations, including real estate or credit rights, (iv) other legal entities domiciled in the country that have as an exclusive objective an equity interest in the aforementioned entities and (v) investment funds in which the entities that compose a prudential conglomerate take or retain substantial risks and benefits (BCB, 2023a).

[13]We used only these four business model category to account for institutions that issue covered deposits under the deposit insurance system in Brazil. The categories include: ($b1$) for commercial banks, universal banks with commercial portfolios, or savings banks; ($b2$) for universal banks without commercial portfolios, investment banks, or foreign exchange banks; and ($n1$) for non-banking credit companies. The member institutions are: (i) multiple banks; (ii) commercial banks; (iii) investment banks; (iv) development banks; (v) Caixa Econômica Federal (Brazilian federal savings bank); (vi) savings banks; (vii) finance and investment companies; (viii) building societies; (ix) mortgage companies savings; and (x) loan associations (FGC, 2023; BCB, 2023a). For more information on FGC-covered deposits, see BCB (2021a).

assets[14] for A, total liabilities to calculate DB, annualized interbank interest rate DI for r, and the annualized standard deviation of the logarithmic returns of adjusted total assets, that is, $\log(A_t/A_{t-1})$, for asset volatility σ_A. Table 1.1 presents the aggregate descriptive statistics for these variables, and all balance sheet accounts are shown in Appendix A.

Table 1.1: Descriptive statistics.

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
ATA^a	44.78	192.25	0.00	0.34	10.29	2,184.86
TL^a	40.77	177.47	0.00	0.26	8.95	2,018.16
Loans^a	18.54	84.21	0.00	0.10	3.93	978.56
DI^b	10.91	5.18	1.90	6.40	14.13	26.23
AV	0.42	0.41	0.00	0.14	0.58	6.03
PD^b	16.20	20.28	0.00	0.02	29.98	99.84
Z-Score	12.22	15.97	−1.04	4.65	15.13	542.71
CAR^b	31.01	40.77	0.38	14.83	30.81	1,151.84
AC^b	0.91	3.12	0.00	0.01	0.24	23.91
EA^b	22.67	21.21	−2.74	10.05	25.03	100.00
ROA^b	2.38	7.75	−113.93	0.69	4.05	125.17
ROE^b	5.29	714.41	−69,934.21	4.31	27.23	277.07
Spread^b	48.33	1,815.16	−22,848.87	9.79	51.32	170,490.10

Notes: The sample period runs from 2000:IV-2022:III for the Brazilian financial system. *ATA* = adjusted total assets; *TL* = total liabilities; *Loans* = loan operations by risk level; *DI* = interest rate; *AV* = assets volatility; *PD* = probability of default; *Z-Score* = $(EA + \mu(ROA))/\sigma(ROA)$; *CAR* = capital adequacy ratio (tier I and II); *AC* = assets concentration; *EA* = equity over total assets; *ROA* = return on assets; *ROE* = return on equity and *Spread* = (net interest income - net loan loss provisions)/gross interest income.

^{a} In BRL billion.

^{b} In percentage.

[14]The adjusted total assets represent a modification of the total assets, accounting for specific adjustments related to netting and reclassification. Netting involves consolidating certain balance sheet items, such as repurchase agreements, interbank relations and relations within branches, the foreign exchange portfolio, and debtors due to litigation. In addition, reclassifications are performed within the foreign exchange and leasing portfolios, which may involve reorganizing or reevaluating these assets according to specific criteria or regulations.

1.4 Results and Discussion

This section outlines the findings concerning the impact of bank capital regulation on the probability of default and Z-Score of financial institutions within the Brazilian financial system from December 2000 to September 2022. As mentioned and detailed in Section 1.2.3, a two-way fixed effects model was used to control the intrinsic characteristics of each FI, along with the influence of macroeconomic factors on the probability of default over time.

The results of the proposed model 1.7, as shown in Table 1.2, are consistent with the regulatory perspective and resonate with the extensive literature from different economies (Valahzaghard and Bahrami, 2013; Thakor, 2014; Vazquez and Federico, 2015; Giordana and Schumacher, 2017; Rosa and Gartner, 2018; Parrado-Martínez et al., 2019; Fidrmuc and Lind, 2020; Le et al., 2020; BCBS, 2021; Jones et al., 2022). Specifically, they show that the higher the CAR of an FI, the lower the associated PD. This relationship holds due to the greater capacity of the institution to absorb losses, a finding that remains robust even after controlling for past CAR[15].

In particular, Table 1.2 shows that, on average, an increase of 1% in CAR reduces PD by 8.4% without considering the influence of lagged CAR. However, when historical capital measures are taken into account, the dynamics become more elaborate. A 1% increase in the current CAR diminishes PD by 3.7%, and the lagged CAR from the previous period contributes further to a reduction of 4.05%. This pattern accentuates the prolonged impact of CAR decisions, underscoring that robust financial strategies in one period can have lasting positive implications for subsequent periods.

[15]Incorporating lagged CAR into the model captures the time-dependent nature of banking decisions and the potential lag in responses due to regulatory compliance. Moreover, this inclusion helps address potential issues of reverse causality (Berger and Bouwman, 2013). To ensure robustness, we controlled for higher levels of lagged CAR up to one year. The results, presented in Table A.3 in Appendix A, support the decision to use only one lagged CAR.

Table 1.2: Effects of capital adequacy ratio on FI's PD and Z-Score.

	Probability of Default		Z-Score	
	(1)	(2)	(3)	(4)
Capital Adequacy Ratio	-8.3943^{**}	-3.7089^{*}	0.4895^{***}	0.3959^{***}
	(3.5424)	(2.1385)	(0.0489)	(0.0435)
Capital Adequacy Ratio$_{t-1}$		-4.0481^{*}		0.1163^{***}
		(2.2651)		(0.0436)
Observations	9,475	9,216	9,583	9,295
R^2	0.0064	0.0057	0.1491	0.1521
Adjusted R^2	-0.0291	-0.0309	0.1188	0.1209
F Statistic	58.8755^{***}	25.2641^{***}	$1{,}620.8760^{***}$	803.8809^{***}

Notes: This table presents the two-way fixed effects estimates of the FI's capital adequacy ratio on their PD and Z-Score. Both dependent and independent variables are in the natural log. Robust standard errors double-clustered are in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

It can be argued that a higher CAR does not always lead to a reduction in an FI's PD, particularly when an excess of capital results in inefficient allocation of resources. Consequently, an excess of misallocated capital can consume an FI's profit margin to the point of increasing its probability of default (Jayadev, 2013; Giordana and Schumacher, 2017). This scenario may arise when financial institutions engage in overly conservative investment strategies or fail to optimize their capital structures, leading to decreased profitability and increased vulnerability to default. To address this case, equation 1.8 was estimated as a robustness check to control for the impact of CAR on PD, considering performance metrics such as return on assets, return on equity, and spread of each financial institution. The results, displayed in Table 1.3, show that PD still decreases (and Z-Score increases) with an increase in CAR, even when controlling for ROA in Regressions (1) and (4), ROE in Regressions (2) and (5), and spread in Regressions (3) and (6)[16]. This finding aligns with the conclusions reached by Valahzaghard and Bahrami (2013), Rosa and Gartner (2018), Parrado-Martínez et al. (2019), and Davydov et al. (2021).

[16]The robustness of these results, including controls for lagged CAR, is confirmed in Table A.4 in Appendix A.

Table 1.3: Effects of CAR and performance on FI's PD and Z-Score.

	Probability of Default			Z-Score		
	(1)	(2)	(3)	(4)	(5)	(6)
ROA	-1.11^{***}			0.07^{***}		
	(0.35)			(0.02)		
ROE		-0.00^{***}			0.00^{***}	
		(0.00)			(0.00)	
Spread			-0.86^{**}			0.06^{***}
			(0.34)			(0.02)
CAR	-3.64^{***}	-4.02^{***}	-3.93^{***}	0.43^{***}	0.49^{***}	0.46^{***}
	(1.11)	(0.84)	(0.94)	(0.06)	(0.05)	(0.06)
Observations	7,946	9,653	7,972	7,937	9,583	7,930
R^2	0.05	0.04	0.05	0.15	0.15	0.16
Adjusted R^2	0.01	0.00	0.01	0.12	0.12	0.12
F Statistic	182.29^{***}	184.31^{***}	189.86^{***}	681.41^{***}	826.24^{***}	702.81^{***}

Notes: This table presents the two-way fixed effects estimates of the FI's capital adequacy ratio and performance on their PD and Z-Score. *CAR* = capital adequacy ratio (tier I and II); *ROA* = return on assets and *ROE* = return on equity. All variables are in the natural log except *PD* and *ROE*. Robust standard errors double-clustered are in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

In addition to the impact that the CAR has on FI's PD and Z-Score, it is also important to explore the effects of the Basel III agreement in Brazil, implemented since October 2013. Designed to strengthen the resilience and risk management practices of financial institutions, Basel III's implementation within the Brazilian banking system is expected to significantly influence bank PD. We have also incorporated a control for lagged CAR in our analysis to account for pre-existing levels and trends in CAR before the implementation of Basel III. Utilizing a difference-in-differences approach (as outlined in equation 1.10), our findings reveal that the Basel III agreement led to a 3.4% reduction in the probability of default for Brazilian banks in Segment 1, as shown in Regression (2) of Table 1.4.

Table 1.4: Effects of Basel III Accord on FI's PD and Z-Score.

	Probability of Default		Z-Score	
	(1)	(2)	(3)	(4)
Basel III:S1	-4.5962^{**}	-3.3930^{**}	0.3097^{**}	0.2210^{*}
	(1.9259)	(1.7308)	(0.1264)	(0.1277)
CAR_{t-1}		-7.0039^{**}		0.4328^{***}
		(3.2094)		(0.0519)
Observations	9,475	9,216	9,583	9,295
R^2	0.0002	0.0053	0.0053	0.1219
Adjusted R^2	-0.0356	-0.0313	-0.0301	0.0896
F Statistic	1.4956	23.4912^{***}	48.8479^{***}	622.0274^{***}

Notes: This table presents the difference-in-difference estimates in a two-way fixed effects model, assuming a value of 1 for FIs classified in Segment 1 after the Basel III Accord. Note that the post-treatment indicator are subsumed by the fixed effects. Basel III was implemented in Brazil in October 2013, guided by four resolutions (No. 4.192 to 4.195) and fifteen circulars (No. 3.634 to 3.648), released on March 1, 2013. These documents delineated the implementation rules for the banking system, detailing a convergence calendar that specified a requirement of 11% of the RWA from October 2013 to December 2015; 9.875% in 2016; 9.25% in 2017; 8.625% in 2018; and 8% from 2019 onward, in addition to particular capital buffers (BCB, 2022b). The dependent variables are in the natural log. Robust standard errors double-clustered are in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

These results in Table 1.4 align with expectations from a regulatory perspective, as Basel III increased the capital requirements for banks, ensuring greater resilience to withstand losses during financial stress. This was achieved by focusing on going-concern loss-absorbing capital, notably in the form of Common Equity Tier 1 (CET1) capital, within Segment 1, enhancing the overall quality of bank regulatory capital. Furthermore, Basel III introduced several macroprudential measures into the regulatory framework. These include (i) capital buffers that are built during favorable economic conditions and can be utilized during stressful periods to reduce procyclicality; (ii) a large exposure regime to mitigate systemic risks associated with interconnectedness among financial institutions and concentrated exposures; and (iii) a specialized capital buffer to address externalities created by systemically important banks, often referred to as Segment 1 in the Brazilian financial market (BCBS, 2017).

It is worth noting that the implementation of the Basel III Accord in Brazil was more rigorous than in other international contexts. While the agreement set a minimum capital requirement of 8%, the BCB increased this requirement to 11% to operate in the market, along with additional specific capital buffers[17]. In addition to the capital requirements, it is important to examine the banking concentration within Brazil. Since 2015, the Concentration Ratio of the Five Largest (RC5) Brazilian banks has exceeded 65% across various banking segments[18] and accounting aggregates, such as assets, deposits, and credit operations. The overall concentration reached an average of 79.2% in 2018, followed by 78.6% in 2019, 75.5% in 2020, 74.2% in 2021, and 73% in 2022 (BCB, 2022a). This gradual decrease in bank concentration in recent years can be attributed to specific market dynamics. Notably, the loss of market share by public banks (from 47.6% in 2019 to 43.7% in 2022), particularly BNDES, and the concurrent advancement of cooperatives (from 4.3% in 2019 to 6.4% in 2022), have contributed to this trend.

These aspects of the Brazilian economy, in conjunction with a robust and well-capitalized banking market, limit the occurrence of banking failure (Liberman et al., 2018). In terms of magnitude, from 2000 to September 2022, FGC, the Brazilian private DIA, recorded only 20 extrajudicial settlements or interventions made by the BCB on the banking market (8.2%). In contrast, the Federal Deposit Insurance Corporation (FDIC) in the USA registered 563 interventions (11.9%), reflecting a different regulatory environment[19].

Thus, the high concentration in Brazil's banking sector raises important questions about its role as a control variable in assessing the impact of CAR on PD, considering that a higher concentration can be indicative of a more consolidated and stable financial system. In this sense, Table 1.5 shows, as expected from a regulatory point of view and consistent with the literature for different economies (Berger and Bouwman, 2013; Rosa and Gartner, 2018; Parrado-Martínez et al., 2019; Karolyi et al., 2023), that an increase of 1 percentage point (pp) in the concentration of banking assets reduces the PD, on average, by 1.4%[20].

[17]For more information, see Appendix A.1.

[18]The BCB (2022a) considers three levels of aggregation to calculate the concentration of RC5: banking and non-banking segment ($b1 + b2 + b3 + b4 + n1$), banking segment ($b1 + b2$), and commercial banking segment ($b1$). The business model category ($b1$) includes commercial banks, multiples with a commercial portfolio and savings banks; ($b2$), multiple banks without a commercial portfolio and investment banks; ($b3$), credit unions; ($b4$), development banks; and ($n1$), non-bank credit institutions.

[19]For more information, see FGC for the list of banking failures in Brazil and FDIC for the list of banking failures in the USA. The percentages were calculated considering the member institutions of the deposit insurance in each country in September 2022.

[20]It is worth mentioning that, as a robustness check, while the Basel III Accord reduced by 3.4% the probability of default of Segment 1 in Brazil, it also increased in 0.4% the banking assets concentration in this same sample.

Table 1.5: Effects of banking assets concentration and CAR on FI's PD.

	Probability of Default	
	(1)	(2)
Banking Assets Concentration	-1.4141^{*}	-0.7292
	(0.7637)	(0.4794)
Capital Adequacy Ratio	-8.4893^{**}	-3.7323^{*}
	(3.5434)	(2.1410)
Capital Adequacy Ratio$_{t-1}$		-4.0815^{*}
		(2.2686)
Observations	9,475	9,216
R^2	0.0067	0.0058
Adjusted R^2	-0.0289	-0.0309
F Statistic	31.0695^{***}	17.1804^{***}

Notes: This table presents the two-way fixed effects estimates of the FI's assets concentration and CAR on their PD. Assets concentration represents the ratio of the assets of a specific bank to the total assets in the banking system at time t. Although the Herfindahl-Hirschman index could be used by squaring assets concentration (Dantas et al., 2012; Fiche et al., 2017; Matt and Andrade, 2019), we opted for the unaltered assets concentration metric as in Rosa and Gartner (2018), to enable a more straightforward interpretation of the coefficient. Regression (2) includes the lag of CAR to control its impact on assets banking concentration. All dependent and independent variables are in the natural log, except Assets Concentration. Robust standard errors double-clustered are in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

In addition to bolstering financial stability and influencing banking concentration, an elevated CAR can have significant implications for the loan market. These effects are especially pronounced during economic downturns marked by heightened default rates and restricted liquidity. When confronted with these challenges, banks might strategically adjust by either reducing assets, leading to a contraction in loan supply, or by raising loan interest rates, subsequently suppressing loan demand (Shim, 2013; Noss and Toffano, 2016; Uluc and Wieladek, 2018; Gropp et al., 2018; Acharya et al., 2018; De Nicolò, 2019; Fraisse et al., 2020; Malovaná and Ehrenbergerová, 2022; Pariès et al., 2022; Alexandre et al., 2022; Bonaccorsi di Patti et al., 2023).

However, it is important to note that the impact of CAR on loan supply is not a consensus in the banking literature. In contrast to the earlier view, a higher CAR could theoretically facilitate an increase in loan supply, especially if this enhanced capital buffer bolsters investor and depositor confidence, encouraging banks to extend more loans. Additionally, when capital buffers make banks safer and more resilient, markets might perceive them as less risky, leading to a reduction in their marginal cost of funding. Consequently, they may be better positioned to assume greater risks, either by relaxing lending standards or increasing the volume of loans in their portfolio (Thakor, 2014; Bassett and Berrospide, 2018; Begenau, 2020; Bahaj and Malherbe, 2020).

Our results, as shown in Table 1.6, align with the prevailing literature, suggesting that heightened capital requirements result in a decreased supply of loans. In particular, a 1% increase in the current CAR is associated with a 3.2% decline in loan operations.

Table 1.6: Effects of banking assets concentration, CAR and PD on FI's loan operations.

	Loan Operations	
	(1)	(2)
Probability of Default	-0.0096^{**}	-0.0107^{**}
	(0.0041)	(0.0046)
Banking Assets Concentration	0.9373^{***}	0.9151^{***}
	(0.0674)	(0.0693)
Capital Adequacy Ratio	-0.4135^{***}	-0.3194^{***}
	(0.0994)	(0.0904)
Capital Adequacy Ratio$_{t-1}$		-0.1155^{**}
		(0.0555)
Observations	8,860	8,563
R^2	0.4345	0.4284
Adjusted R^2	0.4133	0.4062
F Statistic	$2{,}187.0470^{***}$	$1{,}544.2940^{***}$

Notes: This table presents the two-way fixed effects estimates of the FI's probability of default, banking assets concentration, and capital adequacy ratio on their loan operations by risk level. All dependent and independent variables are in the natural log, except probability of default. Robust standard errors double-clustered are in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

The findings from this research underscore the need for a comprehensive examination of the complex trade-offs faced by regulators. Striking the right balance involves careful consideration of factors such as banking concentration, the cost of capital, and comprehensive regulatory requirements designed to prevent systemic crises. This raises a natural question about possible regulatory easing that could foster banking competition, enhance the efficiency of financial intermediation, and promote overall social welfare, while maintaining a stable financial system. Addressing such a multifaceted question requires a comprehensive analysis, incorporating both theoretical insights and empirical evidence, to discern the economic consequences of variations in capital requirements. This complex issue will be explored in more depth and discussed in Chapter 2.

1.5 Final Remarks

This paper examines bank risk using the structural model of Merton (1974) and Z-Score, employing public balance sheet data to evaluate the impact of bank capital regulation on banks' probability of default. To achieve this, we utilized the two-way fixed effects model and a difference-in-difference approach, analyzing data from the Brazilian banking market from December 2000 to September 2022. The results confirm the practical importance of regulatory measures such as Basel III and highlight the need for balanced capital requirements to enhance financial stability.

Our findings indicate that an increase in the capital adequacy ratio results in a decreased probability of default for banks, a relationship that holds even when accounting for performance metrics such as ROA, ROE, Spread, and lagged CAR. Furthermore, our analysis reveals that the implementation of the Basel III agreement has also played a role in reducing the PD of banks. In the context of Brazil, our findings highlight the trade-offs faced by regulators, where a higher CAR is associated not only with a lower PD, but also with an increased banking concentration and a reduction in loan operations. Consequently, striking the right balance requires careful consideration of factors such as banking concentration, cost of capital, and comprehensive regulatory requirements designed to mitigate systemic crises.

The implications of our results should be informative for regulators concerned with the impact of capital regulation on banks' PD. In particular, this study raises questions about how regulatory frameworks can be optimized to enhance banking competition and social welfare, while maintaining a stable financial system. However, addressing such a multifaceted question requires a comprehensive analysis that combines both theoretical insights and empirical evidence. We explore these aspects in more detail in Chapter 2.

Chapter 2

Systemic Risk Measures and Optimized Capital Requirement

2.1 Introduction

The recent Global Financial Crisis has emphasized the importance of policies that improve the overall stability of the financial system. This crisis served as one of the clearest illustrations in history of systemic risk, in which banks and credit play a particularly important role (De Bandt and Hartmann, 2019). Although the topic has generated extensive literature and intense financial interest, many macroeconomists and policymakers recognize a significant gap in understanding the channels of system-wide risk and the contribution of systemic risk by individual financial institutions to the broader economy (Christiano et al., 2018; Fidrmuc and Lind, 2020). Furthermore, the critical nature of this issue has generated a growing consensus among policymakers to adopt a macroprudential approach to regulation and supervision, as it is considered essential to ensure a resilient global financial system (Hannoun, 2010; Borio, 2011; Galati and Moessner, 2013).

Understanding how financial institutions affect systemic risk, whether through their idiosyncratic characteristics or its connections with the rest of the economy, is fundamental for effective action by central banks and policymakers in time of crisis. In that sense, Basel III was introduced as an important part of macroprudential regulation in response to the GFC. Introduced in 2010 by the Basel Committee on Banking Supervision (BCBS) at the Bank for International Settlements (BIS), Basel III comprises a robust set of reform measures aimed to enhance banking regulation, supervision, and risk management. Its goal is to fortify the banking sector and foster financial stability by elevating bank liquidity and diminishing bank leverage. One of the core metrics created for this purpose is the requirement of a minimum capital adequacy ratio (CAR)[1] that banks must maintain to operate in the market, which is an important tool to absorb unexpected losses without requiring the bank to cease its operations (BCBS, 2011).

In addition to Basel III's capital requirements (CR) and capital buffers[2], other regulatory mechanisms have been designed to mitigate systemic risk. In particular, the Financial Stability Board (FSB) introduced the bail-in framework in 2018 as a resolution strategy for handling distressed banks without resorting to taxpayer-funded bailouts. Bail-in prescribes *ex-ante* resolution planning without the use of public funds for solvency support, obliging shareholders and creditors to share the burden of losses (FSB, 2018, 2021). Together, these mechanisms underscore the ongoing commitment of policymakers and central banks to address the critical issue of financial stability.

Although there exists a large literature exploring Basel III's impact on banks' balance sheets and its transmission to the real economy, especially since the GFC, the topic continues to be a subject of active discussion and research. However, a broad consensus supported by substantial evidence recognizes the benefits of Basel III in enhancing financial stability. This has been achieved principally through the reduction in banks' probability of default, as corroborated by numerous studies for different economies (Berger and Bouwman, 2013; Vazquez and Federico, 2015; Giordana and Schumacher, 2017; Rosa and Gartner, 2018; Parrado-Martínez et al., 2019; Le et al., 2020; BCBS, 2021; Jones et al., 2022).

With respect to unresolved questions concerning the effects on capital cost, banking lending, and banking concentration, many works have shown (Maredza, 2016; Gropp et al., 2018; De Nicolò, 2019; Fraisse et al., 2020; Alexandre et al., 2022; Tran et al., 2022; BCBS, 2022; Bonaccorsi di Patti et al., 2023), as well as in Chapter 1, that financial regulators often face conflicting objectives. One of the trade-offs faced by them in establishing the

[1]The capital adequacy ratio is calculated by summing Tier 1 and Tier 2 capital and dividing them by the risk-weighted assets (RWA). Tier 1 capital is the core capital of a bank and is formed by the Common Equity Tier 1 (CET1) and Additional Tier 1 (AT1), which includes equity capital and disclosed reserves. This type of capital, also described as going-concern capital, can absorb losses without requiring the bank to cease its operations. On the other hand, Tier 2 capital, also described as gone-concern capital, has a lower standard than Tier 1 and is used to absorb losses in the event of a liquidation.

[2]Basel III introduces two capital buffers that financial institutions must hold beyond other minimum capital requirements: the capital conservation buffer, providing an extra layer of usable capital to absorb losses, and the countercyclical capital buffer (CCyB), aimed at protecting the banking sector from periods of excess credit growth associated with system-wide risks.

minimum CR, with the goal of both strengthening the financial system and preventing systemic crises, is the potential pitfall that a more secure system might lead to greater banking concentration and higher capital costs. This could undermine the efficiency of financial intermediation, potentially resulting in a social cost that outweighs its benefit. The challenge arises because banks, in response to higher capital requirements, may either reduce assets, thus diminishing loan supply, or elevate the loan interest rate, resulting in a decrease in loan demand[3]. Thus, the design of the financial system's protection mechanism must be finely calibrated, with the goal of maximizing social welfare without unintended consequences.

Regarding the Brazilian context, the implementation of the Basel III Accord was more rigorous than in other international contexts. While the agreement set a minimum CR of 8%, the Central Bank of Brazil increased this requirement to 11% to operate in the market, along with additional specific capital buffers[4]. In addition to the capital requirements, it is important to examine the banking concentration within Brazil. As mentioned in Chapter 1, since 2015, the Concentration Ratio of the Five Largest Brazilian banks has exceeded 65% across various banking segments[5] and accounting aggregates, such as assets, deposits, and credit operations. The overall concentration reached an average of 79.2% in 2018, followed by 78.6% in 2019, 75.5% in 2020, 74.2% in 2021, and 73% in 2022 (BCB, 2022a).

These aspects of the Brazilian economy, in conjunction with a robust and well-capitalized banking market, limit the occurrence of banking failure (Liberman et al., 2018). In terms of magnitude, from 2000 to September 2022, the Brazilian private deposit insurance agency, *Fundo Garantidor de Créditos*, recorded only 20 extrajudicial settlements or interventions made by the BCB on the banking market (8.2%). In contrast, the Federal Deposit Insurance Corporation (FDIC) in the USA registered 563 interventions (11.9%), reflecting

[3]It is important to note, however, that the impact of CAR on loan supply is not a consensus in the banking literature. In contrast to the earlier view, a higher CAR could theoretically facilitate an increase in loan supply, especially if this enhanced capital buffer bolsters investor and depositor confidence, encouraging banks to extend more loans. Additionally, when capital buffers make banks safer and more resilient, markets might perceive them as less risky, leading to a reduction in their marginal cost of funding. Consequently, they may be better positioned to assume greater risks, either by relaxing lending standards or increasing the volume of loans in their portfolio (Thakor, 2014; Bassett and Berrospide, 2018; Begenau, 2020; Bahaj and Malherbe, 2020).

[4]For more information, see Annex A.1.

[5]The BCB (2022a) considers three levels of aggregation to calculate the concentration of RC5: banking and non-banking segment ($b1 + b2 + b3 + b4 + n1$), banking segment ($b1 + b2$), and commercial banking segment ($b1$). The business model category ($b1$) includes commercial banks, multiples with a commercial portfolio and savings banks; ($b2$), multiple banks without a commercial portfolio and investment banks; ($b3$), credit unions; ($b4$), development banks; and ($n1$), non-bank credit institutions.

a different regulatory environment[6].

In light of this scenario, two essential questions arise. The first question concerns the key factors that contribute to bank runs in the Brazilian financial system and the ways in which policymakers can proactively mitigate these risks to ensure financial stability. The second question examines how regulatory frameworks can be optimized to enhance banking competition, efficiency of financial intermediation, and social welfare, all while maintaining a stable financial system. To address these questions, this work has two main objectives. First, we estimate different measures to understand the systemic risk that individual banks contribute to the market. Second, we propose a novel bank run model that simulates the loss distribution (LD), the probability of default of the DIA, and an optimized CR and CAR within the Brazilian banking system.

This paper begins by estimating several well-known measures in bank risk literature, including (i) Systemic Expected Shortfall (SES), (ii) Systemic Risk (SRISK), and (iii) Conditional Value-at-Risk (CoVar) (Adrian and Brunnermeier, 2016; Acharya et al., 2010, 2017; Engle, 2018). However, given that these measures are primarily applicable for listed banks, and our analysis seeks to encompass all financial institutions in the Brazilian banking system, we determine a bank's PD using granular public accounting data based on the structural model of Merton (1974).

After considering these market-based measures, we developed a theoretical framework similar to those used to calculate portfolio risk in banking organizations. Our focus is on proposing a bank run model, first employing it on a reduced sample (RS) of only listed banks, and, subsequently, on a full sample (FS) that encompasses all banks covered by the Brazilian DIA. We also model an optimized CR for the RS, utilizing a heterogeneous CR framework as in Alexandre et al. (2022). Through this approach, our aim is to provide comprehensive insights into the dynamics of the Brazilian banking system.

This work is divided into five sections, beginning with this Introduction. Section 2.2 outlines the theoretical framework of all our analyzes to estimate the systemic risk of each bank, the bank run model, and the optimized CR for the Brazilian financial system. Section 2.3 presents the data used in our models. Section 2.4 offers the results and a discussion of our findings, and Section 2.5 concludes with the final remarks of this paper.

[6]For more information, see FGC for the list of banking failures in Brazil and FDIC for the list of banking failures in the USA. The percentages were calculated considering the member institutions of the deposit insurance in each economy in September 2022.

2.2 Systemic Risk Measures

This section presents the theoretical framework used to estimate each financial institution's contribution to systemic risk. We divided this section into two categories: (i) traditional measures based on market data, and (ii) structural models that use publicly available balance sheet information. Our contribution focuses on the development of a bank run model that considers different channels of contagion used to calculate the loss distribution, the PD of the DIA, and the optimized capital requirement for the Brazilian banking system. To this end, it is important to address the theoretical grounds for understanding systemic risk.

Systemic risk refers to the risk of a financial crisis or market failure that affects the stability of the financial system and has widespread effects on the economy as a whole. This type of risk is of particular concern in financial systems due to the interconnectedness of financial institutions, which can amplify the effects of individual failures and propagate shocks throughout the system as a negative domino effect. In recent years, the concept of systemic risk has gained increasing attention from academics and policymakers, as the GFC highlighted the potential consequences of systemic failures (Brunnermeier and Oehmke, 2013; Adrian and Brunnermeier, 2016; De Bandt and Hartmann, 2019).

Measuring systemic risk in the financial system is, to some extent, related to measuring banking risk. This makes risk measures at the bank level a natural starting point for understanding systemic risk. The purpose of these measures is to reduce a large amount of data into a single meaningful statistic, providing a summary and risk ranking of each financial institution. Over the last two decades, especially after the implementation of Basel II bank regulations and the GFC, a large literature has explored and created metrics to measure and capture systemic risk based on market data such as: (i) Systemic Expected Shortfall (SES), (ii) Systemic Risk (SRISK), (iii) Conditional Value-at-Risk (CoVar), and many others[7] (Brunnermeier and Oehmke, 2013; Adrian and Brunnermeier, 2016; Acharya et al., 2010, 2017; Brownlees and Engle, 2017; Engle, 2018). In order to explore these metrics, the next subsection addresses the foundations and definitions of each concept.

[7]Bisias et al. (2012) categorize and contrast more than 30 systemic risk measures. For more information on the extensive literature on systemic risk and its connection with the current regulatory debate, see also Jobst and Gray (2013), Benoit et al. (2017) and Silva et al. (2017).

2.2.1 Market-Based Metrics

2.2.1.1 Fundamentals

Because we are essentially dealing with market data when constructing the main measures in the literature, we begin by defining some components that they have in common. Consider N financial institutions and let r_t^i be the log return of the daily stock price of bank i at time t, $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Also, let r_t^m be the log return of the daily market index to which all banks participate, which in our case is captured by the Bovespa index (Ibovespa)[8]. Then, the bank and market return processes are given by 2.1.

$$r_t^i = \mu^i + \varepsilon_t^i \qquad \text{and} \qquad r_t^m = \mu^m + \varepsilon_t^m \tag{2.1}$$

In which μ is the expected return and ε_t is a zero-mean white noise. Although serially uncorrelated, the series ε_t does not need to be serially independent and can present conditional heteroskedasticity (Zakoian, 1994). Thus, to model this time-varying volatility, we used the the GJR-GARCH (p, q) model (Glosten et al., 1993) that allows shock asymmetry through γ and assumes a specific parametric form of $\varepsilon_t = \sigma_t z_t$ for this conditional heteroskedasticity, where z_t is a standard Gaussian and the volatility σ_t is given by 2.2.

$$\sigma_t^2 = \omega + \sum_{k=1}^{p} (\alpha_k + \gamma_k I_{t-k})\varepsilon_{t-k}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{2.2}$$

Where

$$I_{t-k} \coloneqq \begin{cases} 0 & \text{if } r_{t-k} \geq \mu, \\ 1 & \text{otherwise} \end{cases} \tag{2.3}$$

Following Brownlees and Engle (2017) and Engle (2018), we used GJR-GARCH (1,1) for all models and also found that this specification best fits our data. Furthermore, all parameters $(\mu, \omega, \alpha, \gamma, \beta)$ were simultaneously estimated by maximizing the log likelihood

[8] Ibovespa is the main performance indicator of the stocks traded in B3 and lists major companies in the Brazilian capital market. It was created in 1968 and, over the last 50 years, has set a benchmark for investors around the world. Ibovespa is reassessed every four months and is the result of a theoretical portfolio of stocks. It is composed of stocks and units of companies listed on B3 that meet the criteria described in its methodology, accounting for about 80% of the number of trades and the financial volume of our capital markets. For more information, see B3 (2023).

and the best model was selected based on the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC). In addition, we use the strictly positive restriction on all parameters beside μ. The assumption that z_t is Gaussian does not imply that the returns are Gaussian. Although their conditional distribution is Gaussian, their unconditional distribution presents excess of kurtosis (fat tails). However, if the true distribution is different, the Quasi-Maximum Likelihood (QML) estimator is still consistent (Glosten et al., 1993).

Despite the assumption that returns are serially uncorrelated, they may present contemporaneous correlation. In other words, from equation 2.1, defining a vector of zero-mean white noise as $\varepsilon_t = r_t - \mu$, $\sum_t := E_{t-1}[(r_t - \mu)(r_t - \mu)']$ may not be a diagonal matrix. Moreover, this contemporaneous variance may be time-varying, depending on past information. Therefore, the correlation between each bank i and the market index is captured by the GARCH-DCC model (Engle, 2002, 2009) and is estimated in two steps.

The first step of the GARCH-DCC model accounts for the conditional heteroskedasticity. It consists of estimating the conditional volatility σ_t^i using a GARCH (1,1) model (Engle, 1982; Bollerslev, 1986) for each one of the N bank series of returns r_t^i. Let $\boldsymbol{D}_t$ be the diagonal matrix with these conditional volatilities, that is, $\boldsymbol{D}_t^{i,i} = \sigma_t^i$ and, if $i \neq j$, $\boldsymbol{D}_t^{i,j} = 0$. Then the standardized residuals with unit conditional volatility, $\boldsymbol{\nu}_t$, are given by Equation 2.4 and the Bollerslev (1990)'s constant conditional correlation (CCC) estimator, $\overline{\boldsymbol{R}}$, is given by Equation 2.5.

$$\boldsymbol{\nu}_t := \boldsymbol{D}_t^{-1}(r_t - \mu) \tag{2.4}$$

$$\overline{\boldsymbol{R}} := \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\nu}_t \boldsymbol{\nu}_t^{\prime} \tag{2.5}$$

The second and final step consists in generalizing Bollerslev (1990)'s CCC to capture dynamics in the correlation, giving origins to the dynamic conditional correlation (DCC). Assuming the standardized residuals are jointly Gaussian and let $\boldsymbol{Q}_t^{i,j}$ be the correlation between r_t^i and r_t^j at time t, the DCC correlations are given by Equation 2.6.

$$\boldsymbol{Q}_t = \overline{\boldsymbol{R}} + \sum_{k=1}^{p} \boldsymbol{\alpha}_k(\boldsymbol{\nu}_{t-k}\boldsymbol{\nu}_{t-k}^{\prime} - \overline{\boldsymbol{R}}) + \sum_{j=1}^{q} \boldsymbol{\beta}_j(\boldsymbol{Q}_{t-j} - \overline{\boldsymbol{R}}) \tag{2.6}$$

Where both parameters, $\boldsymbol{\alpha}, \boldsymbol{\beta} > 0$, are simultaneously estimated by maximizing the log likelihood and the best model was selected based on the BIC and AIC information criterion.

Once these common components of volatility and correlation are defined, the following subsections present the concept and definition of the main metrics of systemic risk that use market data.

2.2.1.2 Systemic Expected Shortfall

The systemic expected shortfall (SES) proposed by Acharya et al. (2010, 2017) measures the expected capital shortfall of a bank conditional on a substantial reduction in the capitalization of the banking system and also provides a ranking for systemically risky banks. The theoretical approach of this model considers that the aggregate capital shortfall of the financial sector imposes a negative externality on the real economy. Thus, in order to estimate the capital shortfall of the financial sector, the first step is to estimate the marginal expected shortfall (MES) of a bank. MES is the expected short-term equity loss of a financial institution conditional on the market taking a loss greater than its Value-at-Risk (VaR) at $\alpha\%$. Taking into account the parameters established in Section 2.2.1.1, the MES is given by 2.7.

$$MES_{it} = E_{t-1}(r_t^i | r_t^m < C) \tag{2.7}$$

Where $C = q^{\alpha}(r_t^m)$ is a threshold corresponding to the tail risk in the market at time t. Note that the definition of Acharya et al. (2010, 2017) considers the market return r_t^m as the value-weighted average of all bank returns in the market, that is, $r_t^m = \sum_{i=1}^{N} w_t^i r_t^i$, where w_t^i denotes the relative market capitalization of the bank i. However, although it would be possible to reconstruct the r_t^m through each r_t^i, we opt to use the main benchmark of the Brazilian stock exchange, Ibovespa, as our r_t^m for purposes of comparability of results with the literature.

Also, define the expected shortfall (ES) of the market as the expected loss in the index conditional on this loss being greater than C, i.e., $ES_t = E_{t-1}(r_t^m | r_t^m < C)$. Thus, note that MES of one bank is the derivative, or the marginal effect, of the market's ES with respect to the bank's market share (or capitalization). Therefore, the MES of a bank in this case can be interpreted as reflecting its contribution and participation to overall systemic risk. The higher the MES, the higher is the individual contribution of the bank to the risk of the financial system. However, it is still possible to define the same statistic even if the observed bank does not participate in the market index. Rather than a measure of how a particular bank's risk adds to the market risk, the MES in this other case should be viewed as a measure of the sensitivity (or resilience) of the bank's stock price to exceptionally bad market events.

After defining the ES and MES, the SES can be understood as an extension of the MES. Specifically, the SES quantifies the extent to which a bank's equity falls below its target level, set as a fraction k of assets, during a systemic crisis where the overall capital is less than a fraction k of the total assets. Formally, the SES is given by 2.8.

$$\frac{SES_{it}}{W_{it}} = kL_{it} - 1 - E_{t-1}\Big(r_t^i \mid \sum_{i=1}^{N} W_{it} < k\sum_{i=1}^{N} A_{it}\Big) \tag{2.8}$$

Where A_{it} denotes the total assets, W_{it} the market capitalization or market value of the equity, and $L_{it} = A_{it}/W_{it}$ the leverage. In particular, in this work we set the prudential capital fraction k according to the prudential segment of each Brazilian bank[9]. Acharya et al. (2010, 2017) show that the conditional expectation term can be expressed as an increasing linear function of MES, given by Equation 2.9.

$$SES_{it} = (kL_{it} - 1 + \theta MES_{it} + \Delta_i)W_{it} \tag{2.9}$$

In which θ and Δ_i are constant terms.

2.2.1.3 Systemic Risk

Taking into account the significant negative externalities that undercapitalization of large financial institutions has on the real economy, Brownlees and Engle (2017) proposed a systemic risk metric called SRISK to measure the capital shortfall of a bank conditional on a severe market decline[10]. Although this contribution is related to the SES measure proposed by Acharya et al. (2010, 2017), the authors argue that SRISK does not require

[9]In accordance with Resolution CMN n.º 4.553/2017, the Central Bank of Brazil segments financial institutions and other licensed entities into five categories: S1 includes universal banks, commercial banks, investment banks, foreign exchange banks, and federal savings banks with a size equal to or exceeding 10% of the GDP or those engaging in relevant international activity; S2 encompasses the same bank categories as S1 but with a size less than 10% and equal to or exceeding 1% of the GDP, and other institutions with a size equal to or exceeding 1% of the GDP; S3 is for institutions with a size below 1% but equal to or exceeding 0.1% of the GDP; S4 consists of institutions with a size less than 0.1% of the GDP; and S5 includes institutions below 0.1% of the GDP that utilize a simplified optional methodology for determining minimum equity requirements, excluding the banks listed in S1 and S2.

[10]For an early report on SRISK, see Acharya et al. (2010, 2012) and Engle (2016).

structural assumptions or observation of the realization of a systemic crisis for estimation, making this a viable *ex-ante* measure with higher predictive power than SES. Furthermore, the authors argue that aggregate SRISK also provides early warning signals of distress in indicators of real activity.

The SRISK calculation is analogous to the stress tests that are applied to financial institutions but only uses market data that are publicly available. The capital shortfall is the variable introduced to measure the distress of a financial institution, which can be defined as the difference between the regulatory capital reserves that the bank needs to hold and the bank's equity. Formally, the capital shortfall of a bank i in time t is given by 2.10.

$$CS_{it} = kA_{it} - W_{it} = k(D_{it} + W_{it}) - W_{it} \tag{2.10}$$

Where D_{it} is the book value of the debt and the other parameters were previously defined in Section 2.2.1.2. The capital shortfall can be understood as the negative of the working capital of the bank. When the capital shortfall is negative, that is, the bank has a capital surplus, the bank functions properly. However, when this metric is positive, the bank experiences distress.

Because the interest is to predict the capital shortfall of a bank in the case of a systemic event, Brownlees and Engle (2017) uses the same concept of a market decline below a threshold C proposed by Acharya et al. (2010, 2017). Thus, the definition of SRISK as the expected capital shortfall conditional on a systemic event is given by 2.11.

$$\begin{aligned} SRISK_{it} &= E_t(CS_{it+h} | r^m_{t+1:t+h} < C) \\ &= kE_t(D_{it+h} | r^m_{t+1:t+h} < C) - (1-k)E_t(W_{it+h} | r^m_{t+1:t+h} < C) \end{aligned} \tag{2.11}$$

In which $\{r^m_{t+1:t+h} < C\}$ is the systemic event with probability α, $r^m_{t+1:t+h}$ is the multiperiod arithmetic market return between periods $t+1$ and $t+h$, and $SRISK_{it} \geq 0$. In order to compute this expectation, the authors assume that, in the case of a systemic event, debt cannot be renegotiated, which implies that $E_t(D_{it+h} | r^m_{t+1:t+h} < C) = D_{it}$. Finally, using this assumption in Equation 2.11 results in the final Equation 2.12 for SRISK.

$$SRISK_{it} = \max\{0\ ;\ kD_{it} - (1-k)W_{it}(1 - LRMES_{it})\} \tag{2.12}$$

Where $LRMES_{it} = -E_t(r^i_{t+1:t+h}|r^m_{t+1:t+h} < C)$ is the Long Run MES, that is, the expectation of the bank equity multiperiod arithmetic return conditional on the systemic event, in which $r^i_{t+1:t+h}$ is the multiperiod arithmetic return of bank equity between periods $t+1$ and $t+h$. In other words, when a stress scenario occurs, the equity decreases by a rate called the LRMES. Note from Equation 2.12 that SRISK is higher for banks that are larger, more leveraged, and with higher sensitivity to market declines.

To estimate LRMES, Engle (2018) proposed a direct approach in which $LRMES_{it} = 1 - \exp\left[(\widetilde{\beta}^i_t)\log(1-\theta)\right]$, where $\widetilde{\beta}^i_t$ is the nested dynamic conditional beta (DCB) and θ is the expected drop in the market during a financial distress event[11]. The initial DCB approach described by Engle (2016) considers that beta is the product of a correlation between the bank return and the market return, $\rho^{i,m}_t$, times the standard deviation of the bank return, σ_{it}, and market return, σ_{mt}. Because all of these three values are potentially time-varying, the author estimates them through GJR-GARCH and GARCH-DCC models. Formally, the DCB estimates of beta is given by 2.13.

$$\widehat{\beta}^i_t = \rho^{i,m}_t \left[\frac{\sigma_{it}}{\sigma_{mt}}\right] \tag{2.13}$$

We also follow Engle (2018) by building an artificial model that nests both the constant beta and DCB through Equation 2.14.

$$r^i_t = (\phi_1 + \phi_2\widehat{\beta}^i_t)r^m_t + \varepsilon^i_t \tag{2.14}$$

The estimates of both coefficients are made by assuming a MA(1) GJR-GARCH error term to construct a weighted least squares (WLS) model[12]. We would expect $\phi_2 = 0$ for a constant beta or $\phi_1 = 0$ for DCB, but because both hypotheses can be rejected, it is preferable to consider some combination of constant and time-varying beta (Engle, 2018). Then, the nested DCB used to estimate LRMES and SRISK is given by 2.15.

$$\widetilde{\beta}^i_t = (\widehat{\phi}_1 + \widehat{\phi}_2\widehat{\beta}^i_t) \tag{2.15}$$

[11]Works like Acharya et al. (2012) and Engle (2018) consider $\theta = 40\%$ over 6 months for the US market, a measure that is further discussed in Section 2.2.2.2
[12]The weights are the inverse of the variance of the MA(1) GJR-GARCH model of ε^i_t (Engle, 2016).

2.2.1.4 Conditional Value-at-Risk

The conditional value-at-risk (CoVaR) is a concept proposed by Adrian and Brunnermeier (2016) and represents the $p\%$ quantile of the market return. This quantile is determined based on the distribution when a specific bank's return matches its VaR at $q\%$, given that the preceding state variables are set at M. In other words, it is a measure of how sensitive the overall market is to a decline in a particular financial institution. First, the definition of VaR of bank i with probability q, such as 5%, is given by 2.16.

$$P(r_t^i < -VaR_{i,t}^q \mid M_{t-1}) = q \tag{2.16}$$

The CoVaR of the banking system in quantile p when a particular bank i has a market decline equal to its $VaR_{i,t}^q$ is given by 2.17.

$$P(r_t^m < -CoVaR_{m|i,t}^{p,q} \mid r_t^i = -VaR_{i,t}^q, M_{t-1}) = p \tag{2.17}$$

Adrian and Brunnermeier (2016) define the risk contribution of the bank i to the overall market as the incremental change in its risk relative to its median state, that is, $\Delta CoVaR_{m|i,t}^q = CoVaR_{m|i,t}^{q,q} - CoVaR_{m|i,t}^{q,0.5}$. In addition, they also proposed the use of quantile regression for its efficiency and simplicity in estimating CoVaR, evaluating the estimated equation with the independent variables of interest, which is given by Equation 2.18 and Equation 2.19.

$$r_t^m = \alpha_i^q + \beta_i^q r_t^i + \varepsilon_t \tag{2.18}$$

$$CoVaR_{m|i,t}^q = \widehat{\alpha}_i^q + \widehat{\beta}_i^q VaR_{i,t}^q \tag{2.19}$$

Therefore, the $\Delta CoVaR_{m|i,t}^q$ is then given by Equation 2.20.

$$\Delta CoVaR_{m|i,t}^q = \widehat{\beta}_i^q (VaR_{i,t}^q - VaR_{i,t}^{0.5}) \tag{2.20}$$

2.2.2 Bank Run Model

In Section 2.2.1, we presented the main systemic risk measures based on market data, which are limited to listed banks and pose significant challenges for developing economies where few banks are publicly traded. Specifically, in the Brazilian context, these measures can be applied to only 13.19% of banks, raising concerns about the representativeness and completeness of evaluating the systemic risk of the whole financial system[13]. To address this limitation, in this section we present the main contribution of our paper by proposing a bank run model that accounts for single, or idiosyncratic, probability of default of banks based on their public balance sheet structure and a systemic risk process in which additional defaults occur through different channels of contagion.

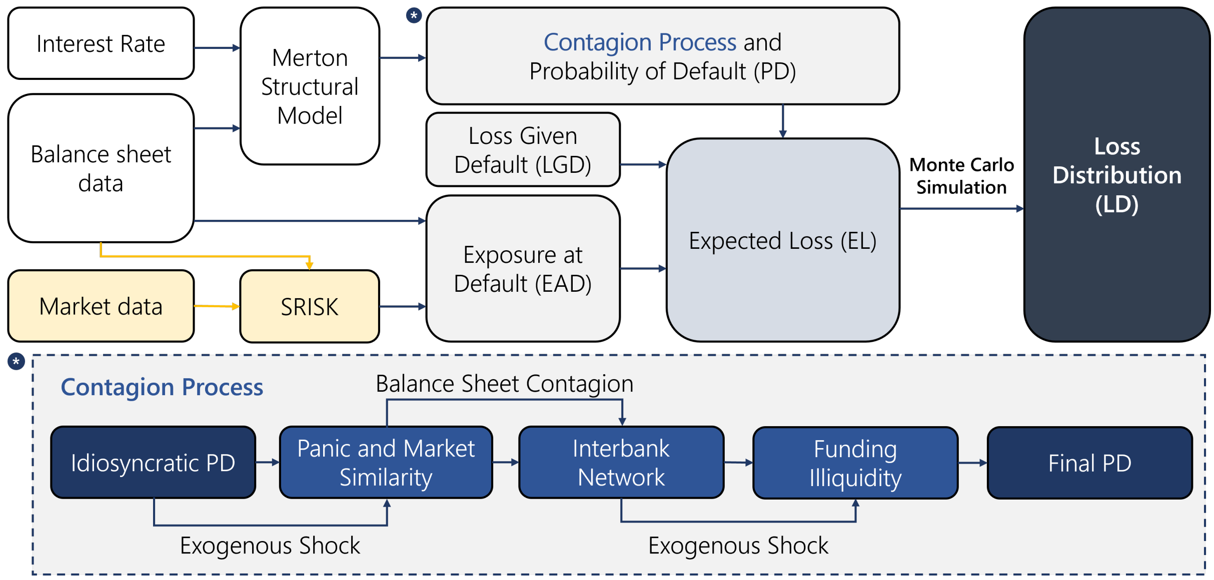
The two fundamental channels of contagion considered by the literature of financial stability and systemic risk are: (i) the exposure channel, and (ii) the informational channel (Greenwood et al., 2015; Paltalidis et al., 2015; Hurd, 2016; Souza et al., 2016; De Bandt and Hartmann, 2019; Jackson and Pernoud, 2021; Radev, 2022). These fundamental channels can work in conjunction as well as independently, and we model them through three different processes: (i) panic due to deposit withdrawals and market similarity, (ii) interbank network and (iii) funding illiquidity. Several works in the literature focus on the banking contagion process through one of these three channels[14], but few studies have shown the significant role they all play together in understanding the full effect of contagion (Paltalidis et al., 2015; Glasserman and Young, 2016; Anderson et al., 2019; Jackson and Pernoud, 2021). It is with this last understanding that we propose our model.

Our approach follows a similar theoretical framework in which portfolio risk is calculated in banking organizations and how banking losses are estimated in deposit insurance schemes (Lehar, 2005; Gupton et al., 2007; De Lisa et al., 2011; Bellini, 2017; O'Keefe and Ufier, 2017; Parrado-Martínez et al., 2019; Matt and Andrade, 2019; Fernández-Aguado et al., 2022). We implement this model in two distinct scenarios. First, we use a reduced sample consisting only of listed banks. Second, we utilize a comprehensive sample that includes all banks covered by the Brazilian DIA. Due to these different samples, the inputs that determine the exposure at default (EAD) and the contagion process in the probability of default vary accordingly. The diagram of the structure and steps of our model is shown in Figure 2.1.

[13]Because the Brazilian financial market is highly concentrated, as discussed in this paper, although only 13.19% of covered member institutions are listed, they represent 79.89% of total assets. This concentration is commonly used as an argument to make approximate inferences about the entire system.

[14]See Martin (2006), Gorton (2010), Greenwood et al. (2015), Robatto (2019) and Gertler et al. (2019) for panic due to deposit withdrawals and market similarity, Acemoglu et al. (2015), Cabrales et al. (2017) and Alexandre et al. (2022) for the interbank network and Ferrara et al. (2019) and Ardekani et al. (2020) for funding illiquidity.

Figure 2.1: Diagram of the structure and steps of the Bank Run model.



* Note: The contagion process amplifies the idiosyncratic probability of default from the Merton (1974)'s structural model through the mechanisms and shocks described in Section 2.2.2.2. The model input for EAD varies based on the sample: yellow arrows for listed banks and blue for covered member institutions. The loss given default (LGD) parameter is set to one for all Brazilian banks, as discussed in Section 2.2.2.3.

In Section 2.2.2.1, we detail the model used to calculate the idiosyncratic probability of default, which is the same as the one presented in Chapter 1. Then, Section 2.2.2.2 introduces our model designed to account for various channels of contagion. Finally, Section 2.2.2.3 presents how the loss distribution is computed taking into account the components of the bank run model.

2.2.2.1 Probability of Default

Following Chapter 1 of this thesis, we utilize the structural model of Merton (1974) to estimate the idiosyncratic PD for each financial institution (Souza et al., 2015, 2016; Guerra et al., 2016; Coccorese and Santucci, 2019; da Rosa München, 2022). Recalling previous definitions and using the Black and Scholes (1973)'s model, the option's payoff for the equity holder at time T is given by 2.21.

$$E_{it} = max(A_{it}\, \mathcal{N}(d_{1it}) - DB_{it}\, e^{-r_t T} \mathcal{N}(d_{2it}), 0) \tag{2.21}$$

Where A_{it} is the asset value, r_t is the risk-free interest rate, $\mathcal{N}(.)$ is the cumulative normal distribution function,

$$d_{1it} = \frac{ln(\frac{A_{it}}{DB_{it}}) + (r_t + \frac{\sigma^2_{Ait}}{2})T}{\sigma_{Ait}\sqrt{T}} \quad \text{and}$$

$$d_{2it} = d_{1it} - \sigma_{Ait}\sqrt{T} = \frac{ln(\frac{A_{it}}{DB_{it}}) + (r_t - \frac{\sigma^2_{Ait}}{2})T}{\sigma_{Ait}\sqrt{T}},$$

in which σ_{Ait} denotes the volatility of the assets. Thus, the idiosyncratic PD_{it} of a FI in a time horizon T, calculated in $t = 0$, is given by 2.22.

$$\begin{aligned} PD_{it} &= P(DB_{it} \geq A_{it}) \\ &= P(\ln DB_{it} \geq \ln A_{it}) \\ &= \mathcal{N}(-d_{2it}) \\ &= \mathcal{N}\left[-\frac{ln(\frac{A_{it}}{DB_{it}}) + (r_t - \frac{\sigma^2_{Ait}}{2})T}{\sigma_{Ait}\sqrt{T}}\right] \end{aligned} \tag{2.22}$$

Note that the probability of default is the area under the default barrier, that is, a fraction of total liabilities.

2.2.2.2 Contagion

The following subsections present our approach for the two fundamental channels discussed in the literature, which are (i) the exposure channel and (ii) the informational channel (Greenwood et al., 2015; Paltalidis et al., 2015; Hurd, 2016; Souza et al., 2016; De Bandt and Hartmann, 2019; Jackson and Pernoud, 2021; Radev, 2022). As in Diamond and Dybvig (1983), the runs we consider are runs in the entire banking system and not on a single bank. In addition, as in Abergel et al. (2013), we do not consider the possibility of partial default in our model[15].

We begin by establishing three channels that are used to incorporate the contagion effect in the event of a bank default: (i) panic due to deposit withdrawals and market similarity, (ii) interbank network, and (iii) funding illiquidity. Note that these channels are only activated after a single bank defaults, although they do not necessarily generate a banking contagion process if the rest of the system is resilient[16]. The generation of a

[15]In a model with partial default, represented by a default level on liabilities, it is necessary to introduce a different mechanism in which the non-defaulting banks have to sell part of their assets in order to satisfy some solvability ratio constraints. For more information, see Eisenberg and Noe (2001).

[16]Note that a run on an individual bank may not have aggregate effects if depositors simply shuffle their

banking contagion process or a systemic risk event occurs when the default of an individual bank or multiple idiosyncratic defaults deteriorates the other banks in the system through these channels to the point of generating additional defaults.

An important component to be defined in the simulation of our bank run model is the deterioration of the probability of default when there is a contagion process or financial crisis. Engle (2018) estimates a 40% deterioration over 6 months in the PD when there is a financial distress event for the US market and Greenwood et al. (2015) estimated a 28% deterioration over 18 months in the market capitalization of European banks after the financial crisis. For this work, utilizing the structural model by Merton (1974), as described in 2.2.2.1, we estimated the aggregate PD of the Brazilian financial system and found a 27% deterioration during periods of financial distress[17]. Therefore, we calibrate our shocks so that all three contagion processes, as defined by our model, amplify the average bank default rate by 27%.

2.2.2.2.1 Panic

The extensive literature on banking panics, beginning with Diamond and Dybvig (1983), emphasizes the phenomenon where sudden withdrawals, triggered by the expectation of defaults, can force the bank to liquidate many of its assets at a loss and, ultimately, fail. The subsequent losses worsen the conditions of banks, including those that are financially strong, reinforcing the flight to liquidity and making the process self-fulfilling. In this scenario, it is the forced liquidation at fire sale prices during a run that pushes these banks into bankruptcy, categorizing a situation of short-term illiquidity (Martin, 2006; Gorton, 2010; Greenwood et al., 2015; Kiss et al., 2018; Allen et al., 2019; Robatto, 2019; Anderson et al., 2019; Gertler et al., 2019).

During such forced short-term liquidation, there is also a contagion effect on asset prices for other banks that hold assets or portfolios similar to those that suffer from these fire sales, leading to a deterioration of their balance sheets. If the loss of an affected bank is so severe that it is unable to meet its minimum capital requirement, then the bank will

funds from one bank to the others in the system (Gertler et al., 2019). Our model captures this dynamic when the shocks from a contagion process over the PD of banks that do not idiosyncratically default are not strong enough to lead to additional defaults.

[17]We constructed a weighted PD considering the value of total deposits of each FI from December 2000 to September 2022. The 27% is the average deterioration during periods of economic recessions dated by CODACE (2023). For more information, see Section 3.2.1 and Figure 3.2 of Chapter 3.

have to sell some of its assets with a haircut, increasing the downward spiral in market prices (Nier et al., 2007; Cont et al., 2013; Huang et al., 2013; Glasserman and Young, 2016; Caccioli et al., 2018; Pichler et al., 2021).

Considering that panic behavior is fundamentally a problem of depositor expectation and liquidity risk, with the potential to affect the overall stability of similar banks within the industry, this matter could be addressed from the complete information on the connection that each depositor and bank has with all other banks in the system (Brown et al., 2016; Anginer and Demirgüç-Kunt, 2019; Jackson and Pernoud, 2021). In this context, after one bank idiosyncratically defaults, the cross-information of depositors and assets could help anticipate which banks are most susceptible to the panic effect and estimate the magnitude of this illiquidity shock. However, given the private and confidential nature of this information, we choose to model this behavior using a cluster approach based on balance sheets and market data, in which the utilization of information varies depending on the sample under consideration. Through this approach, our objective is to simulate a short-term panic shock affecting similar banks within the industry.

We choose to utilize the K-means clustering (Lloyd, 1957; Jancey, 1966; MacQueen, 1967) to estimate the clusters in the banking system. K-means clustering is an unsupervised learning algorithm for finding K pre-specified clusters and cluster centers (i.e. centroid) in a set of unlabeled data. Let $C_1, \ldots, C_K$ be the sets containing the indices of the observations in each banking cluster. These sets satisfy both (i) $C_1 \cup C_2 \cup \cdots \cup C_K = \{1, \ldots, n\}$ and (ii) $C_k \cap C_{k'} = \emptyset$ for all $k \neq k'$. Thus, (i) means that each observation belongs to at least one of the K clusters and (ii) means that the clusters are non-overlapping, i.e., no observations belong to more than one cluster (Hastie et al., 2009).

For instance, if the ith observation is in the kth cluster, then $i \in C_k$. Thus, a good clustering is one for which the within-cluster variation, $W(C_k)$, is as small as possible. Therefore, the objective of the K-means algorithm is to solve 2.23.

$$\min_{C_1,\ldots,C_K} \left\{ \sum_{k=1}^{K} W(C_k) \right\} \tag{2.23}$$

In order to solve 2.23, we first need to define the specification of the within-cluster variation. The most common choice to define this concept in the literature is the Hartigan and Wong (1979)'s algorithm, which defines the total within-cluster variation as the sum of squared Euclidean distances between items and the corresponding centroid. This specification is given by 2.24.

$$W(C_k) = \sum_{\mathbf{x}_i \in C_k} (\mathbf{x}_i - \mu_k)^2 \tag{2.24}$$

Where $\mathbf{x}_i$ denotes a vector belonging to the cluster C_k, with $\mathbf{x}_i = \{x_{1i}, x_{2i}, \ldots, x_{Ni}\}$, where x_{ji} is attribute j of bank i, and μ_k is the mean value of the points assigned to the cluster C_k. The algorithm that solves the optimization problem when combining equations 2.23 and 2.24 can be decomposed into two major steps: (i) initially assigns a random number, from 1 to K, to each of the observations and (ii) iterates until the cluster assignments stop changing. In the second step, (a) for each of the K clusters, the algorithm computes the cluster centroid and (b) assigns each observation to the cluster whose centroid is the closest in terms of Euclidean distance. This algorithm is used iteratively until the local optimum is reached. (Hastie et al., 2009).

Because the K-means technique requires a pre-specification of the expected number of clusters, there are a variety of other direct and statistical methods used to define the within-cluster variation in order to find the optimal number of clusters. The two most common are the silhouette and the gap statistics. The silhouette statistic (Kaufman and Rousseeuw, 1990) measures how well an observation is clustered and estimates the average distance between clusters. A high average silhouette width indicates a good clustering and the optimal number of clusters K is the one that maximizes the average silhouette over a range of possible values for K. For observation i, let $a(i)$ be the average distance to other points in its cluster, and $b(i)$ the dissimilarity between i and its closest clusters to which it does not belong. Then, the silhouette statistics, $s(i)$, is defined by 2.25.

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \tag{2.25}$$

On the other hand, the gap statistic (Tibshirani et al., 2001) compares the total within-cluster variation for different values of K with their expected values under the null reference distribution of the data. The estimate of the optimal clusters will be the value that maximizes the gap statistic, in which $\log(W_k)$ falls the farthest below this reference curve. Thus, considering $E_n[\log(W_k)]$ the expectation under a sample of size n, the gap statistic is defined by 2.26.

$$\text{Gap}_n(k) = E_n[\log(W_k)] - \log(W_k) \tag{2.26}$$

Once the clustering technique and the optimal number of clusters have been defined, we select some balance sheet and market data, depending on the sample under consideration, to distinguish these clusters and simulate the panic behavior in the Brazilian banking system. It is important, however, that these clusters have an internal consistency with the prudential segmentation of S1 to S5 and the business model category established by the BCB (2023a).

The variables used for clustering the reduced sample of listed banks are: (i) total deposits over total assets, (ii) loan, lease and other credit operations by risk level over total assets, (iii) RWA, (iv) SRISK, and (v) ΔCoVaR. On the other hand, the variables used for clustering the full sample are: (i) total deposits over total assets, (ii) loan, lease and other credit operations by risk level over total assets, (iii) RWA, (iv) CAR, (v) proxy of the net stable funding ratio, and (vi) prudential segment.

In our simulation of the banking panic process, when one bank within a cluster idiosyncratically defaults, the other banks in that cluster receive a shock in their probability of default, as discussed in Section 2.2.2.2. This dynamics continues until all three contagion processes established by our model converge and increase the default of banks, on average, by 27%.

2.2.2.2.2 Interbank Network

In addition to the panic effect, several works in the literature emphasize the significant role of interbank networks in the contagion process of systemic risk, suggesting that the probability of default cascades increases with the magnitude of interbank exposures (Anand et al., 2015; Bardoscia et al., 2015; Acemoglu et al., 2015; Silva et al., 2016; Anand et al., 2018; Anderson et al., 2019; Ferrara et al., 2019). However, it is important to note that network interconnectedness can be understood by both correlated portfolios, through common asset holdings among banks, and counterparty risk, through direct bilateral exposures between banks (Nier et al., 2007; Cont et al., 2013; Huang et al., 2013; Glasserman and Young, 2016; Pichler et al., 2021; Jackson and Pernoud, 2021). We model these two phenomena by (i) considering market and asset correlation in the first short-term contagion channel described in Section 2.2.2.2.1, and (ii) considering the interbank market as a second mid-term contagion channel, which is addressed in this section.

In the interbank market, banks lend to each other based on interbank interest rate. In Brazil, the interbank deposit, known as DI, is a private fixed income instrument traded exclusively among financial institutions. It plays an important role in assisting banks with cash management, either as a mechanism for raising funds or allocating surplus resources. These securities, often referred to as CDI, have high liquidity, no incidence of taxes on profitability, and carry a very low risk, usually associated with the creditworthiness of the participating banks in the market. The negotiation between banks determine the DI rate, which serves as a reference for most of the fixed income securities offered to investors. This interbank interest rate is the main benchmark of the market and is obtained by calculating the weighted average of the rates of the prefixed, extragroup (different conglomerates), and one-day transactions between financial institutions. Furthermore, it is worth noting that, in the interbank market, banks can also trade one-day repurchase agreement contracts backed by federal securities with similar purpose of interbank deposit, but shifting the counterparty risk for sovereign risk (B3, 2023).

Along with its capacity as the monetary, regulatory, and supervisory authority, the Central Bank of Brazil is also responsible for controlling and monitoring the liquidity of the banking system. In line with its role as lender of last resort (LOLR), the BCB's provision of liquidity support contributes to the credibility of domestic currency and to the financial system's stability. The BCB's liquidity facility (LFL) includes discount window lending operations, which are based on non-government issued securities with financial institutions maintaining a reserve or settlement account at the BCB. The LFL's operations can be a short-term standing facility (LLI), normally intraday and overnight, or a long-term facility (LLT), through discretionary approval and enhanced operational process. The loans are secured against high-quality assets as collateral, which may not present immediate liquidity (BCB, 2021b).

Considering that the complete information about the Brazilian interbank network is private and only known by the Central Bank of Brazil, we reconstruct the interbank network based on the following main properties: (i) density, (ii) average degree, and (iii) assortativity. Among the several methods available to estimate the matrix of bilateral exposures[18], we employed an adapted version of the minimum density (MD) method

[18] See, for instance, Anand et al. (2018) for a comprehensive survey of different estimation methods.

proposed by Anand et al. (2015). We chose this approach because it best aligns with the actual, known properties of the Brazilian interbank network from March 2010 to September 2015. These known properties are: (i) a density range of [0.03,0.07], (ii) an average degree between [4.6,7.8] for all banks and [21,26] for large banks, and (iii) an assortativity values between [-0.31, -0.54][19] (Castro Miranda et al., 2014; BCB, 2016; Souza et al., 2016; Silva et al., 2016; Anand et al., 2018; Alexandre et al., 2022).

The minimum density method proposed by Anand et al. (2015) is a heuristic procedure for allocating links that combines elements from information theory with economic incentives to produce networks that preserve the realistic characteristics of interbank activity. The authors argue that the MD approach is suitable for sparse networks, which is appropriate for the Brazilian financial market, and is able to reconstruct it by minimizing its cost of linkages. Let c be the fixed cost of establishing a link, $\mathbb{X} \in [0, \infty)^{N \times N}$ the matrix of bilateral aggregated exposure values, X_{ij} the unknown exposure of bank i to bank j and N the number of banks. The total observable interbank assets of bank i are $A_i = \sum_{j=1}^{N} X_{ij}$ and the total observable liabilities of bank j are $L_j = \sum_{i=1}^{N} X_{ij}$. Then, the MD problem is given by 2.27.

$$\min_{\mathbb{X}} c \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbf{1}[X_{ij} > 0] \qquad \text{s.t.} \tag{2.27}$$

$$\sum_{j=1}^{N} X_{ij} = A_i \quad \forall i = 1, 2, \ldots, N$$

$$\sum_{i=1}^{N} X_{ij} = L_j \quad \forall j = 1, 2, \ldots, N$$

$$X_{ij} \geq 0 \quad \forall i, j$$

Where the integer function $\mathbf{1}$ equals one only if bank i lends to bank j, and zero otherwise. This problem is equivalent to finding the network with lowest average degree, i.e., the lowest number of edges, under given constraints.

However, because 2.27 is computationally expensive to solve, the authors proposed a

[19]The last official release of the assortativity and degree metric for the Brazilian interbank network made by the BCB was in April 2016, covering data up to September 2015. All other data come from studies using the real Brazilian interbank network. For more information, see BCB (2016).

smooth value function, $V(\mathbb{X})$, which is high whenever the network $\mathbb{X}$ has a few links and satisfies the asset and liability constraints. First, the authors soften the constraints by assigning penalties for deviations from the marginal of each bank, which is given by 2.28.

$$AD_i \equiv \left(A_i - \sum_{j=1}^{N} X_{ij}\right) \qquad \text{and} \qquad LD_j \equiv \left(L_j - \sum_{i=1}^{N} X_{ij}\right) \tag{2.28}$$

In which LD_j measures bank j's current deficit, i.e., how much its bilateral borrowing falls short of the total amount it needs to raise, L_j, and AD_i is measures bank i's current surplus. When they are introduced into the objective function 2.27, the problem becomes a maximization given by 2.29.

$$V(\mathbb{X}) = -c\sum_{i=1}^{N}\sum_{j=1}^{N} \mathbf{1}[X_{ij} > 0] - \sum_{i=1}^{N}(\alpha_i^2 AD_i) - \sum_{j=1}^{N}(\delta_j^2 LD_j) \tag{2.29}$$

Where α_i are the weights for assets deviations and δ_j are the weights for liabilities deviations. Note that sparse networks $\mathbb{X}$ that minimize marginal deviations are more efficient and achieve higher values in the objective function $V(\mathbb{X})$.

In addition to being sparse, interbank networks are typically disassortative, i.e. small banks seek to match their lending and borrowing needs through relationships with larger banks that are well placed to satisfy those needs. Such behavior contributes to sparsity, since most small banks can satisfy their needs with a single large counterparty. The authors capture this information through the set of probabilities $Q \equiv \{Q_{ij}\}$ for the relative relationships between i and j. The probability that i lends to j increases if either i is a large lender to a small borrower j or i is a small lender to a larger borrower j. This process is given by 2.30.

$$Q_{ij} \propto \max\left\{\frac{AD_i}{LD_j}, \frac{LD_j}{AD_i}\right\} \tag{2.30}$$

To ensure that the most likely network solutions are disassortative, the authors propose a probability distribution, $P(\mathbb{X})$, that should be close to the prior Q. This mechanism is given by 2.31.

$$\max_{P} \sum_{\mathbb{X}} P(\mathbb{X})V(\mathbb{X}) + \theta R(P \parallel Q) \tag{2.31}$$

Where the scaling parameter θ emphasizes the weight placed on finding solutions with characteristics similar to the prior matrix Q and $R(P \parallel Q) = \sum_{\mathbb{X}} P(\mathbb{X}) \log(P(\mathbb{X})/Q(\mathbb{X}))$ is the relative entropy between P and Q. The solution to this problem can be obtained from the first-order conditions given by 2.32, stating that a candidate $\mathbb{X}$ has a higher likelihood of being chosen than the prior Q specifies if the departure from Q raises the value of the maximization problem given by 2.29.

$$P(\mathbb{X}) \propto Q(\mathbb{X}) e^{\theta V(\mathbb{X})} \tag{2.32}$$

Note that while the prior Q codifies the probabilities for picking links, there are no restrictions to the values one should allocate to selected links. However, because each bank has a maximum exposure limit of 25% of its Tier 1 capital that it can have with another bank in the Brazilian financial market since the publication of Resolution 4.667 in July 2018 (BCB, 2018a), we adapted the MD method proposed by Anand et al. (2015) by including one more step in the iteration process. After the end of each iteration of the MD procedure, we checked whether the solution exceeded the maximum exposure limit of each bank's assets in relation to its Tier 1 capital. Only in cases where the solution have exceeded is that we limit the asset's exposure of each bank to 25%, and then we update the simulated network until the total interbank market volume has been allocated. The optimal solution is the one that satisfies all these restrictions and produces topological features that match the moments of the Brazilian interbank network.

In terms of the topological features, the three most important are: (i) density, (ii) average degree, and (iii) assortativity[20]. Density is the number of undirected links as a percentage of the total number of links (excluding self-loops), which can also be seen as the sparsity of a network. The degree, or valency, is a strictly local measure that corresponds to the number of counterparties each bank connects in the financial network. Thus, degree can be interpreted as a proxy of bank's portfolio diversification inside the financial network. Since this is a bank-level network measurement, it is common to report the average value of all participating banks (Souza et al., 2016; Anand et al., 2018).

Assortativity is a global measure of the network that presents the correlation between the number of counterparties of pairs of banks that have operations with each other. If this measure is positive, it means that, in the assessed network, banks with many counterparties usually carry out operations with banks that have many counterparties. On the other

[20] For more information on topological information of a interbank network, see Silva et al. (2016).

hand, a negative measure, denoted disassortativity, reveals the predominance of financial operations between pairs of banks with different total numbers of financial operations. Usually, large banks, which have many operations and act as money centers, interconnect with small banks, which have few operations and act as investors or borrowers in the interbank. Furthermore, financial networks with negative assortativity show the existence of core-periphery structures. In Brazil, the nucleus is mainly composed of the group of large banks, whose members are highly connected to other members of the core and also intermediate operations between members of the periphery (Souza et al., 2016; Silva et al., 2016; Anand et al., 2018).

Once the interbank network is reconstructed considering all these properties, we proceed to the estimation of the impact of shock scenarios. This includes considering potential contagion effects and evaluating the systemic importance of a bank within the network. Our aim is to understand how stress dynamics within the banking system can resonate throughout the entire network. Among the different methods proposed by the literature for this purpose, we opt for the differential DebtRank algorithm (Bardoscia et al., 2015) for its effective shock propagation dynamics (Souza et al., 2016; Silva et al., 2017; Poledna et al., 2021).

The original DebtRank algorithm first proposed by Battiston et al. (2012) is inspired by the feedback centrality measure[21] and assumes that losses are linearly propagated between connected banks. Suppose a network of mutually exposed banks in which each of these banks has assets and liabilities, among which a fraction is related to the counterparties within the network, and a capital buffer. If a bank suffer assets losses greater than its capital buffer, it becomes insolvent and will not be able to honor any of its short-term liabilities, a scenario in which the bank defaults. On the other hand, if the losses are lower than its capital buffer, suppose 90%, the bank will be in distress and will not pay its creditors a proportional part (10%) of its liabilities, which characterizes a stress measure. In the first case, the creditors of the default bank, in turn, will suffer losses and undergo through the same dynamics. This feedback process continues until the whole system converges.

[21]Feedback centrality measures are those in which the centrality of a node, or vertex, depends recursively on the centrality of its neighbors. The recursiveness criterion effectively forces the centrality of each node to depend on the entire network structure through feedforward/feedback mechanism. In this sense, the original DebtRank of Battiston et al. (2012) is not formally a feedback centrality measure because it does not propagate a second- and high-order round of impacts that come from cycles or multiple routes in the network. These components are considered in the differential DebtRank of Bardoscia et al. (2015).

Formally, the original DebtRank method models the interbank market as a direct network $\mathcal{G} = \langle \mathcal{B}, \mathcal{E} \rangle$, in which the banks compose the vertex $\mathcal{B}$ and the exposures between them compose the set of edges $\mathcal{E}$. Again, these links are represented by a weighted adjacency matrix $\mathbb{X}$, where the (i, j)th entry, X_{ij}, represents the amount bank i lends to bank j, i.e., the exposure of bank i to bank j. In a similar notation, the total value of the asset invested by i in funding activities is $A_i = \sum_{j \in \mathcal{B}} X_{ij}$ and the relative economic value of bank i is given by $\varphi_i = A_i / \sum_{j \in \mathcal{B}} A_j$, $\varphi_i \in [0, 1]$, which is the fraction of i's assets with respect to the total assets in the interbank market.

Also, each bank i has a capital buffer against shocks, E_i, which is represented by the Common Equity Tier 1 (CET1)[22]. If $E_i \leq \gamma$, where $\gamma > 0$, the bank defaults. If vertex j defaults, all of the neighbors i will suffer losses corresponding to their exposure towards j, given by X_{ij}. When $X_{ij} > E_i$, then vertex i defaults. The local impact of j on i is $W_{ij} = \min(1, V_{ij})$, where $V_{ij} = X_{ij}/E_i$ is the bank's stress level, so that if i's losses exceed its capital, the local impact is 1. Intermediate values within the interval $(0, 1)$ for W_{ij} lead i into distress, but not into default.

The presence of cycles in the network inflates the computed impacts by counting the local impact of a node on another more than once. To avoid the distortion caused by this double-counting, the original DebtRank algorithm evaluates the additional stress caused by some initial shock using a dynamic system, allowing only a single impact propagation per each node. It maintains two state variables for each bank $i \in \mathcal{B}$: (i) $h_i(t) \in [0, 1]$ and (ii) $s_i(t) \in \{U, D, I\}$. $h_i(t)$ is the stress level of i and $s_i(t)$ is a categorical variable that denotes the state of i. U, D, and I stand for undistressed, distressed, and inactive, respectively. The update rules of the dynamic system are given by 2.33 and 2.34.

$$h_i(t) = \min\left\{1, h_i(t-1) + \sum_{j \in \mathcal{D}(t)} W_{ij} h_j(t-1)\right\} \tag{2.33}$$

[22]Although Battiston et al. (2012) uses the Tier 1 (the sum of the Common Equity Tier 1 and the Additional Tier 1) as the capital buffer against shocks, we used only the CET1 as it is the component of highest quality capital and can absorb losses immediately as they occur (BCBS, 2011). CET1 comprises the core capital of a bank and consists mostly of issued equity and retained earnings, being used by investors to assess a bank's solvency. Even considering the capacity to absorb losses of some AT1 instruments, which includes noncumulative, nonredeemable preferred stock and related surplus and qualifying minority interest, they have a lower quality compared to the CET1 instruments, especially in periods of financial distress (Ramirez, 2017; Couaillier, 2021).

$$
s_i(t) = \begin{cases} D, & \text{if } h_i(t) > 0 \text{ and } s_i(t-1) \neq I, \\ I, & \text{if } s_i(t-1) = D, \\ s_t(t-1), & \text{otherwise} \end{cases} \tag{2.34}
$$

In which $t \geq 0$ and $\mathcal{D}(t) = \{j \in \mathcal{B} \mid s_j(t-1) = D\}$. Note that the sum of 2.33 occurs only for those distressed banks in the previous iteration. However, once distressed, they become inactive in the next iteration due to 2.34. Thus, they are not able to propagate further stress. Observe that the algorithm must converge for a sufficiently large number of steps $T \gg 1$ due to the min(.) operator, which places upper bounds on the bank's stress levels, and the non-decreasing property of $h_i(t)$, derived from the non-negative entries of the vulnerability matrix V_{ij}. We compute the resulting DebtRank due to the initial shock scenario $h(0)$ using equation 2.35.

$$
DR(h(0)) = \sum_{i \in \mathcal{B}} \Big[h_i(T) - h_i(0) \Big] \varphi_i \tag{2.35}
$$

Note that, by removing the initial stress $h(0)$ from the DebtRank computation in 2.35, it conveys the notion of additional stress given an initial shock scenario. However, the great drawback of this formulation is that it prevents banks from diffusing second- and high-order rounds of stress. This means that, once a vertex propagates stress, it will not be able to propagate additional stress due to other subsequent impacts that it receives, which can lead to a severe underestimation of the stress levels of banks.

To address the limitations of the original DebtRank algorithm, Bardoscia et al. (2015) introduced an enhancement known as the differential DebtRank. This improved method still captures cycles or multiple routes in the vulnerability network and, therefore, prevents double-counting of stress by employing stress differentials between successive iterations. As a result, at each iteration, banks are only allowed to propagate the stress increase that they received from the previous iteration. Using this mechanism, financial stress is never double-counted because differentials are always innovations from one iteration to another. Again, once a bank default at time t, it no longer propagates financial stress during the dynamic process for $t + k$, $k > 0$. Therefore, substituting the stress levels in 2.33 by stress differentials results in equation 2.36.

$$h_i(t) = \min\left\{1, h_i(t-1) + \sum_{j\in\mathcal{B}} W_{ij}\Delta h_j(t-1)\right\} \tag{2.36}$$

Where $\Delta h_j(t-1) = h_j(t-1) - h_j(t-2)$ is the stress differential of the bank j in the previous iteration $t-1$ and $h(t) = 0 \; \forall \; t < 0$. In the beginning of the iteration process, $h(0)$ is an ex-ante input that denotes the initial stress scenario, or the list of shocks based on the bank's probability of default. Thus, we then compute the differential DebtRank value of an initial shock scenario in 2.35 using the convergent stress values of 2.36.

There are two important dinstinctions between the differential DebtRank of Bardoscia et al. (2015) compared to the original formulation of Battiston et al. (2012). First, the sum index in 2.36 runs through all the banks, such that there is no need to maintain states in the dynamic system. Second, instead of only one propagation immediately after the shock has been received, they could propagate shocks until all connected banks in the network default, which makes a formally feedback centrality measure. The dynamics now reaches global equilibrium only when the direct and indirect neighborhoods of each bank are considered, taking into account multiple routes and network cycles when establishing the final stress levels of banks.

2.2.2.2.3 Funding Illiquidity

It is well known in the literature, especially after the GFC, as discussed in this work, that banks suffer from liquidity and funding risk and the importance of these channels for the process of contagion and systemic risk. Bank asset and liability structures proved to be highly vulnerable to deposit runs, market shocks and breakdowns in funding markets (Acemoglu et al., 2015; Paltalidis et al., 2015; Venkat and Baird, 2016; Ferrara et al., 2019; Wen et al., 2023).

However, it is important to note the difference between short-term and long-term liquidity risk and to distinguish their respective roles in the financial contagion process. Short-term liquidity risk (less than one year) can arise from various sources, such as panic behavior and asset fire sales, as mentioned in Section 2.2.2.2.1, and the direct exposures in the interbank network, as mentioned in Section 2.2.2.2.2. On the other hand, long-term liquidity risk (greather than one year) is related to sufficiently stable sources of funding or the inability of banks to raise funds when needed, such that longer-term liabilities are assumed to be more stable than short-term liabilities to mitigate the risk of future funding stress (BCBS, 2014; Venkat and Baird, 2016; Ardekani et al., 2020; Wieser, 2022).

Some mechanisms have been created to prevent and mitigate all the risks addressed so far. The design of a deposit insurance scheme (DIS), for instance, constitutes an integral part of the financial safety net provided to the banking system and is intended to prevent runs on individual banks by depositor. If the DIS is credible and depositors expect that they will receive their money back from the insurance fund, regardless of what other depositors do or whether they are last in line for reimbursement, then they no longer have incentives to run and withdraw their funds. In the event of bank failure, it also limits losses to depositors and reduces the risk that a run on one bank might undermine confidence in others through contagion effects. Thus, the existence of a credible DIS contributes to the reduction of the funding cost, especially long-term, of banks (Diamond and Dybvig, 1983; Allen et al., 2011; Anginer and Demirgüç-Kunt, 2019; Freixas and Parigi, 2019).

The most fundamental deposit insurance scheme is the paybox mandate in which the deposit insurer is only responsible for the reimbursement of insured deposits. Most countries with an established DIS have improved its legal and operational characteristics over time, usually by expanding the mandate and powers and strengthening the role of the DIS within the financial safety net. (Ognjenovic, 2017; Kerlin, 2017). In Brazil, for instance, the deposit insurance agency, FGC, is a privately held company and exercises a paybox plus mandate, in which the additional resolution function (e.g., financial support) is also attributed to the deposit insurer (BCB, 2018b).

Regarding the Brazilian case, the situation that a private deposit insurer has limited resources raises the question about the DI's ability to withstand a strong systemic risk event. In worst-case scenarios where all available fund resources are consumed, even with the framework of lender of last resort played by the central bank in order to provide a potential source of liquidity for banks, the uncertainty about the credibility of the entire system in this case makes it difficult for financial institutions with poor long-term liquidity to raise short-term funding in the market, a combination that further increases their probability of default (Allen et al., 2011; Vazquez and Federico, 2015; Diamond and Kashyap, 2016; Ognjenovic, 2017; Kerlin, 2017; Ebrahimi Kahou and Lehar, 2017; Bouwman, 2019).

The Basel Committee on Banking Supervision also created other mechanisms in Basel III to reduce short-term liquidity risk and long-term financing risk. The two proposed quantitative liquidity standards are the Liquidity Cover Ratio (LCR) and the Net Stable Funding Ratio (NSFR). LCR reflects short-term liquidity soundness and requires banks to hold sufficient high-quality liquid assets (HQLA) to offset the net cash outflows in a liquidity stress scenario over 30 days. On the other hand, NSFR requires a minimum amount of available stable funding (ASF) relative to the required stable funding (RSF) over a one-year horizon. Both liquidity ratios have a minimum regulatory of 100%. Note that the implementation of LCR encourages a substitution from long-term illiquid assets to short-term liquid assets, which consequently eases bank runs. Furthermore, under NSFR the bank needs to finance illiquid assets with long-term funding, which can alter the bank's incentive to use less runnable deposits (BCBS, 2013, 2014; Diamond and Kashyap, 2016; Ebrahimi Kahou and Lehar, 2017).

As mentioned previously, considering that the different aspects of short-term liquidity risk are covered in Section 2.2.2.2.1 and Section 2.2.2.2.2, we address the long-term liquidity risk through the NSFR concept in all scenarios where the available resources from DI are consumed[23]. The NSFR equation is given by 2.37.

$$NSFR = \frac{ASF}{RSF} \tag{2.37}$$

Where available stable funding (ASF) is defined as the portion of capital and liabilities expected to be reliable over the time horizon and the required stable funding (RSF) is a function of the liquidity characteristics and residual maturities of the various assets held by the bank, as well as those of its off-balance sheet exposures (Chiaramonte et al., 2013; BCBS, 2014).

Therefore, in order to comply with a minimum regulatory of 100% for NSFR, banks can either increase their ASF or reduce their RSF. A natural option to increase ASF is to increase the proportion of long-term funding in the whole portfolio or to increase the Common Equity Tier 1 (CET1), which is the sum of Common Equity Tier 1 and Additional Tier 1. On the other hand, a natural option to reduce RSF is to shrink its balance sheet by changing the composition of its investments and loans or to change its assets to such combination that would result in a lower weight factor.

Considering that the NSFR is subject to national discretion to reflect jurisdiction-

[23]In those scenarios where banks suffer from the contagion channels of panic behavior and interbank network, but the DI still has enough resources to maintain the credibility and confidence in the financial system, then it can be argued that the remaining banks will not suffer from a long-term liquidity shock.

specific conditions, the Central Bank of Brazil defined the ILE (structural liquidity ratio) as its equivalent concept that has been in effect since October 2018 through the Resolution CMN 4.616 for banks in the S1 prudential segmentation (BCB, 2015, 2022b). However, because the construction of the ILE requires private and confidential data from each bank, we follow Takeuti (2020) to create a proxy for our NSFR based on publicly available balance sheet information[24]. Tables 2.1 and 2.2 detail each balance sheet data from BCB (2023a) used to calculate the proxy's for ASF and RSF.

Table 2.1: Balance sheets accounts used for calculating the ASF proxy.

Factor	Composition	Description
1.00	[60000002]	Equity
1.00	[70000009]	Gross Revenues
1.00	[80000006]	Gross Expenses
0.90	[41100000]	Demand Deposits
0.90	[41200003]	Saving Deposits
0.90	[41500002]	Time Deposits
0.60	[43000005]	Mortgage, real estate and others
0.60	[46000002]	Onlending
0.50	[42000006]	Repurchase Agreements

Source: Adapted from Takeuti (2020) and the methodological concepts presented in BCB (2016).

Table 2.2: Balance sheets accounts used for calculating the RSF proxy.

Factor	Composition	Description
1.00	[19000008]	Other Assets
1.00	[14000003]	Interbank Transactions
1.00	[15000002]	Interbranches Transactions
1.00	[20000004]	Fixed Assets
1.00	[23000001]	(-) Leased Assests
1.00	[18000009]	Other Receivables
0.85	[16000001]	Loans
0.65	[23000001]	Leased Fixed Assets
0.65	[17000000]	Leases
0.40	[13000004]	Securities and Derivatives

Source: Adapted from Takeuti (2020) and the methodological concepts presented in BCB (2016).

[24]The two works that proposes a proxy for ILE in Brazil, for the best of our knowledge, are Cardoso et al. (2019) and Takeuti (2020), but we evaluate that the work of Takeuti (2020) is more accurate in terms of the concept of NSFR.

Once the ASF and RSF metrics have been calculated to determine the NSFR proxy for each bank, we proceed to the simulation of the funding illiquidity risk. The dynamics of this third and final channel of contagion in the model is such that, in any scenario where the available resources from Brazilian DIA[25] are consumed, either through the reimbursement of covered deposits from banks that default for idiosyncratic reasons or through the first two channels described in Section 2.2.2.2.1 and Section 2.2.2.2.2, the remaining banks with NSFR less than one experience an additional shock to their probability of default. Finally, these three contagion mechanisms collectively contribute to a 27% average increase in the bank default rate, as described in Section 2.2.2.2.

2.2.2.3 Loss Distribution

As mentioned in the beginning of Section 2.2.2, our approach follows a similar theoretical framework in which portfolio risk is calculated in banking organizations and how banking losses are estimated in deposit insurance schemes (Lehar, 2005; Gupton et al., 2007; De Lisa et al., 2011; Bellini, 2017; O'Keefe and Ufier, 2017; Parrado-Martínez et al., 2019; Matt and Andrade, 2019; Fernández-Aguado et al., 2022). Thus, the last part of our model consists in estimating the loss distribution of the banking system taking into account the probability of default, the loss given default, and the exposure at default. To this end, we first define the expected loss (EL) of a given bank i, $i = 1, \ldots, N$, through equation 2.38.

$$EL_{it} = PD_{it}^{c} \times LGD_{it} \times EAD_{it} \tag{2.38}$$

In which $PD_{it}^{c} \in [0,1]$ represents the final probability of default after the contagion process has been applied to the idiosyncratic PD_{it}, as defined in Section 2.2.2.1 through equation 2.22. The $EAD_{it} \in [0, \infty)$ term is given by SRISK in Section 2.2.1.3 for listed banks in the reduced sample and by a share of covered deposits for the full sample. Lastly, $LGD_{it} \in [0,1]$ is assumed to be one for all banks in the Brazilian banking system[26].

[25] In our bank run model, the available resources of the DIA correspond to its equity capital. For more information, see FGC (2023).

[26] The concept of loss given default can also be understood as the proportion of non-recovery assets, i.e., $LGD_{it} = 1 - RR_{it}$, in which RR_{it} is the recovery rate. Note that in this framework the LGD is absorbed by the central bank or the deposit insurer of the economy. The assumption of absence of recovery on a one-year horizon when a bank default in Brazil can be argued considering the following aspects: (i) the high historical interest rate in Brazil diminishes the present value of futures recoveries, (ii) during the process of extrajudicial settlements there are mostly poor quality and illiquid assets left, which increases the time in line for recoveries, and (iii) in periods of financial distress it is even more difficult to liquidate assets without incurring great losses due to fire sales and market conditions. Therefore, considering

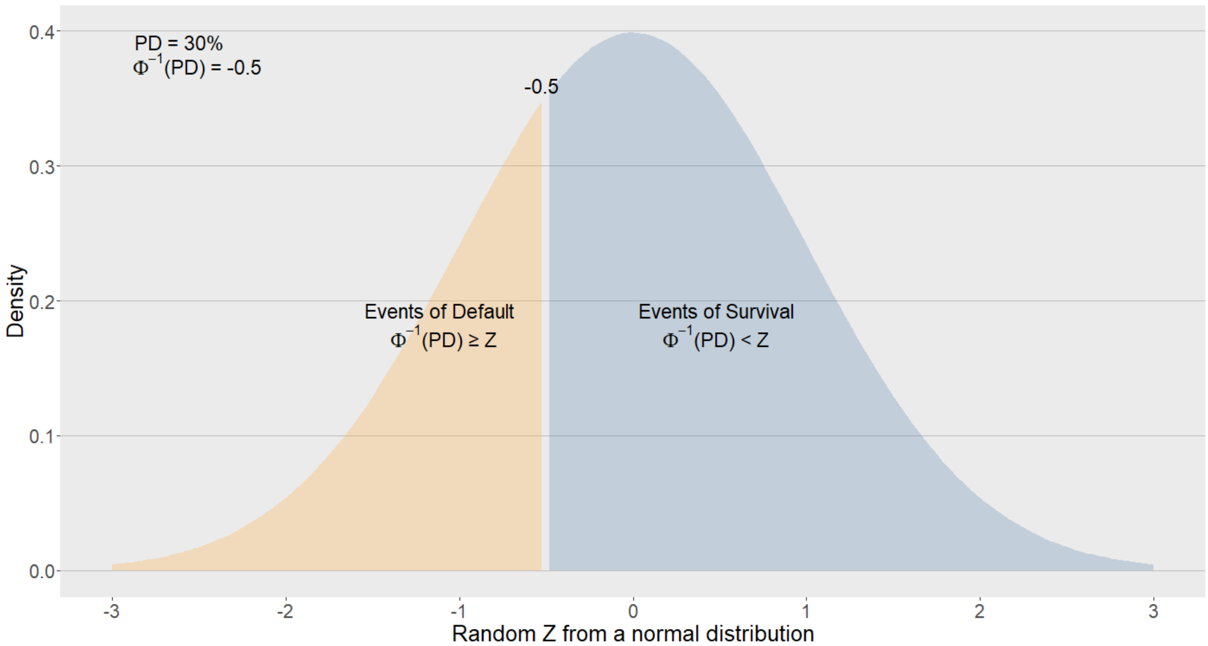
The construction of the economy's LD_t is the result of a Monte Carlo simulation over $\mathcal{L}_t^s$ for each scenario s through the bank run model, where $\mathcal{L}_t^s = \sum_{i=1}^{N} EL_{it}^s$, $s = 1, \ldots, \mathcal{S}$ and $t = 1, \ldots, T$. The value of $\mathcal{S}$ must be large enough to achieve convergence, which in our case is equal to 100.000. Let N be the number of banks, $\mathbf{\Phi}^{-1}(PD^c) \coloneqq [\Phi^{-1}(PD_{1t}^c), \ldots, \Phi^{-1}(PD_{Nt}^c)]$ the vector of the probit function, or the inverse of the standard normal cumulative distribution (quantile) function, of each bank's PD, and $\mathcal{Z}^s \coloneqq (z_1^s, \ldots, z_N^s) \sim \mathscr{N}(0,1)$ a vector of random variables for $\mathcal{S}$ different scenarios. Under the assumption of normality over bank asset values (Crouhy et al., 2000; Lehar, 2005; De Lisa et al., 2011; Guerra et al., 2016; Souza et al., 2016; Bellini, 2017; O'Keefe and Ufier, 2017; da Rosa München, 2022), the indicator function that generates the vector of bank's default $\mathcal{P}_t^s$ used to estimate $\mathcal{L}_t^s$ for each scenario s is given by 2.39.

$$\mathcal{P}_t^s \coloneqq \begin{cases} 1, & \text{if } z_i^s \leq \Phi^{-1}(PD_{it}^c) \quad \forall i = 1, \ldots, N, \\ 0, & \text{otherwise} \end{cases} \tag{2.39}$$

In this framework, the vector $\mathcal{Z}^s$ can be understood as the vector of initial shocks that gives origins to the idiosyncratic default PD_{it}, which will have a different combination of initial defaults for each s. After this initial shock, all three contagion processes described in Section 2.2.2 have their own dynamic according to the established criteria. Therefore, the final vector $\mathcal{L}_t^s$ of size $1 \times \mathcal{S}$ has different values for each s because the shock vector $\mathcal{Z}^s$ is also different for each s. For example, consider that bank i has $PD_{it} = 30\%$ and its associated probit function $\Phi^{-1}(PD_{it}) = -0.5$. If $z_i^1 = -0.6$ and $z_i^2 = -0.4$, both randomly drawn from a normal distribution through Monte Carlo simulations, then bank i in time t will idiosyncratically default in scenario 1 and will not idiosyncratically default in scenario 2. However, if bank i receives one shock of 27% during the contagion process, which gives $PD_{it}^c = 38\%$ and $\Phi^{-1}(PD_{it}^c) = -0.3$, then it will default in both scenarios at the end of the simulation to construct the loss distribution. Figure 2.2 gives a synthetic illustration of this default dynamics for each bank in each scenario.

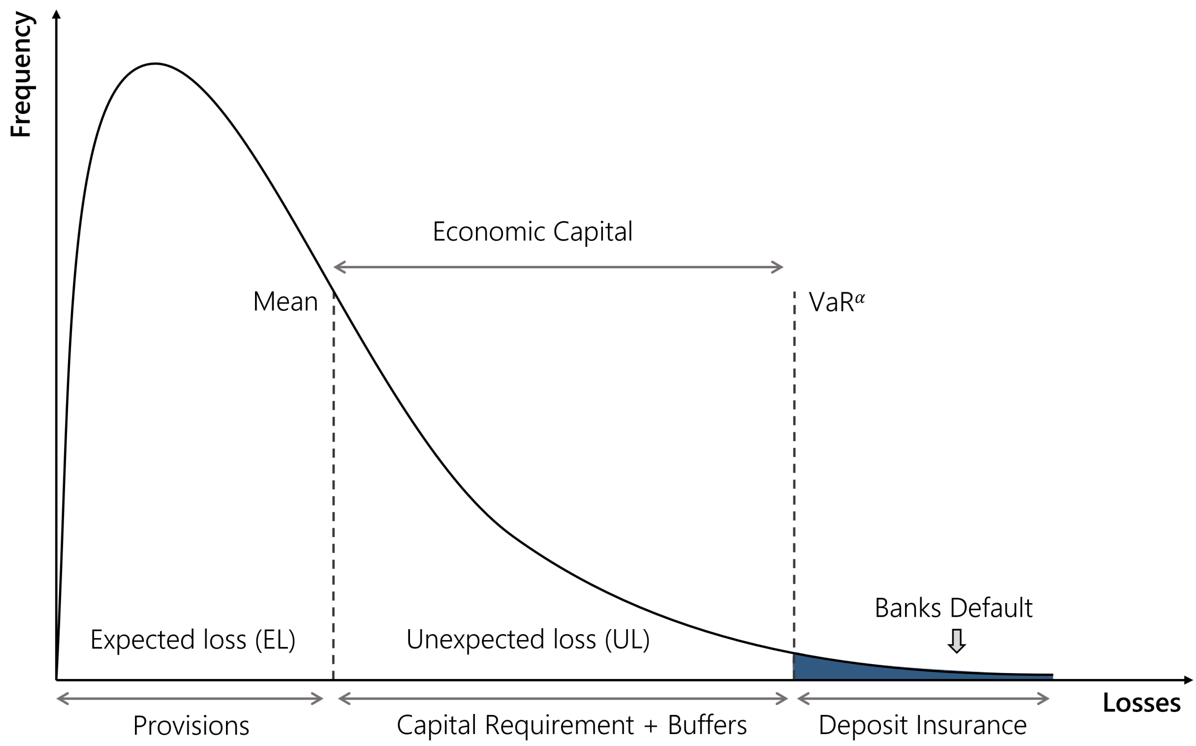
that we are modeling extreme distress events, the combination of all these factors further reduces the present value of recoveries on a one-year horizon, which underlies the proxy of one for the LGD of the Brazilian banking system. However, it is important to mention that if one would like to consider asset recoveries in the event of bank default, we can use the Merton (1974)'s model framework to calculate $LGD_{it} = 1 - (1-\varphi)\left[\frac{A_{it}}{DB_{it}} \frac{\mathscr{N}(-d_1)}{\mathscr{N}(-d_2)}\right] \exp{(r_t T)}$, in which φ represents administrative costs (Guerra et al., 2016).

Figure 2.2: Events of default from a normal distribution.



Once $\mathcal{L}_t^s$ is constructed for each s, the LD_t is obtained by ordering its values. The expected loss of LD_t corresponds to the average of the simulated scenarios and $VaR^{\alpha}_{LD_t}$ is calculated as the quantile α of the distribution. In contrast, the unexpected loss is derived as the difference between the $VaR^{\alpha}_{LD_t}$ and the expected loss. Because we are dealing with extreme events, the vector LD_t usually has a positive skew distribution, or a right-skewed distribution, and leptokurtosis, i.e., fat tails with excess of kurtosis (Gupton et al., 2007; Bellini, 2017). The representation of a hypothetical loss distribution of the banking system is shown in Figure 2.3.

Figure 2.3: Banking system loss distribution.



2.3 Data

For the first sample of listed banks, we utilized quarterly data from December 2000 to September 2022 for all 26 publicly traded Brazilian banks, resulting in a unbalanced panel data with 1.586 observations. On the other hand, for the second sample of covered banks, we utilized quarterly data from December 2000 to September 2022 for 244 Brazilian financial institutions, yielding an unbalanced panel data with 9,653 observations [27]. All balance sheet data employed in this study are publicly provided by the Central Bank of Brazil (BCB, 2023a), while the daily stock price and market capitalization data were obtained from Bloomberg for the same period.

Following the data structure presented in Chapter 1, the data set considers financial conglomerates and independent institutions until December 2014, and the prudential

[27]Although information is available from 2000:I-2000:III in the database, we used these first three quarters to calculate assets volatility since bank capital information is available from December 2000. Brazilian banks (with national headquarters) that are traded on foreign markets were also considered. Also, it is important to mention that, although we have a data set for 244 Brazilian financial institutions, which would result in 9.653 observations as in Chapter 1, we work only with a restricted subset that has market data available.

conglomerates and independent institutions before March 2015 with the business model category of $b1$, $b2$, $b4$, and $n1$, provided there are at least six valid observations in the studied period. The final data set of listed banks represents 79.89% of total assets, 80.46% of total credit, 89.71% of total deposits, and 13.19% of total member institutions in September 2022. For the interest rate, we used public data provided by B3, the Brazilian financial market infrastructure company (B3, 2023).

To estimate the probability of default on a one-year horizon for each FI using the Merton (1974)'s structural model, we applied the following variables: adjusted total assets for A, total liabilities to calculate DB, annualized interbank interest rate DI for r, and the annualized standard deviation of the logarithmic returns of adjusted total assets, that is, $\log(A_t/A_{t-1})$, for asset volatility σ_A. Table 2.3 presents the aggregate descriptive statistics of these variables for listed banks, Table 2.4 presents the aggregate descriptive statistics for the covered banks, and all balance sheet accounts are shown in Appendix B.1.

Table 2.3: Descriptive statistics for listed financial institutions.

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
ATA[a]	211.72	418.76	0.04	5.71	137.19	2,184.86
TL[a]	194.04	387.44	0.00	5.05	123.11	2,018.16
Loans[a]	89.23	185.98	0.00	2.44	39.73	978.56
DI[b]	11.19	5.11	1.90	6.93	14.13	26.23
AV	0.28	0.34	0.00	0.09	0.30	3.42
PD[b]	9.87	16.72	0.00	0.00	13.54	94.93
SRISK[a]	12.97	36.46	0.00	0.00	2.44	186.96
SES[a]	9.69	32.37	0.00	0.00	0.64	161.89
Beta	0.86	0.69	−0.59	0.45	1.09	6.01
MES	0.05	0.15	0.00	0.02	0.04	1.42
LRMES	0.30	0.15	−0.30	0.18	0.38	0.93
CoVaR	0.04	0.04	0.00	0.03	0.04	0.28
ΔCoVaR	0.01	0.03	0.00	0.00	0.01	0.24
II[a]	34.27	84.11	0.00	0.40	17.74	634.07
ID[a]	1.54	4.83	0.00	0.02	0.69	41.48
TD[a]	72.68	147.03	0.00	1.71	53.11	854.76
CET1[a]	10.26	26.70	0.00	0.00	2.37	142.78
TRC[a]	20.44	38.70	0.00	0.60	14.29	180.30
RWA[a]	123.00	234.02	0.00	3.62	88.56	1,225.17
DoA[b]	36.82	20.47	0.00	21.86	52.35	93.64
CoA[b]	42.04	21.06	0.00	29.30	54.53	97.87
LR[b]	4.15	8.31	0.00	0.00	7.11	100.00
CAR[b]	20.89	25.04	0.00	14.25	18.73	542.27
NSFR	1.61	19.92	0.15	0.88	1.12	793.98

Notes: The sample period runs from 2000:IV-2021:IV for 26 publicly traded Brazilian banks. *ATA* = adjusted total assets; *TL* = total liabilities; *Loans* = loan operations by risk level; *DI* = interest rate; *AV* = assets volatility; *PD* = probability of default; *SRISK* = systemic risk metric; *SES* = systemic expected shortfall; *Beta* = market beta; *MES* = marginal expected shortfall; *LRMES* = long-run marginal expected shortfall; *CoVaR* = conditional value-at-risk; $\Delta CoVaR$ = delta conditional value-at-risk; *II* = interbank investments; *ID* = interbank deposits; *TD* = total deposits; *CET1* = common equity tier I; *TRC* = total regulatory capital; *RWA* = risk-weighted assets; *DoA* = total deposits over total assets; *CoA* = total credit over total assets; *LR* = leverage ratio; *CAR* = capital adequacy ratio (tier I and II) and *NSFR* = net stable funding ratio.

[a] In BRL billion.

[b] In percentage.

Table 2.4: Descriptive statistics for covered financial institutions.

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
ATA[a]	44.78	192.25	0.00	0.34	10.29	2,184.86
TL[a]	40.77	177.47	0.00	0.26	8.95	2,018.16
Loans[a]	18.54	84.21	0.00	0.10	3.93	978.56
DI[b]	10.91	5.18	1.90	6.40	14.13	26.23
AV	0.42	0.41	0.00	0.14	0.58	6.03
PD[b]	16.20	20.28	0.00	0.02	29.98	99.84
II[a]	7.07	36.69	0.00	0.02	1.14	634.07
ID[a]	0.43	2.13	−0.00	0.00	0.15	41.48
TD[a]	13.48	65.37	0.00	0.02	2.16	854.76
CET1[a]	2.51	13.22	−0.05	0.00	0.52	164.12
TRC[a]	4.69	19.55	0.00	0.08	1.25	201.00
RWA[a]	26.85	109.74	0.00	0.27	6.61	1,225.17
DoA[b]	23.86	22.93	0.00	2.21	39.20	93.64
CoA[b]	45.23	31.33	0.00	17.28	69.83	263.99
LR[b]	10.11	150.30	−0.06	0.00	10.10	14,640.04
CAR[b]	31.01	40.77	0.38	14.83	30.81	1,151.84
NSFR	1.19	8.13	0.05	0.75	1.20	793.98

Notes: The sample period runs from 2000:IV-2022:III for 244 covered Brazilian banks. *ATA* = adjusted total assets; *TL* = total liabilities; *Loans* = loan operations by risk level; *DI* = interest rate; *AV* = assets volatility; *PD* = probability of default; *II* = interbank investments; *ID* = interbank deposits; *TD* = total deposits; *CET1* = common equity tier I; *TRC* = total regulatory capital; *RWA* = risk-weighted assets; *DoA* = total deposits over total assets; *CoA* = total credit over total assets; *LR* = leverage ratio; *CAR* = capital adequacy ratio (tier I and II) and *NSFR* = net stable funding ratio.

[a] In BRL billion.

[b] In percentage.

2.4 Results and Discussion

This section summarizes and discusses the empirical results obtained in three different exercises. First, we present some statistics on the banks' risk measures, the contagion process, and the estimation of the loss distribution of the banking system considering a reduced sample of only listed banks. Second, we compare the results of the RS with a full sample that considers all banks covered by the Brazilian DIA. Then, we propose and discuss an optimized capital adequacy ratio based on this framework considering the RS.

2.4.1 Banks' Probability of Default and Contagion Process in the Reduced Sample

Table 2.5 presents summary statistics for the risk measures estimated to build the LD of the Brazilian banking system in September 2022 with a RS of the 24 listed banks. The first risk measure is the PD given by the Merton (1974)'s structural model (structural probability of default) used to compute the idiosyncratic PD of each bank. The second measure is the banks' PD that represents the default rates in the Monte Carlo simulation. As expected and reflecting the construction of our model, the mean of the structural PD and the Banks PD without contagion are practically the same and equal to 12.4%. In addition, also note that the contagion process described in Section 2.2.2.2 increases the average probability of bank default by 27%.

Table 2.5: Summary statistics of structural PD and banks' PD with and without contagion in the reduced sample.

		Structural PD of Merton model (%)	Banks' probability of default (%)	
			Panel A: without contagion process	Panel B: with contagion process
Mean		12,45	12,42	15,75
Std. Deviation		21,40	5,22	6,49
Minimum		0,00	0,00	0,00
Maximum		83,37	37,50	45,83
Percentiles:	25%	0,00	8,33	12,50
	50%	1,99	12,50	16,67
	75%	13,34	16,67	20,83

Notes: This table shows statistics of PD estimates for 24 Brazilian banks in September 2022. Structural PD represents the estimates of the Merton (1974)'s model used as an input for the idiosyncratic PD of each bank. Banks' PD are the default rates of the sample banks in the Monte Carlo simulation. In details, Panel A shows summary statistics of Banks' PD with only idiosyncratic default (without contagion process). Panel B presents the same statistics when all the contagion process described in Section 2.2.2.2 are considered.

The clusters of the 24 Brazilian banks for September 2022 are presented in Table 2.6. We specify four clusters based on the statistics described in Section 2.2.2.2.1 and taking into account the relative consistency of their size, operation, regulation, and market similarity. These results were used in the first contagion process (panic) when the idiosyncratic default of a bank within a cluster causes a shock to the other banks' PD within the cluster. Note that banks classified in the S1 prudential segment were clustered in the same group.

Table 2.6: Cluster segmentation of the Brazilian banking system.

Cluster	Bank	Ticker	DoA[b]	CoA[b]	RWA[a]	SRISK[a]	ΔCoVaR
1	Itaú	ITUB4 BZ	38,7	40,3	1225,17	94,44	0,0148
	Bradesco	BBDC4 BZ	35,9	39,2	988,41	67,18	0,0144
	Santander	SANB11 BZ	39,7	47,0	637,46	48,44	0,0114
	BB	BBAS3 BZ	34,7	40,5	1039,39	185,10	0,0136
	CEF	CXSE3 BZ	34,9	62,7	704,62	168,25	0,0039
2	BTG	BPAC11 BZ	26,6	26,3	300,75	0,00	0,0117
	XP	XP US	17,0	15,5	37,38	0,00	0,0077
	B3	B3SA3 BZ	3,8	0,0	0,17	0,00	0,0174
	BR Partners	BRBI11 BZ	17,4	3,5	2,68	0,00	0,0032
	Nordeste	BNBR3 BZ	16,2	22,1	78,78	0,46	0,0026
3	Porto Seguro	PSSA3 BZ	1,5	96,8	14,76	0,00	0,0093
	Alfa	BRIV4 BZ	24,4	58,2	19,37	1,67	0,0025
	ABC	ABCB4 BZ	19,1	49,5	41,35	1,44	0,0114
	Amazônia	BAZA3 BZ	14,4	52,2	36,72	2,03	0,0044
4	Nubank	NU US	86,7	33,7	24,22	0,00	0,0131
	Inter	BIDI11 BZ	48,8	48,0	24,04	1,18	0,0027
	BMG	BMGB4 BZ	53,1	52,4	22,99	3,06	0,0076
	Modal	MODL11 BZ	45,7	19,6	5,51	0,00	0,0023
	Pine	PINE4 BZ	47,9	30,8	7,23	1,26	0,0055
	Banrisul	BRSR6 BZ	57,1	41,9	51,56	7,64	0,0099
	Banestes	BEES4 BZ	54,3	19,3	12,78	2,45	0,0027
	Mercantil	BMEB4 BZ	71,7	71,3	8,55	0,25	0,0004
	Mercado Crédito	MELI34 BZ	84,4	22,5	0,89	0,00	0,0018
	Est. Sergipe	BGIP4 BZ	77,4	46,6	4,97	0,56	0,0029

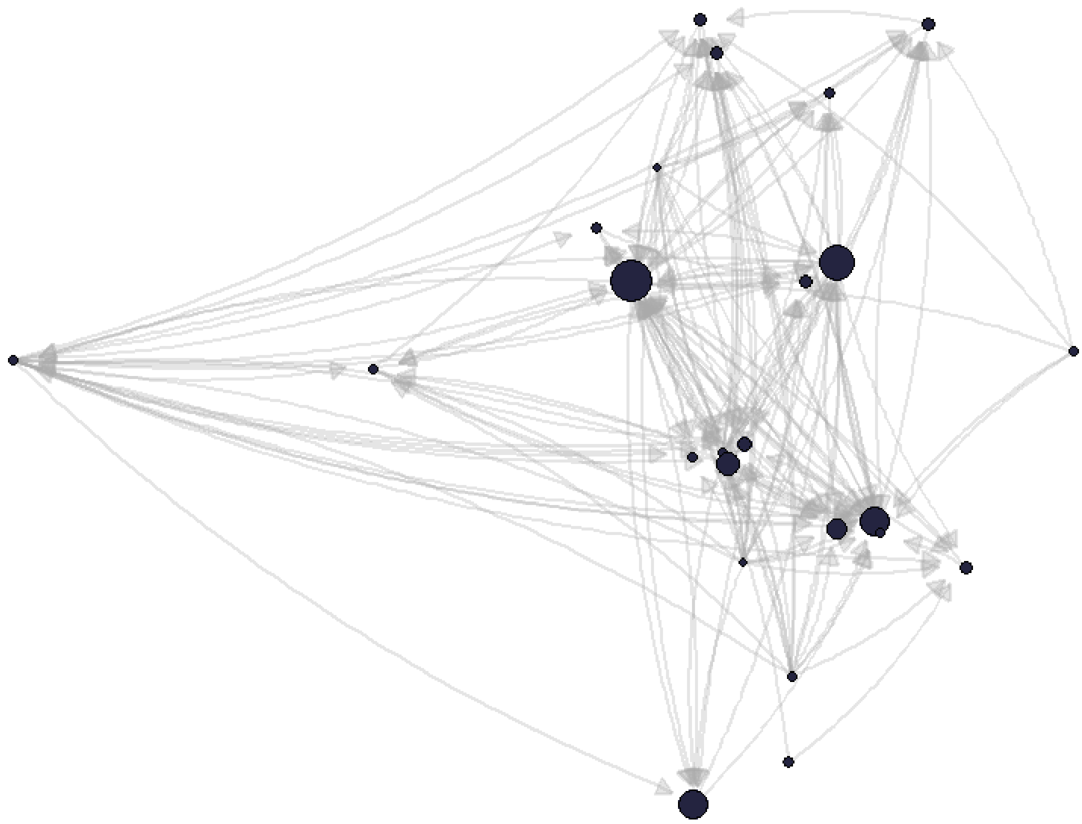
Notes: This table shows the four estimated clusters for 24 Brazilian banks in September 2022. *DoA* = total deposits over total assets; *CoA* = total credit over total assets; *RWA* = risk-weighted assets; *SRISK* = systemic risk metric; $\Delta CoVaR$ = delta conditional value-at-risk.
[a] In BRL billion.
[b] In percentage.

The estimated interbank network for the Brazilian economy in September 2022 is shown in Figure 2.4. The main statistics of our estimated network, such as density, assortativity, r, and average degree, are 0.31, -0.54 and 14.4, respectively. In terms of comparison, the density, assortativity and average degree reported by BCB and other works using the true Brazilian interbank network vary between [0.03,0.07], [-0.31,-0.54] and [4.6,7.8] (or [21,26] for large banks), respectively, during March 2010 and September 2015[28] (Castro Miranda et al., 2014; BCB, 2016; Souza et al., 2016; Silva et al., 2016; Anand et al., 2018; Alexandre et al., 2022). Note that even considering 13.19% of FIs, the adapted minimum density

[28]The last release of the assortativity and degree metric of the Brazilian interbank network made by the BCB was in April 2016 with data until September 2015. All other data comes from works that use the true Brazilian interbank network. For more information, see BCB (2016).

method was still able to simulated the expected properties of the Brazilian interbank network.

Figure 2.4: Estimated interbank network for the Brazilian financial system.



Notes: The size of each financial institution represented in the interbank network reflects the sum of its interbank investments, interbank deposits and CET1 (Common Equity Tier I) in September 2022. This is the same weight used by the differential DebtRank algorithm when calculating the process of bank stress due to interbank contagion.

Because $r < 0$, the interbank market network is disassortative, indicating that the Brazilian financial system has highly connected FI's that are frequently connected to others with very few connections. This result follows the conclusion of Silva et al. (2016) and Alexandre et al. (2022), indicating the existence of money centers in which a few large banks have several connections with the market. This network topology makes the onset of a default in these money centers directly affect a large portion of the system. Thus, vulnerable neighbors to these money centers may, in turn, default, leading to a contagion process throughout the network. Evidence of negative assortativity in financial networks has also been reported in other empirical studies (Bottazzi et al., 2020).

Taking into account all the steps of our modeling process, the results of the LD described in Section 2.2.2.3 for the RS are shown in Tables 2.7 and 2.8. Table 2.7 shows how many of the 100.000 simulated scenarios have at least one idiosyncratic default (default scenarios) and how many have at least one contagion default (systemic risk scenarios). This analysis is separated into systemically important financial institution (SIFI) and non-systemically important financial institution (N-SIFI)[29]. This result shows the importance of the contagion process for our model, as it is present in 57% of the simulated scenarios with all banks.

Table 2.7: Summary statistics of the loss distribution of the banking system in the reduced sample.

	All banks	SIFI	N-SIFI
N. of default scenarios	98.820	24.717	74.103
N. of systemic risk scenarios	56.593	14.081	42.512

Notes: This table shows statistics of the loss distribution for 24 Brazilian banks in September 2022. SIFI stands for systemically important financial institution and N-SIFI for non-systemically important financial institution. The number of default scenarios represents the number of scenarios in which at least one bank idiosyncratically default. On the other hand, the number of systemic risk scenarios represents the number of scenarios in which at least one bank default due to the contagion process described in Section 2.2.2.2.

In detail, Table 2.8 shows the summary statistics of the LD of the Brazilian banking system in the RS. This exercise can be seen as the capacity of the Brazilian DIA to execute its additional resolution function (e.g., financial support) in the paybox plus the mandate. Panel A shows that, without contagion, the maximum cost of capital shortfall is close to BRL 131 billion and the one-year probability of Brazilian DIA default is close to 0.2% when considering only this operation. In the scenario where the DIA reimburses the covered deposits for N-SIFI in default under its paybox mandate, the maximum cost approaches BRL 72 billion, with the DIA's resources being sufficient to accommodate

[29]The SIFI is defined by the prudential segment 1 and the N-SIFI are defined by the prudential segment 2 up to 5. In accordance with Resolution CMN n.º 4.553/2017, the Central Bank of Brazil segments financial institutions and other licensed entities into five categories: S1 includes universal banks, commercial banks, investment banks, foreign exchange banks, and federal savings banks with a size equal to or exceeding 10% of the GDP or those engaging in relevant international activity; S2 encompasses the same bank categories as S1 but with a size less than 10% and equal to or exceeding 1% of the GDP, and other institutions with a size equal to or exceeding 1% of the GDP; S3 is for institutions with a size below 1% but equal to or exceeding 0.1% of the GDP; S4 consists of institutions with a size less than 0.1% of the GDP; and S5 includes institutions below 0.1% of the GDP that utilize a simplified optional methodology for determining minimum equity requirements, excluding the banks listed in S1 and S2. For more information, see BCB (2017).

the reimbursements of every simulation. Lastly, when considering the cost of all eligible deposits, the maximum cost is close to BRL 1.46 trillion and the PD is close to 50.6%.

On the other hand, Panel B shows that, with contagion, the maximum cost of capital shortfall is close to BRL 310 billion and the PD of the DIA is close to 11%. Note that even the contagion process increases banks' PD by 27% on average, as reported in Table 2.5, it increases 56 times the DIA's PD. In scenarios considering the reimbursement of the covered deposits of the N-SIFI, the maximum cost is approximately BRL 78 billion, with the DIA maintaining sufficient resources for all simulated scenarios. Finally, when considering all eligible deposits, the maximum cost is close to BRL 2.2 trillion and the PD of the DIA is close to 56.3%.

Table 2.8: Summary statistics of the loss distribution of the banking system with and without contagion in the reduced sample.

	Cost of capital shortfall			Cost of eligible deposits			Cost of covered deposits		
	All banks	SIFI	N-SIFI	All banks	SIFI	N-SIFI	All banks	SIFI	N-SIFI
Panel A: without contagion process									
Mean	12,14	11,82	0,32	199,82	149,62	50,20	104,58	78,31	26,27
Percentiles:									
0%	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
25%	0,00	0,00	0,00	67,29	23,54	12,20	35,22	12,32	6,38
50%	0,00	0,00	0,00	113,50	55,10	67,29	59,40	28,84	35,22
75%	42,94	41,69	0,00	415,52	344,02	71,50	217,48	180,06	37,42
90%	42,94	42,56	1,25	501,35	411,68	89,66	262,41	215,48	46,93
95%	44,20	42,94	1,64	591,35	497,11	90,79	309,51	260,19	47,52
99%	45,98	42,94	3,03	615,10	514,73	101,19	321,94	269,41	52,96
99,9%	127,50	123,22	4,50	1.324,12	1.209,54	116,49	693,04	633,07	60,97
100,0%	131,16	124,40	7,07	1.459,99	1.325,37	137,56	764,16	693,70	72,00
N. of LOLR scenarios	0.194	0.194	0.00	50.556	41.016	0.721	29.831	26.706	0.00
Panel B: with contagion process									
Mean	31,74	31,32	0,43	306,79	239,26	67,54	160,58	125,23	35,35
Percentiles:									
0%	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
25%	0,00	0,00	0,00	68,46	22,37	67,29	35,83	11,71	35,22
50%	0,00	0,00	0,00	180,79	107,42	68,41	94,63	56,23	35,81
75%	42,94	41,69	0,39	432,93	343,27	89,66	226,60	179,67	46,93
90%	220,94	218,30	1,44	1.270,74	1.179,70	91,04	665,10	617,45	47,65
95%	220,94	219,31	2,40	1.337,24	1.237,74	99,44	699,91	647,84	52,05
99%	223,35	219,93	3,65	1.362,15	1.251,13	111,86	712,95	654,84	58,55
99,9%	305,89	301,22	5,12	2.175,13	2.051,32	123,30	1.138,47	1.073,66	64,54
100,0%	309,79	302,72	7,07	2.220,18	2.071,09	149,09	1.162,04	1.084,01	78,03
N. of LOLR scenarios	10.901	10.901	0.00	56.291	50.871	2.107	36.091	26.853	0.00

Notes: This table shows statistics of the loss distribution for 24 Brazilian banks in September 2022. SIFI stands for systemically important financial institution and N-SIFI for non-systemically important financial institution. Cost of capital shortfall uses the SRISK as the EAD in the process described in Section 2.2.2.3, while cost of eligible deposits uses the total deposits and cost of covered deposits uses a share θ over the total deposits. Because the exact covered deposits are not publicly available for each bank, but only the total covered and eligible deposits of the system, we used the share $\theta = 52.34\%$ over the total deposits as a proxy. Panel A shows summary statistics of the banking system LD with only idiosyncratic default (without contagion process). Panel B presents the same statistics when all the contagion process described in Section 2.2.2.2 are considered. The number of LOLR scenarios represents the number of scenarios in which the consumption of resources due to default exceeded the equity of the Brazilian deposit insurance in all simulations. All values are in BRL billion except for the number of LOLR scenarios.

2.4.2 Banks' Probability of Default and Contagion Process in the Full Sample

This section details the exercise of calculating the economy's expected loss using the full sample, and compares these findings with those from the reduced sample, as explored in Section 2.4.1. While the RS considers all 24 listed banks in September 2022, a necessary condition for calculating the main systemic risk measures in the literature, but with the downside of representing only 13.19% of the total member institutions in Brazil, this FS considers all 182 member institutions in its analysis. Since we cannot calculate measures such as SRISK and others to discuss the capital shortfall in time of crisis, we focus on comparing the results of LD in terms of eligible and covered deposits.

Table 2.9 presents summary statistics for the risk measures estimated to build the LD of the Brazilian banking system with a FS of all member institutions. Compared to Table 2.5, we can observe, as expected, that our model maintains the average default of the structural model and preserves the increase of 27% in the average PD, but with a lower standard deviation. Furthermore, the PD without contagion for this FS, at 13.3%, closely aligns with the 12.4% in the RS, underscoring the latter's capacity to capture key structural elements of the PD across the broader banking system.

Table 2.9: Summary statistics of structural PD and banks' PD with and without contagion in the full sample.

		Structural PD of Merton model (%)	Banks' probability of default (%)	
			Panel A: without contagion process	Panel B: with contagion process
Mean		13,34	13,34	16,97
Std. Deviation		17,17	2,19	2,41
Minimum		0,00	5,00	6,11
Maximum		83,37	23,33	27,78
Percentiles:	25%	0,01	11,67	15,56
	50%	4,24	13,33	17,22
	75%	25,14	15,00	18,33

Notes: This table shows statistics of PD estimates for 182 covered banks in September 2022. Structural PD represents the estimates of the Merton (1974)'s model used as input for the idiosyncratic PD of each bank. Banks' PD are the default rates of the sample banks in the Monte Carlo simulation. In detail, Panel A shows summary statistics of Banks' PD with only idiosyncratic default (without contagion process). Panel B presents the same statistics when considering all the contagion process described in Section 2.2.2.2.

Regarding the estimation of the interbank network, the FS database produces a more adherent result with the moments of the true Brazilian interbank network. While in the RS the statistics for the density, assortativity and average degree are 0.31, -0.54 and 14.4, respectively, the statistics in the FS are 0.02, -0.58 and 4.8, respectively. Again, in terms of comparison, the density, assortativity and average degree reported by BCB and other works using the true Brazilian interbank network vary between [0.03,0.07], [-0.31,-0.54] and [4.6,7.8] (or [21,26] for large banks), respectively, during March 2010 and September 2015. Thus, in the FS we are able to better capture the level of density and average degree when we compare to the true Brazilian interbank network.

Taking into account the results of the LD described in Section 2.2.2.3 for the FS, Table 2.10 shows how many of the 100.000 simulated scenarios have at least one idiosyncratic default (default scenarios) and how many have at least one contagion default (systemic risk scenarios). Compared with Table 2.7, which focuses on listed banks, our findings indicate that the dynamics of default and contagion in the FS are more pronounced across the banking system, particularly for N-SIFI. This aligns with the principles of systemic importance and the extensive interconnectedness within the banking sector.

Table 2.10: Summary statistics of the loss distribution of the banking system in the reduced sample.

	All banks	SIFI	N-SIFI
N. of default scenarios	100.000	3.296	96.704
N. of systemic risk scenarios	99.985	3.295	96.690

Notes: This table shows statistics of the loss distribution for 182 Brazilian banks in September 2022. SIFI stands for systemically important financial institution and N-SIFI for non-systemically important financial institution. The number of default scenarios represents the number of scenarios in which at least one bank idiosyncratically default. On the other hand, the number of systemic risk scenarios represents the number of scenarios in which at least one bank default due to the contagion process described in Section 2.2.2.2.

Lastly, considering that the RS represents 13.19% of the total member institutions but 89.71% of the total deposits, we analyze whether the LD constructed using the RS maintains the proportions of total deposits with respect to the FS. Comparing the results of the LD for eligible and covered deposits in the FS in Table 2.11 with the results of the RS in Table 2.8, we can verify that the RS was able to capture only the tails of the LD (above the 99th quantile), especially for the SIFI, but was unable to correctly capture the beginning and middle of the distribution. This is because the SIFI's are common in both databases, but the RS considers only a smaller portion of the N-SIFI's.

Note that this exercise in Table 2.11, different from the RS in Table 2.8, which tests the capacity of additional resolution functions of the DIA, can be seen as the capacity of the DIA to execute its paybox mandate, that is, to reimburse all covered deposits in the system. In that sense, Panel A shows that, without contagion, the maximum cost for the DIA to reimburse the covered deposits of the N-SIFI in default is BRL 156 billion and its associate PD is close to 1.1%. On the other hand, Panel B shows that, with contagion, the maximum cost of the N-SIFI is close to BRL 168 billion, where the DIA would default in 2.9% of all simulated scenarios.

Table 2.11: Summary statistics of the loss distribution of the banking system with and without contagion in the full sample.

	Cost of elegible deposits			Cost of covered deposits		
	All banks	SIFI	N-SIFI	All banks	SIFI	N-SIFI
Panel A: without contagion process						
Mean	257,91	149,92	107,99	134,99	78,47	56,52
Percentiles:						
0%	4,54	0,00	4,54	2,37	0,00	2,37
25%	111,73	35,53	76,21	58,48	18,59	39,89
50%	170,67	59,08	111,58	89,33	30,92	58,40
75%	461,07	325,08	135,99	241,32	170,15	71,18
90%	570,47	412,95	157,52	298,58	216,14	82,45
95%	633,98	461,95	172,03	331,83	241,79	90,04
99%	702,01	500,10	201,91	367,43	261,75	105,68
99,9%	1.378,93	1.145,08	233,85	721,73	599,34	122,39
100,0%	1.579,64	1.282,31	297,33	826,79	671,16	155,62
N. of LOLR scenarios	78.080	38.090	56.067	43.777	26.706	1.069
Panel B: with contagion process						
Mean	868,01	745,49	122,52	454,32	390,19	64,12
Percentiles:						
0%	12,91	0,32	11,60	6,76	0,17	6,07
25%	820,08	728,88	91,20	429,23	381,49	47,73
50%	920,85	795,79	125,06	481,97	416,51	65,46
75%	1.030,12	878,58	151,54	539,16	459,85	79,31
90%	1.334,48	1.159,94	174,54	698,47	607,11	91,35
95%	1.394,46	1.204,82	189,63	729,86	630,60	99,25
99%	1.457,43	1.236,48	220,95	762,82	647,17	115,65
99,9%	2.131,00	1.875,45	255,55	1.115,37	981,61	133,76
100,0%	2.318,86	1.997,66	321,20	1.213,69	1.045,58	168,12
N. of LOLR scenarios	93.036	84.587	66.352	84.227	83.987	2.893

Notes: This table shows statistics of the loss distribution for 182 Brazilian banks in September 2022. SIFI stands for systemically important financial institution and N-SIFI for non-systemically important financial institution. Cost of eligible deposits uses the total deposits and cost of covered deposits uses a share θ over the total deposits. Because the exact covered deposits are not publicly available for each bank, but only the total covered and eligible deposits of the system, we used the share $\theta = 52.34\%$ over the total deposits as a proxy. Panel A shows summary statistics of the banking system LD with only idiosyncratic default (without contagion process). Panel B presents the same statistics when all the contagion process described in Section 2.2.2.2 are considered. The number of LOLR scenarios represents the number of scenarios in which the consumption of resources due to default exceeded the equity of the Brazilian deposit insurance in all simulations. All values are in BRL billion except for the number of LOLR scenarios.

2.4.3 Optimized Capital Requirement

In order to simulate an optimized capital requirement for the Brazilian financial system, we used a similar theoretical framework in which portfolio risk is calculated in banking organizations. We also adopted a heterogeneous CR regime in which we have different CRs for each bank depending on the prudential segment as in Alexandre et al. (2022). Banks can also hold different levels of capital adequacy ratio based on their own strategy and balance structure. Through this analysis, we used granular and public balance sheet information to estimate the PD of each FI as presented in Section 2.2.2.1, the amount of capital expected in times of crisis (SRISK) presented in Section 2.2.1.3, and the systemic risk contagion process presented in Section 2.2.2, which models the banking segmentation cluster to capture bank runs due to panic and market similarity, the interbank network to estimate the financial contagion in the banking system, and the net stable funding ratio to capture long-run liquidity shortages.

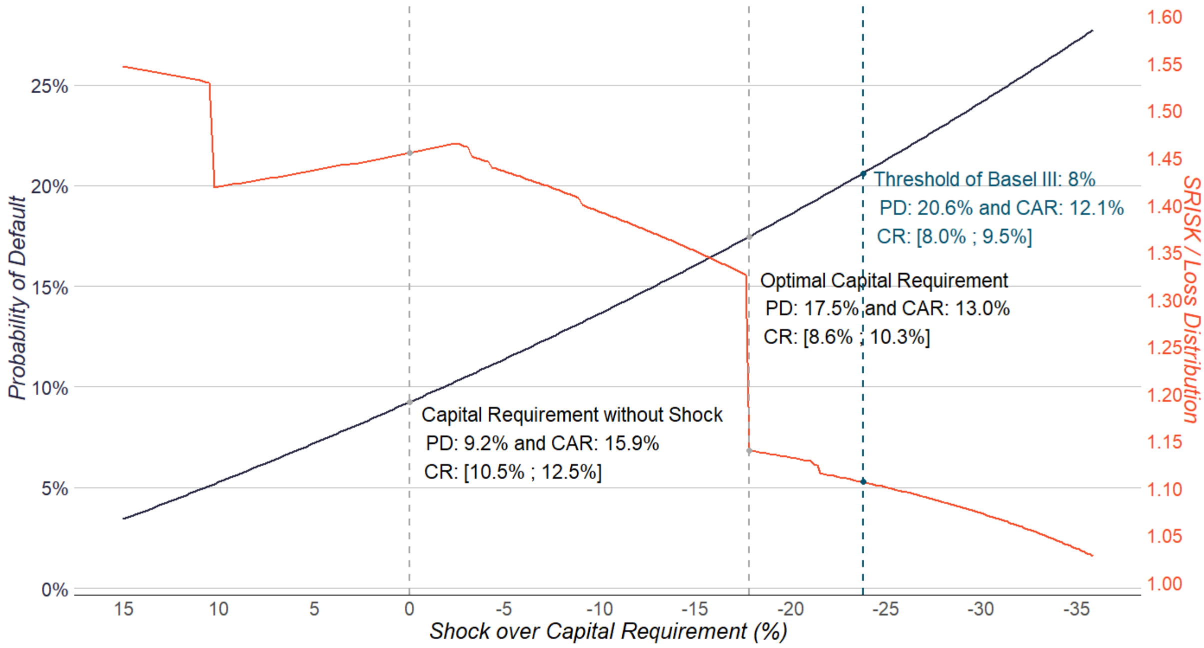
Thus, we estimated the minimum CR that maintains stable the relationship between SRISK and the economy's loss distribution and that takes into account the international framework of Basel III. The fundamentals of this approach over the LD consider that (i) a lower CR decreases the CAR hold by banks and the amount of SRISK necessary in the economy, which also decreases the total cost of bailout by the central bank and the LD, and (ii) increased the bank's probability of default, which increases, on the other hand, the frequency and costs of extrajudicial settlements, and the LD. Therefore, because there are effects of costs and benefits acting in opposite directions through PD and SRISK, respectively, the maintenance of the relation between SRISK and LD through several shocks characterizes an optimized trade-off scenario. To this purpose, we calculated the PD of the whole system as a weight average of all banks' PD in relation to its total deposits.

Because the CAR of each bank is the result of total regulatory capital over the risk-weighted assets, we calculated a two-way fixed effects model to estimate the impact that the probability of default has over these two variables to reconstruct the individual and aggregate capital adequacy ratio at every new shock on the bank's PD. We also controlled this effect by considering loan operations by risk level, since it is expected that banks may try to meet a higher capital requirement by either reducing assets, which decreases loan supply, or increasing the loan interest rate, which leads to a reduction in loan demand (Thakor, 2014; Alexandre et al., 2022). It is important to mention that we also control this reduction in loan demand on its reflection on total deposits. Thus, we are able to construct a link between our interest variables and also check the individual consistency of the simulated CAR with the minimum required by Basel III. The results of our equation estimate that, on average, 1% increase in the bank's PD reduces total regulatory capital by 0.10%, increases RWA by 0.12%, and increases loan operations by 0.35%, which also increases the total deposits by 0.13%[30].

Through Figure 2.5 it is possible to observe a more accelerated drop in SRISK in relation to the LD of the economy as we decrease the CR and CAR of banks. Thus, the moment when a next marginal drop maintains this relationship relatively stable for the next simulations and it is in accordance with Basel III framework is characterized as the optimized CR and CAR. Additional decreases from this point in CR would not bring long-term benefits in reducing the LD and would continue to increase the burden of a greater PD for the financial system. Note that this optimization problem is subject to binding regulatory constraints such that an indefinite reduction in CR would be incompatible with international accords.

[30]The complete results are shown in Table B.2 and Table B.3 in Appendix B. It is important to mention that, although we are simulating the optimized capital requirement for 26 banks that has available market data, we estimated the average effect considering all 244 Brazilian financial institutions between December 2000 and September 2022, which results in 9.653 observations.

Figure 2.5: Impact of the capital requirement shock on SRISK/LD and probability of default.



Notes: CAR stands for capital adequacy ratio and *CR* for minimal capital requirement, which varies depending on the bank prudential segment. The blue dashed line represents the threshold in which the capital requirement of at least one bank would be bellow the minimum required of 8% by Basel III.

Our results show that the optimized capital adequacy ratio for the Brazilian financial system in September 2022 is close to 13%. Also, the optimized interval for the minimum capital requirement depending on the prudential segment varies between 8.6% and 10.3%. Considering that the CAR calculated in our model is 15.9% and the reported by the BCB (2023b) is 16.15%[31], this reduction of 18% in our model implies an optimized CAR in the entire financial system of 13.2%. Because a lower CAR is expected to lead to a financial system more fragile and susceptible to systemic risk, the probability of default of the entire system calculated by our model increases from 9.2% to 17.5%, a variation of 89%.

Taking into account that the percentage of extrajudicial settlements or interventions made by the BCB on the Brazilian banking market is 8.2%, this positive variation of 89% in the PD implies a probability of default of the entire financial system of 15.5%. Considering that the PD of the US financial system is 11.9% and the lowest historical

[31]This indicator measures the capital adequacy of financial institutions in the Brazilian financial system and is based on the definitions used in the Basel Capital Accord. The scope of the data coverage commercial banks, universal banks, investment banks, savings banks or any financial conglomerate comprising any of these entities. For more information, see BCB (2023b) series 21424 (I001 - Regulatory Capital to Risk-Weighted Assets).

level of CAR in Brazil since 2000 was 12.9% in May 2001, the level of 15.5% for the Brazilian PD as a consequence of a CAR of 13.2% can be argued as reasonable in view of the benefits of a financial system with lower cost for the central bank in a financial crisis scenario, greater loan supply and lower credit cost[32].

In terms of impact of this new CAR, because the simulated capital requirement is lower, the total amount of capital needed in an extreme financial crisis event would also be lower. For our data set, which covers 79.89% of total assets but only 13.19% of total member institutions, we estimate a total cost of bailout in the base scenario (without shock) between 3.9 and 5.7 times the size of the Brazilian deposit insurance agency. Considering the scenario with a 18% shock in the CAR, the total cost of bailout would be between 3.4 and 3.9[33]. Therefore, one of the estimated benefits of this new CAR and CR would be a reduction between 13% and 32% in the cost of bailout. In addition, another positive estimated impact would be an increase in loan operations by 31.1% and in total deposits by 11.6%.

It is important to note that the optimized CAR of the system does not necessarily mean the optimized minimum CR established by the regulator. Banks can hold levels of regulatory capital ratio higher than the minimum demanded by the Central Bank for strategic reasons or for expectations about its own portfolio or future macroeconomic conditions. Thus, considering that in our model we allow a heterogeneous CR regime, while the official minimum regulatory capital ratio in September 2022 varies between 10.5% and 12.5% depending on the bank prudential segments, the minimum regulatory capital ratio with a 18% shock would result in a range between 8.6% and 10.3%, which is compatible with the 8% minimum established by Basel III (BCBS, 2011).

[32]It is worth to mention that our modeling process share similar premises presented by Alexandre et al. (2022), but our results do not generate unrealistic minimum CAR for some banks (while the authors mention results close to 1%, our minimum is close to 9.1%). Furthermore, we also consider several balance sheet variables of each bank (used to construct our probability of default, systemic risk metrics, clusters, interbank network, and NSFR) in our model instead of calibrating or assuming baseless values for these parameters.

[33]The lower bound of the range is composed by the maximum value of the loss distribution and the upper bound is composed by the sum of SRISK in the specific scenario. Note that the reduction in the gap between the lower and upper bound after the shock is a consequence of the increase in PD in the simulation.

2.5 Final Remarks

This paper estimates different measures to understand how much systemic risk each bank brings to the Brazilian market and proposed a bank run model that accounts for idiosyncratic probability of default of banks and a systemic risk process in which additional defaults occur through different channels of contagion.

Our approach follows a similar theoretical framework in which portfolio risk is calculated in banking organizations and in which banking losses are estimated in deposit insurance schemes. Through this analysis, we used granular and public balance sheet information to estimate the PD of each FI and the systemic risk contagion process captured through the channels of (i) panic due to deposit withdrawals and market similarity, (ii) interbank network, and (iii) funding illiquidity.

Through the application of our model to a reduced sample of 24 listed banks and a full sample of 182 covered banks for September 2022, we estimate the loss distribution, the PD of the deposit insurance agency, and the optimized capital adequacy ratio of the Brazilian banking system. We find that the DI would be able to bailout 80% the system without contagion in 99.8% of the simulated scenarios and 89.1% when considering the contagion process in the reduced sample. However, if banks were liquidated and the DI needed to reimburse covered deposits in worst-case scenarios, the PD of the DI with contagion in the reduced sample for all banks would be 36.1% and, considering only the N-SIFI, the DI would be able to reimburse all covered deposits. In the full sample, considering also the reimbursement of all covered deposits, the DI's PD would be 84.3% for all banks and 2.9% for the N-SIFI.

Regarding the optimized capital adequacy ratio, our results show a value close to 13.2% with a minimum capital requirement interval varying between 8.6% and 10.3% depending on the prudential segment. This is 18% lower than the practiced in the Brazilian financial market, but compatible with the 8% minimum established by Basel III. This reduction would increase loan operations by 31.1% and total deposits by 11.6%, but would also increase the PD of the banking system from 8.2% to 15.5%.

In terms of application, the results of this paper could also be used to determine the optimal fund size for a DIA to effectively address widespread bank failure. Furthermore, it is also possible to customize the framework to incorporate country-specific factors, capturing the distinct realities and challenges of each economy. This could be done by adjusting each trigger and process of the contagion channels. Finally, this paper also shows the need to consider different channels of contagion and a broader coverage of the banking system as key elements when designing the overall financial safety net.

Chapter 3

Bootstrap Estimator Approach to Financial Stability

3.1 Introduction

After more than a decade since the recent Global Financial Crisis, despite considerable efforts made to understand the sources of financial stability and systemic risk, a significant gap still remains in the field (Christiano et al., 2018; De Bandt and Hartmann, 2019). This is specially the case in developing economies, since these regions often feature a limited number of publicly traded banks and restricted information, generating challenges in analyzing the probability of default across all financial institutions.

As explored in Chapters 1 and 2, taking the Brazilian scenario as a case study, which offers insights applicable to other economies, there is a lack of research in understanding the determinants of PD in the banking system, encompassing both listed and non-listed banks (Souza et al., 2015, 2016; Guerra et al., 2016). Considering the composition of the banking system as of September 2022, where only 24 (13.2%) of the member FIs covered by the private deposit insurance agency, *Fundo Garantidor de Créditos*, are publicly listed, the inclusion of non-listed banks to address the whole system becomes particularly relevant.

To address this gap, the present study introduces the Bootstrap Estimator with Variable Selection (BEVS), a method that allows for an integrated and broad analysis of the impact of macroeconomic variables on PD. Through this approach, we aim to provide deeper insight into the determinants of PD that could inform the development of strategies and policies designed to improve financial stability. In relation to the broader literature, our work is closely related to the strand that integrates different techniques to enhance accuracy in time series analysis (Petropoulos et al., 2018; Li et al., 2023; Wang et al., 2023).

In our proposed method, we combine techniques such as Lasso regression, Loess smoothing, and bootstrap aggregation (bagging), showing that this integrated approach yields improved results compared to those obtained through their individual performance. In particular, we demonstrate that BEVS outperforms the single Lasso model set as a benchmark. Our findings indicate that BEVS not only refines the estimate of PD but also offers a nuanced view of the impact of macroeconomic factors over the study period, such as a distribution of coefficients and a measure of variable significance through the number of appearances.

This work is structured into five sections, beginning with this Introduction. Section 3.2 details the theoretical framework with the foundations of BEVS that is employed to estimate the determinants of the probability of default in the Brazilian banking system. Section 3.3 describes the data used in our models. Section 3.4 presents the results and a discussion of our findings, and Section 3.5 concludes with the final remarks of this paper.

3.2 Theoretical Framework

To understand the determinants of the probability of default in the Brazilian financial system as a case study, in this section, we introduce our Bootstrap Estimator with Variable Selection. In addition, we also present the construction of the PD and delve into key concepts of methods such as Lasso, Loess, and bagging.

3.2.1 Probability of Default

Following Chapters 1 and 2 of this thesis, we utilize the structural model of Merton (1974) to estimate the idiosyncratic probability of default (IPD) for each financial institution. Using these IPDs, we then compute the aggregate PD for the Brazilian banking system, weighted by the deposits of individual FIs[1] (Souza et al., 2015, 2016; Guerra et al., 2016; Coccorese and Santucci, 2019; da Rosa München, 2022).

[1]By weighting the IPD based on individual FI's deposits, we gain insights into systemic risk and an institution's influence within the Brazilian banking sector. Deposits are core liabilities for FIs and delineate the potential burden on the Brazilian deposit insurance in the event of default.

Recalling previous definitions and using the Black and Scholes (1973)'s model, the option's payoff for the equity holder at time T is given by 3.1.

$$E_{it} = max(A_{it}\, \mathscr{N}(d_{1it}) - DB_{it}\, e^{-r_t T} \mathscr{N}(d_{2it}), 0) \tag{3.1}$$

Where A_{it} is the asset value, r_t is the risk-free interest rate, $\mathscr{N}(.)$ is the cumulative normal distribution function,

$$d_{1it} = \frac{ln(\frac{A_{it}}{DB_{it}}) + (r_t + \frac{\sigma^2_{Ait}}{2})T}{\sigma_{Ait}\sqrt{T}} \quad \text{and}$$

$$d_{2it} = d_{1it} - \sigma_{Ait}\sqrt{T} = \frac{ln(\frac{A_{it}}{DB_{it}}) + (r_t - \frac{\sigma^2_{Ait}}{2})T}{\sigma_{Ait}\sqrt{T}},$$

in which σ_{Ait} denotes the volatility of the assets.

Thus, the IPD_{it} of a FI in a time horizon T, calculated in $t = 0$, is given by 3.2.

$$\begin{aligned} IPD_{it} &= P(DB_{it} \geq A_{it}) \\ &= P(\ln DB_{it} \geq \ln A_{it}) \\ &= \mathscr{N}(-d_{2it}) \\ &= \mathscr{N}\left[- \frac{ln(\frac{A_{it}}{DB_{it}}) + (r_t - \frac{\sigma^2_{Ait}}{2})T}{\sigma_{Ait}\sqrt{T}} \right] \end{aligned} \tag{3.2}$$

Note that the probability of default is the area under the default barrier, that is, a fraction of total liabilities. Finally, to calculate our aggregate PD for the entire Brazilian financial system, we utilize the following expression given by 3.3.

$$PD_t = \frac{\sum_{i=1}^{N} IPD_{it} \times \text{Deposits}_{it}}{\sum_{i=1}^{N} \text{Deposits}_{it}}, \quad \forall t \in \{1, \ldots, T\} \tag{3.3}$$

3.2.2 Least Absolute Shrinkage and Selection Operator

The Least Absolute Shrinkage and Selection Operator (Lasso) is a linear regression method proposed by Tibshirani (1996, 2011) that performs both variable selection and regularization, thereby improving the prediction accuracy and interpretability of traditional linear models. It shrinks some of the coefficients and sets others to zero, combining the advantages of subset selection and ridge regression.

Consider the data $(\mathbf{x}^i, y_i)$, $i = 1, 2, \cdots N$, where $\mathbf{x}^i := (x_1, \cdots, x_p)^T$ are the predictor variables and y_i are the responses. Letting $\hat{\boldsymbol{\beta}} := (\hat{\beta}_1, \cdots, \hat{\beta}_p)^T$, the objective of the Lasso is to solve 3.4.

$$\underset{\beta_0, \boldsymbol{\beta}}{\arg\min} \left\{ \sum_{i=1}^{N} \Big(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \Big)^2 \right\} \qquad s.t. \quad \sum_{j=1}^{p} |\beta_j| \leq t \tag{3.4}$$

Here, $t \geq 0$ s a tuning parameter that determines the amount of regularisation, that is, the amount of shrinkage that is applied to the estimates. We can also write the Lasso problem in the equivalent Lagrangian form given by 3.5.

$$\hat{\beta}^{lasso} = \underset{\beta_0, \boldsymbol{\beta}}{\arg\min} \left\{ \frac{1}{2} \sum_{i=1}^{N} \Big(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \Big)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \tag{3.5}$$

Note that Lasso regression employs an l_1 penalty, denoted as $\|\boldsymbol{\beta}\|_1 = \sum |\beta_j|$, which enforces certain coefficient estimates to be precisely zero when the tuning parameter λ is sufficiently large. Hence, similar to the best subset selection, Lasso also performs variable selection. Consequently, models derived from Lasso tend to be more interpretable (sparse models) compared to those generated by Ridge regression.

Carefully selecting the regularization parameter, denoted as λ, is an important aspect of utilizing Lasso regression. Making an informed choice for this parameter is essential for optimizing the model's performance in terms of prediction accuracy and model interpretability, as it controls the strength of shrinkage and variable selection. However, if regularization becomes too strong, important variables may be left out of the model, and coefficients may shrink excessively, which can reduce both predictive power and inference. Considering this, information criteria such as the Bayesian Information Criterion and the Akaike Information Criterion may be favored for cross-validation, since they offer the advantage of faster computation while exhibiting greater stability in small sample sizes. An information criterion selects the estimator's regularization parameter by optimizing a model's in-sample accuracy while penalizing its effective number of parameters or degrees of freedom.

3.2.3 Locally Weighted Regression

Locally weighted regression (Loess), also known as Locally Estimated Scatterplot Smoothing, is a non-parametric method that enables the fitting of multiple regressions within local neighborhoods of a dataset. Introduced by Cleveland (1979) and further developed by Cleveland and Devlin (1988), this technique combines the simplicity of linear least squares regression with the flexibility of non-linear regression.

The Loess method constructs a function that interprets the deterministic portions of data variability, analyzing point by point by fitting simple models to localized data subsets. Consequently, there is no requirement to specify a global function to fit a model to the data. Instead, it focuses on representing individual data segments, promoting a granular understanding of data distributions.

In this approach, at each point in the dataset, a low-degree polynomial is fitted to a subset of the data, using explanatory variable values near the point for which the response is estimated. The fitting process employs weighted least squares, assigning higher weights to nearby data points and lower weights to those farther away. The value of the regression function for each data point is determined by evaluating the local polynomial using the values of the specific explanatory variables associated with that data point. The Loess fitting process concludes once the values of the regression function have been computed for each of the n data points.

For the selection of data subsets in weighted least squares fits, a nearest-neighbors algorithm is employed. The 'bandwidth' or 'smoothing parameter', denoted as α, is a user-defined input that controls the amount of data used in each local polynomial fit. Specifically, α represents the fraction of the total n data points used in each local fit. These data points are selected on the basis of their explanatory variable values, with a preference for those closest to the point for which the response is estimated. Since a polynomial of degree k requires at least $(k+1)$ points for a fit, the smoothing parameter α must be between $(\lambda+1)/n$ and 1, with λ denoting the degree of the local polynomial.

In practice, irregularly spaced local regressions are common when using a fixed span h. This results in some local estimates (e.g., x_0) being based on many points, while others rely on only a few points. For this reason, it is beneficial to employ a nearest-neighbor strategy to determine the span for each target of local regressions. To achieve this, we calculate $\Delta_i(x_0) = |x_0 - x_i|$ based on the smoothing parameter α and define the span as

$h(x_0) = \Delta_{(n\times\alpha)}(x_0)$. In this context, a span equal to 0.75 of α, for example, implies that for each local fit, our goal is to utilize 75% of the data defined by α.

The variable α is known as the smoothing parameter because it controls the flexibility of the Loess regression function. Larger values of α result in a smoother function that is less sensitive to data fluctuations. As α decreases, the regression function becomes increasingly aligned with the data. However, using an excessively small value for the smoothing parameter is not advisable, as it can lead the regression function to capture random errors in the data.

The local polynomials fit to each subset of the data are typically either of first or second degree, meaning they are either locally linear or locally quadratic. Using a zero-degree polynomial transforms Loess into a weighted moving average. Although it is theoretically possible to employ higher degree polynomials, doing so would result in models that deviate from the core principles of Loess. Loess operates on the premise that any function within a small neighborhood can be adequately approximated using a low-order polynomial. This preference for simplicity aligns with the ease of fitting the data, as high-degree polynomials tend to overfit and introduce numerical instability.

The weight function assigns the highest weight to the data points closest to the point of estimation and the lowest weight to those farthest away. This weighting scheme is rooted in the concept that points in close proximity within the explanatory variable space are more likely to exhibit a simple relationship than those that are distant. Consequently, data points closely aligned with the local model exert a more substantial influence on model parameter estimates, while those less likely to conform to the local model have a diminished impact on these estimates. In this context, Loess traditionally uses the tri-cube weight function, defined as 3.6.

$$W(x) = \begin{cases} (1-|d|^3)^3, & \text{for } |d| < 1, \\ 0, & \text{for } |d| \geq 1 \end{cases} \tag{3.6}$$

where d represents the distance of a given data point from the point on the curve being fitted, scaled to fall within the range of 0 to 1. However, any other function that meets the criteria listed in Cleveland (1979) can also be employed[2]. The weight assigned to a

[2]Let W be a weight function with the following properties:

1. $W(x) > 0$ for $|x| < 1$;

particular point within a localized data subset is determined by evaluating the distance weight function in such a way that the maximum absolute distance among all points in the data subset is normalized to exactly one.

3.2.4 Bagging

Bagging, also called bootstrap aggregation, is a machine learning ensemble algorithm proposed by Breiman (1996) designed to improve the stability and accuracy of regression and classification algorithms. In general, this method is used for fitting multiple versions of a prediction model and then combining (or ensembling) them into an aggregated prediction. In other words, bagging is an algorithm in which b bootstrap copies of the original training data are created and new predictions are made by averaging the predictions of the individual base learners.

Recall that for a set of n independent observations $Z_1, \cdots, Z_n$, each with a variance of σ^2, the variance of their mean $\bar{Z}$ is σ^2/n. This indicates that the variance is reduced when averaging a group of observations. Consequently, a straightforward method to decrease the variance and thereby enhance the prediction accuracy of a statistical learning approach is to create multiple training sets from the population, construct a separate predictive model for each set, and then average these predictions. In other words, by calculating $\hat{f}^1(x), \hat{f}^2(x), \cdots, \hat{f}^B(x)$ using B distinct training sets and then averaging these, we can obtain a single, low-variance statistical learning model, as indicated by Equation 3.7.

$$\hat{f}_{avg}(x) = \frac{1}{B}\sum_{b=1}^{B}\hat{f}^b(x) \tag{3.7}$$

However, this approach is not feasible in most cases, as we typically do not have access to multiple training sets. Instead, we can employ bootstrapping, which involves drawing repeated samples from the single available training dataset, doing so with replacement. In this method, we generate B distinct bootstrapped training datasets. For each of these, labeled as the bth set, we train our model to obtain $\hat{f}^{*b}(x)$. By averaging all these predictions, we arrive at a final model as described in Equation 3.8.

---

2. $W(-x) = W(x)$;

3. $W(x)$ is a non-increasing function for $x \geq 0$;

4. $W(x) = 0$ for $|x| \geq 1$.

$$\hat{f}_{bag}(x) = \frac{1}{B}\sum_{b=1}^{B}\hat{f}^{*b}(x) \tag{3.8}$$

This method is particularly effective with unstable, high-variance base learners, which are algorithms that show significant output variation in response to minor changes in the training data. However, for more stable algorithms or those with high bias, bagging tends to yield less improvement in predictions because of their inherent lower variability.

3.2.5 Bootstrap Estimator with Variable Selection

Building on the established frameworks of Lasso, Loess, and bagging techniques, we integrate these concepts to formulate the Bootstrap Estimator with Variable Selection, or BEVS. The contribution of BEVS lies in the combination of robust and established methods that could offer improved results compared to those obtained through their individual performance[3]. Thus, in this section, we outline the principal components and procedures of the BEVS approach, demonstrating its efficiency in analyzing the determinants of the probability of default in the Brazilian financial system.

There are notable works in the literature that also propose to combine different techniques to enhance accuracy in time series analysis (Petropoulos et al., 2018; Wang et al., 2023). For instance, the work of Bergmeir et al. (2016) presents a bagging approach that first transforms the data using Box-Cox, then decomposes it into trend, seasonal, and remainder components. They bootstrap the remainder with the Moving Block Bootstrap (MBB), reintegrate the series, and apply an inverse Box-Cox transformation. This process generates a random pool of similar bootstrapped time series, each fitted with an optimal exponential smoothing model selected via bias-corrected AIC, culminating in a median aggregation of forecasts.

Before exploring in detail the BEVS procedure, Algorithm 1 provides an overview of its algorithmic structure, which will guide the subsequent discussion.

[3]We demonstrate this by comparing the performance of Lasso in relation to BEVS.

Algorithm 1 Bootstrap Estimator with Variable Selection – BEVS

Data: Time series of the probability of default y, number of bootstrap iterations, bootstrap block size b, maximum number of parameters in Lasso, and size of the dimensionality reduction;

`// 1.` **`Pre-processing`**

Compute the smoothed series, y^s, by applying Loess smoothing on the original PD series y using cross-validation;

Construct deviations from the smoothed series, d^s, as: $d^s = y - y^s$;

Transform d^s for Circular Block Bootstrapping:

for $i > N$ **do**

  $X_i = X_{i \bmod N}$

$X_0 = X_N$

`// 2.` **`Bootstrap Process`**

for *`iteration = 1 to Number of bootstrap iterations`* **do**

  Generate Bootstrapped Subseries:

  for $i = 1$ **to** N **do**

    $\mathcal{B}_i = (X_i, \ldots, X_{i+b-1})$

  Combine y^s with random error blocks from $\{\mathcal{B}_1, \ldots, \mathcal{B}_N\}$ to obtain augmented series y^a;

  Apply Lasso regression on y^a incorporating bounded constraints and optimizing the penalty parameter λ using cross-validation to obtain y^l;

  Assess the model fit y^l using the deviance ratio and null deviance.

`// 3.` **`Post-Bootstrap Processing`**

From the individual Lasso models y^l, construct an ensemble model y^e by averaging coefficients with non-zero values over bootstrap iterations;

Examine the distribution of coefficients across all bootstrap samples to identify patterns or trends;

Initiate a dimensionality reduction process on y^e with an appearance threshold of 10% to filter significant features;

while *not all variables meet threshold* **do**

  Discard models with variables below threshold

  Adjust threshold

`// 4.` **`Residual Analysis`**

Compute the residuals, r, as: $r = y - y^e$;

Perform the following analyses on the residuals y^e to validate the model's performance:

- Goodness-of-Fit tests;
- Error metric calculations;
- Distribution tests;
- Autocorrelation tests.

Delving into the BEVS procedure, at the beginning we employ the Loess smoothing technique to delineate the underlying trends in the time series data, which in our case is the PD. This non-parametric method uses a time trend to recover the underlying dynamics in the series, capturing specifically the low-frequency variations of the data. In this step, the smoothing parameter is determined using the generalized cross-validation (GCV) criterion, which optimizes the bias-variance trade-off to minimize the predictive error on a validation set.

Following the determination of the optimal smoothing parameter via the Loess technique, the next step in the BEVS procedure is to construct a residual series for the bootstrap process. This is achieved by subtracting the smoothed data, derived through the Loess method, from the original dataset representing the PD. To adapt it for the Circular Block Bootstrap (CBB) approach (Politis and Romano, 1992), this error term vector, denoted as X_i, $i = 1, \ldots, N$, where N is the length of the error series, is transformed to form a circular series by appending a segment of its initial part to the end. Mathematically, for $i > N$, the series wraps around such that $X_i \equiv X_{i(\bmod N)}$, and, at the starting point, it holds that $X_0 \equiv X_N$. This definition ensures a continuous and seamless transition, forming a loop where the end reconnects to the beginning, maintaining the intrinsic structure and dependencies present in the original series.

After creating the circular series, we systematically generate a collection of potential subseries, each denoted by $\mathcal{B}_i = (X_i, \ldots, X_{i+b-1})$, where b represents the bootstrap block size holding a uniform number of observations[4]. This iterative process spans the entire length of the dataset, assembling a pool of subseries to construct new series based on the original data. This is achieved by augmenting the Loess smoothed series with error blocks randomly sampled with replacement from the set of potential subseries $\{\mathcal{B}_1, \ldots, \mathcal{B}_N\}$. Here, we apply the CBB concept, utilizing the circular nature of the error series to maintain the temporal dependencies and structures observed in the original data.

In each iteration, the Lasso regression is applied to the newly bootstrapped series to identify significant predictors, utilizing a penalty to induce sparsity in the parameter estimates. To embed theoretical reasoning into the regression, the lower and upper bounds for each independent variable are defined, guiding the estimation within plausible and

[4]In this study, the PD series comprises 60 observations, leading us to segment it into 7 bootstrap blocks for a balanced and efficient analysis.

theoretically grounded ranges[5]. The regression's tuning parameter, λ, is optimized through cross-validation to ensure optimal predictive performance. This process is repeated for a predetermined number of series[6], aiming to capture a robust representation of potential outcomes and maintain stability in the results. It is important to note that the new series have a lag adjustment, which involves incorporating lagged values of the smoothed series into the analysis.

In addition, we derive a set of goodness-of-fit (GOF) metrics in each bootstrap iteration to evaluate each model performance based on the optimal λ determined through cross-validation. The central element in this analytical process is the evaluation of the log-likelihood, derived from the deviance ratio and the null deviance of the dataset (Hastie et al., 2015). These informations are used to calculate key statistical criteria including the Akaike Information Criterion (AIC), corrected AIC (AICc), and Bayesian Information Criterion (BIC)[7]. These criteria incorporate the number of parameters (non-zero coefficients at the chosen λ value) and the number of observations, thereby providing a comprehensive view of the model fit.

After completing the iterative process, we advance to the next phase of the BEVS procedure, which involves aggregating all individual Lasso models created during the bootstrapping process into a unified bagged (ensemble) model. This strategy aims to retain only those coefficients that consistently appear with non-zero values across all the iterations, thereby accentuating the variables that significantly influence the dependent variable. Furthermore, we calculate the average coefficient value and analyze a range of percentiles to understand the distribution of each coefficient across the bootstrap samples,

[5]To guide the directional relationships between the dependent variable and each of the independent variables based on theoretical reasoning, we implement bounded constraints on the coefficients during the Lasso regression. When a negative relationship is expected, we assign a lower bound of negative infinity and an upper bound of zero to the coefficient estimates, restricting them to non-positive values. Conversely, for a expected positive relationship, the bounds are established at zero and positive infinity, ensuring only non-negative estimates. In instances where there is no prior theoretical directional expectation or where our objective is to empirically determine the sign of the relationship, we opt for a more unrestricted approach by setting the bounds to negative and positive infinity, allowing the analysis to freely estimate the optimal coefficient values.

[6]We found that 1,000 simulations is sufficient to maintain stability in our results. We also perform a robustness check varying the number of simulations to see its effects on the results, which can be seen in Table 3.4.

[7]Both AIC and BIC serve to assess the model's fit, each from a slightly different theoretical premises. AIC aims to balance goodness-of-fit with model complexity, penalizing models that have too many parameters to prevent overfitting. The adjusted version of AIC, called AICc, is more unbiased, making it advantageous when working with smaller sample sizes. Conversely, BIC favors parsimonious models, imposing stricter penalties to models with a large number of parameters.

enabling us to account for the potential pathways the series could follow. This analytical step includes counting the frequency of non-zero coefficients for each variable in the ensemble, thereby providing a quantitative measure of its significance in the model.

Following the aggregation process, the BEVS procedure initiates a dimensionality reduction phase to further optimize the model. During this iterative process, the prevalence of each variable across the ensemble of Lasso models is assessed, retaining only those variables that exceed a predefined threshold of appearance in the remaining models[8]. This process is conducted iteratively with each cycle discarding the variables that fall below the threshold and recalibrating the threshold based on the newly reduced model dimension. The procedure continues until all variables in the model satisfy the appearance threshold, resulting in a more condensed, yet effective set of predictors. This approach not only enhances the robustness and efficiency of the predictive framework but also fosters a model that is both parsimonious and retains substantial predictive power by concentrating on the most consistently influential variables.

In the final step of the BEVS approach, we employ a detailed analysis of the residuals derived from the difference of the original series, which in our case is the PD, and the bagged ensemble model. This important phase involves extensive analysis on both the complete and the reduced model to evaluate whether the dimensionality reduction process has created a parsimonious, yet effective model that retains reliable results for the residuals. We analyze the following robustness pillars[9]: (i) GOF tests, (ii) error metrics, (iii) distribution tests, and (iv) autocorrelation tests.

For the GOF tests, we consider several metrics including D^2, AIC, AICc, BIC, R^2, and average R^2[10]. The metric D^2 represents the fraction of deviance explained[11]. For

[8]The threshold was set at 10% to ensure that only the most consistently significant variables were retained.

[9]We do not test for cointegration between the modeled PD series and its derivative ensemble model because both series are inherently related by design, making the identification of a common stochastic trend more a reflection of the model's construction than an expected property. Cointegration typically implies a long-run equilibrium relationship between non-stationary series, but here, it merely underscores the ensemble's dependency on the PD. If the PD were observed rather than modeled, then testing for cointegration would be more relevant, as it would assess the long-term consistency of predictions with real-world data.

[10]In the context of our analysis, R^2 is the traditional coefficient of determination calculated using residuals from the original series versus the ensemble model. In contrast, the average R^2 represents the mean of the R^2 values computed for each of the 1,000 individual bootstrapped series, providing an aggregated insight into their collective performance.

[11]The name D^2 is by analogy with R^2, the fraction of variance explained in regression. Its expression is given by $D^2 = (Dev_{null} - Dev_{\lambda})/Dev_{null}$, where Dev_{λ} is defined as minus twice the difference in log-likelihood between a model fit with parameter λ and the fully parameterized model, while Dev_{null} is the null deviance computed for the constant model. For more information, see Hastie et al. (2015).

the error metrics, we calculate the Mean Absolute Scaled Error (MASE) and the Root Mean Square Error (RMSE), both of which offer distinct perspectives on the discrepancies between our predictions and the actual observations. For the distribution tests, we utilize the Kolmogorov-Smirnov (KS) test to access whether the residuals conform to a normal distribution. Lastly, for autocorrelation tests, we employ the Ljung-Box test up to the fourth lag to examine any potential autocorrelation in the residuals.

3.3 Data

To calculate the individual PD and construct the aggregate PD, we utilized quarterly data from December 2007 to September 2022 for 226 Brazilian financial institutions, yielding an unbalanced panel data with 7,556 observations. All balance sheet data employed in this study are publicly provided by the Central Bank of Brazil (BCB, 2023a).

Following the data structure presented in Chapters 1 and 2, the dataset considers financial conglomerates and independent institutions until December 2014, and the prudential conglomerates and independent institutions before March 2015 with the business model category of $b1$, $b2$, $b4$, and $n1$, provided there are at least six valid observations in the studied period. The final dataset represents 99.82% of total assets and 99.75% of total credit of covered member institutions in September 2022, with an average of 98.94% and 99.07% throughout the period, respectively. For the interest rate, we used public data provided by B3, the Brazilian financial market infrastructure company (B3, 2023).

To estimate the probability of default on a one-year horizon for each FI using the Merton (1974)'s structural model, we applied the following variables: adjusted total assets for A, total liabilities to calculate DB, annualized interbank interest rate DI for r, and the annualized standard deviation of the logarithmic returns of adjusted total assets, that is, $\log(A_t/A_{t-1})$, for asset volatility σ_A. Once all IPD are constructed, we build our PD according to equation 3.3 utilizing the deposits of individual FIs, yielding our final time series of 60 observations. Table 3.1 presents the aggregate descriptive statistics for these variables, Figure 3.1 presents the correlation matrix, and all balance sheet accounts and code variables are shown in Appendix C.3.

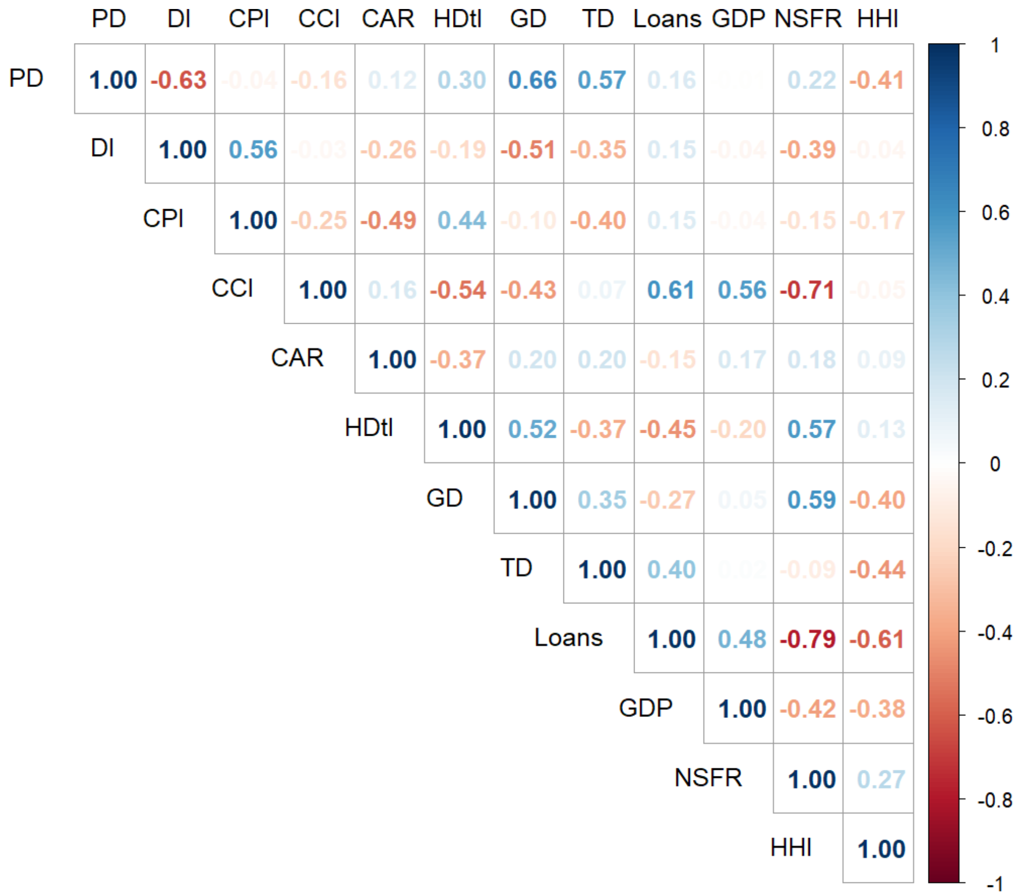
Table 3.1: Descriptive statistics of the dependent and independent variables.

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Dependent Variable						
ATA[a]	52.86	215.48	0.00	0.31	11.03	2,184.86
TL[a]	48.17	198.95	0.00	0.23	9.53	2,018.16
TD[a]	15.63	73.16	0.00	0.01	2.29	854.76
DI[b]	9.05	3.83	1.90	6.39	12.60	14.14
AV	0.39	0.36	0.00	0.14	0.55	3.73
IPD[b]	14.98	19.07	0.00	0.01	27.59	94.93
PD[b]	5.73	1.10	3.91	4.89	6.30	8.38
Independent Variables						
DI[b]	9.34	3.44	1.90	6.77	11.93	14.14
CPI[b]	5.97	2.28	2.13	4.50	6.72	11.89
CCI[b]	126.11	23.76	85.53	107.36	147.99	164.42
CAR[b]	16.85	0.80	15.42	16.33	17.36	18.65
HDtI[b]	37.54	5.80	24.96	36.02	39.86	49.86
GD[b]	31.02	9.55	18.88	22.77	39.17	47.98
TD[c]	12.80	11.64	−2.95	4.81	15.57	42.78
Loans[c]	13.28	9.04	−3.46	6.31	18.16	34.10
GDP[c]	1.71	3.84	−10.10	−0.40	3.62	12.40
NSFR	1.04	0.07	0.90	0.98	1.09	1.14
HHI[b]	0.16	0.01	0.12	0.15	0.17	0.18

Notes: The sample period runs from 2007:IV-2022:III for the Brazilian financial system. *ATA* = adjusted total assets; *TL* = total liabilities; *TD* = total deposits; *DI* = interest rate (CDI); *AV* = assets volatility; *IPD* = idiosyncratic probability of default; *PD* = weighted probability of default; *CPI* = broad national consumer price index (IPCA) in 12 months; *CCI* = consumer confidence index; *CAR* = capital adequacy ratio; *HDtI* = household debt to income; *GD* = net public debt (federal government and Central Bank in terms of GDP); *Loans* = credit operations outstanding; *GDP* = gross domestic product at market prices (real growth rate) ; *NSFR* = proxy for the net stable funding ratio[d] and *HHI* = Herfindahl-Hirschman index for deposits concentration.
[a] In BRL billion.
[b] In percentage.
[c] In year-over-year (YoY) transformation.

[d]See Section 2.2.2.2.3 of Chapter 2 for more details.

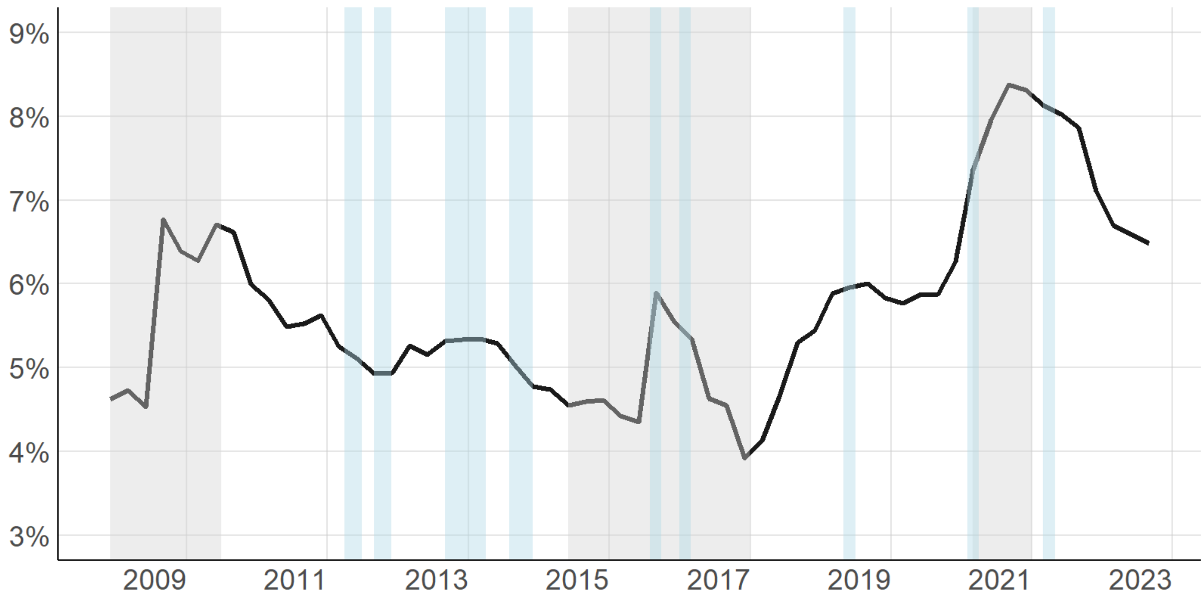
Figure 3.1: Correlation matrix of the dependent and independent variables.



3.4 Results and Discussion

This section investigates the determinants of the PD of the Brazilian banking system from December 2007 to September 2022, utilizing the BEVS procedure for this purpose. As detailed in Section 3.2, the PD is calculated using Equation 3.3 by weighting the IPD of individual FIs based on their deposits, and the application of the BEVS procedure is outlined in 3.2.5. Figure 3.2 presents the PD series, highlighting both periods of economic recession, as classified by CODACE (2023), and instances of extrajudicial settlements or interventions conducted by the BCB.

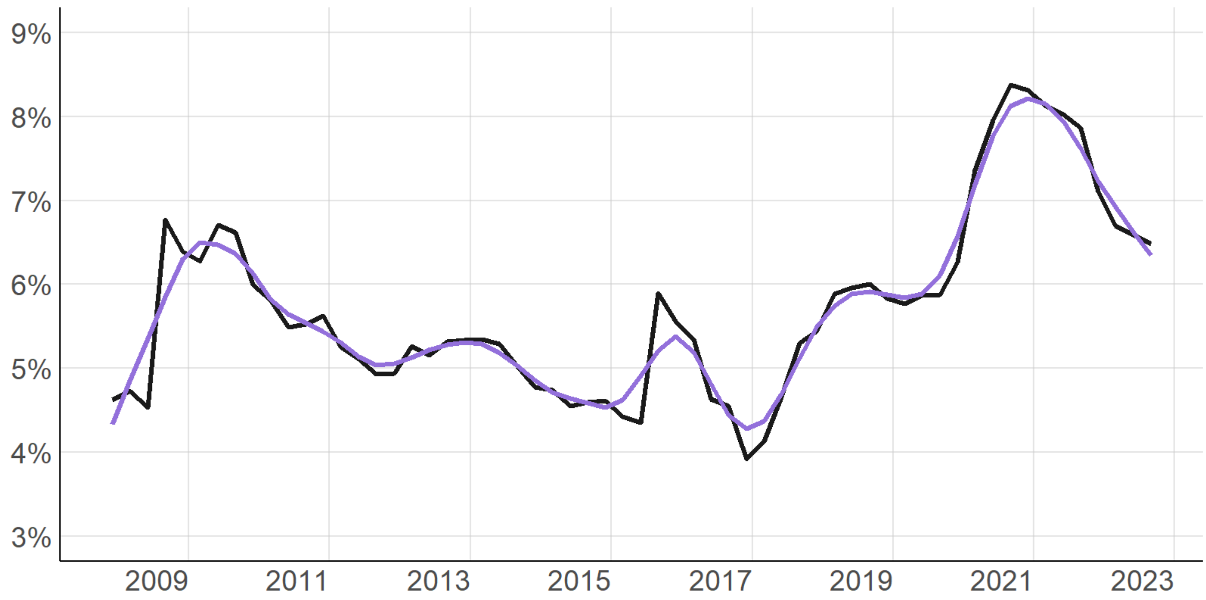
Figure 3.2: Probability of default of the Brazilian banking system.



Notes: Areas shaded in gray indicate periods of economic recession as dated by CODACE (2023), while areas shaded in blue represent periods of extrajudicial settlements or interventions made by the BCB in the banking sector.

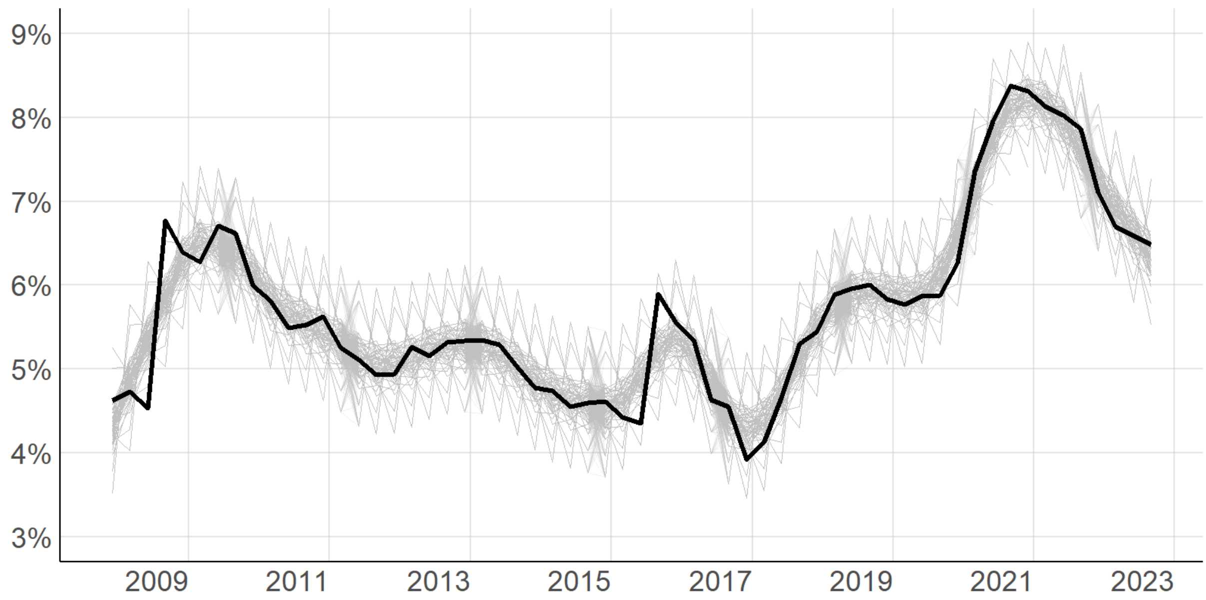
In the initial stage of implementing the BEVS framework, we first apply a Loess fit to the PD series using cross-validation to capture its underlying trends and specifically its low-frequency variations, as illustrated in Figure 3.3. Subsequently, we employ the circular block bootstrap technique to generate 1,000 bootstrapped series, thus preserving the temporal dependencies and structures inherent in the original PD series. These bootstrapped series serve as the basis for each Lasso fit and the final bagged model, and are presented in Figure 3.4.

Figure 3.3: Loess fit of the probability of default.



Notes: The line in purple represents the Loess fit of the PD in black.

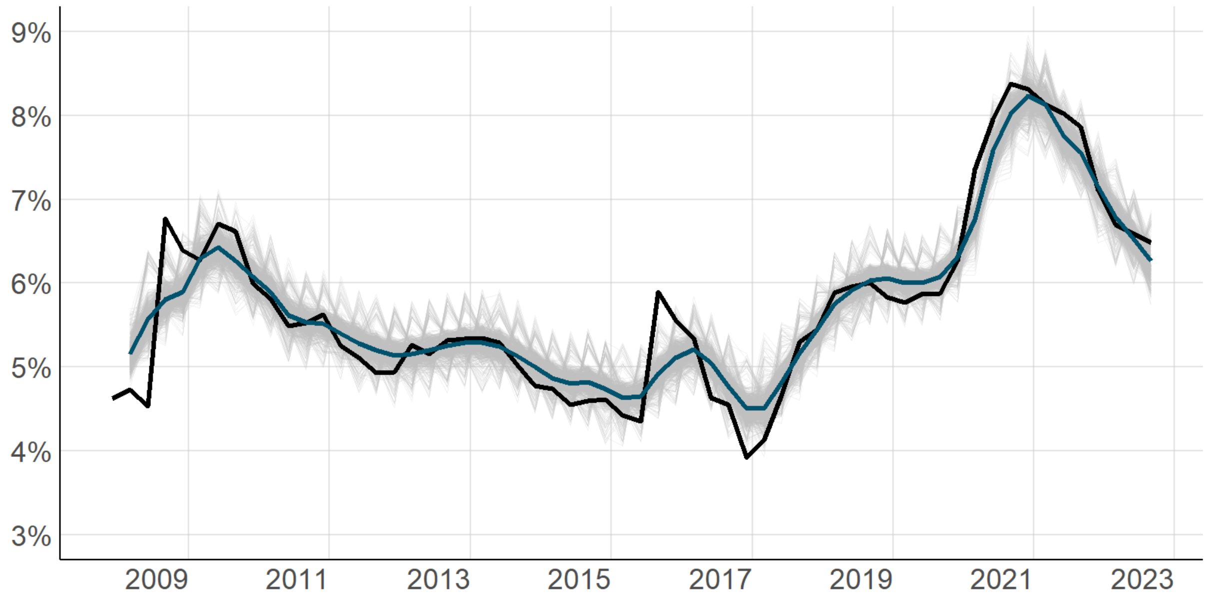
Figure 3.4: Bootstrapped series of the probability of default.



Notes: The lines in gray represent the bootstrap series of the PD in black.

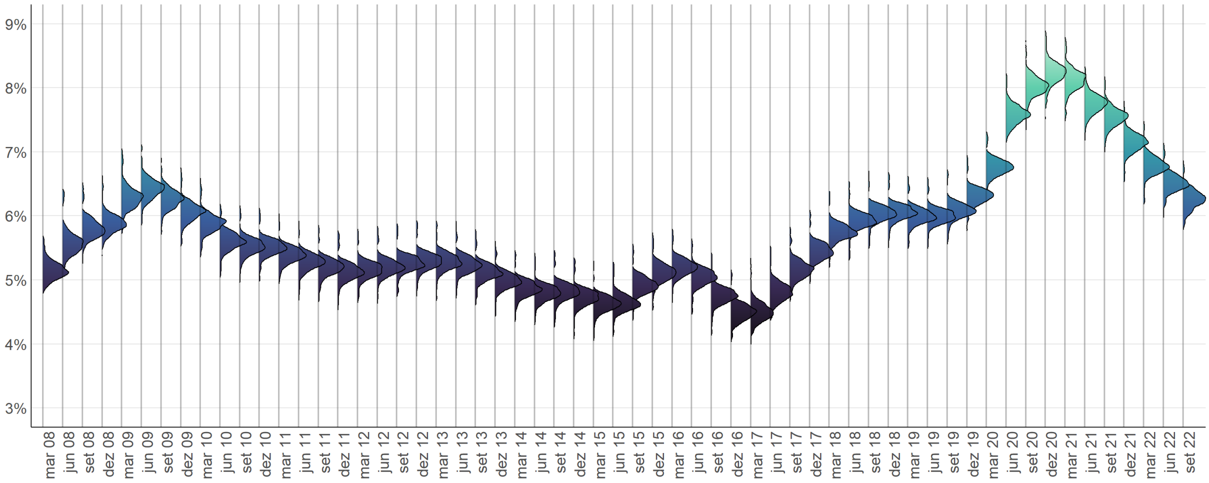
Upon completing the bootstrapping and Loess fitting phases, we construct an ensemble model named the 'BEVS model.' This model aggregates information from all individual Lasso models generated from the 1,000 bootstrapped series by averaging the non-zero coefficients, thus mitigating the uncertainty and risk associated with selecting a single model, as shown in Figure 3.5. The density distribution of these coefficients, presented in Figure 3.6, serves as an additional measure to understand the influence of the variables and the reliability of the coefficients. Specifically, each point on the density plot represents an estimate derived from one of the 1,000 individual Lasso models. A greater dispersion around the mean value indicates greater uncertainty in the coefficient estimates, while a narrower dispersion indicates increased reliability.

Figure 3.5: BEVS and Lasso fit of the probability of default.



Notes: The line in blue represents the BEVS fit of the PD (in black), which is the bagged model of all Lasso fit in the bootstrapped series (in gray) of the PD.

Figure 3.6: Density level of BEVS model.



Notes: Each point in time shows the distribution of the BEVS estimation for that particular point. The color of the distribution is related to the absolute value of the PD, in which the darker the color, the lower is the related value.

To assess the efficiency of the BEVS model, we compare it with a benchmark single Lasso model fitted on the original PD series. The coefficient values for both the benchmark Lasso model and the BEVS models before and after dimensionality reduction are shown in Table 3.2. This table is organized into three panels: Panel A contains the coefficients and values for the benchmark Lasso model; Panel B presents the BEVS model before dimensionality reduction; and Panel C shows the BEVS model after dimensionality reduction. The results of the residuals and other statistical tests for these models are presented in Table 3.3. This comparative analysis underscores the performance advantages and statistical robustness achieved by the BEVS approach.

In Table 3.2, it can be observed that, although the average coefficients of the variables present in both the benchmark Lasso and BEVS models are similar, the BEVS model has the distinct advantage of showing a distribution of possible coefficient values, illustrated by the percentile ranges. This feature not only enhances the model's statistical robustness but also allows for a more nuanced understanding of each variable's impact. Specifically, the range enables us to identify whether a variable's effect is consistently positive, consistently negative, or varies in sign, thereby broadening the scope for economic interpretation. Additionally, the number of appearances column in the BEVS models serves as a quantitative measure of variable significance. In particular, variables such as the autoregressive PD_{t-1}, the interest rate, and total deposits appear consistently across all bootstrapped iterations, reinforcing their importance for the estimation process.

When comparing the benchmark Lasso with the BEVS model, it is clear that each employs a distinct approach to variable selection, generating implications for model robustness and interpretability. While Lasso produces a single optimal set of coefficients based on minimizing the residual sum of squares across the entire dataset, BEVS leverages multiple bootstrap iterations to create an ensemble of models. This ensemble approach makes BEVS more sensitive to variables with smaller, although non-zero, impacts on the outcome variable, allowing it to capture marginal effects that may be overlooked by Lasso. Additionally, BEVS averages out the influence of data outliers or noise, resulting in a more stable set of variables. Importantly, this stability extends to the distribution of each coefficient across bootstrap samples, in contrast to Lasso's single-point estimate approach. By considering the coefficient distribution, BEVS not only ensures statistical robustness but also offers a nuanced understanding of the variability in each variable's impact.

Regarding the economic interpretation of the BEVS model, Table 3.2 shows how macroeconomic factors exert influence over the PD in the Brazilian banking system. As expected, we found an increase in the default risks during adverse economic conditions, as shown by the dynamics of the GDP, inflation, the interest rate, consumer confidence, household debt, and government debt, with great emphasis on the interest rate due to its recurrent selection in all models. Additionally, increases in the growth of total deposits and loans are associated with higher PD, indicating that the acceleration of these portfolios may reflect a deterioration in the overall risk profile of banks and should be closely monitored. We also observe results consistent with those detailed in Chapter 1, where a stronger capital adequacy ratio and higher market concentration are associated with lower PD. Finally, we also find that the persistence of the PD is significant, ranging from 41% up to 74% in all simulations.

In examining the number of appearances, it is important to address the counterintuitive impact of NSFR in the BEVS model without dimensionality reduction. Specifically, NSFR has a positive coefficient of 0.337[12] and appears in 8.3% of all simulations, in accordance with its positive correlation of 22.1% as shown in Figure 3.1. This result is counterintuitive because, from a regulatory perspective, a higher NSFR should contribute to a safer financial system. However, in terms of correlation, it is important to note that the NSFR metric

[12]Note that while the average and 50th percentile coefficients suggest a positive impact of NSFR, a negative coefficient is observed in less than 5% of the simulations. .

became a regulatory requirement in Brazil in October 2018 (BCB, 2022b), and all values before this date are constructed based on a proxy[13]. When the correlation is examined specifically for the period from December 2018 to September 2022, it changes to -22.2%, aligning more closely with the expected influence of NSFR on financial stability. Section 3.4.2 delves into the theoretical restrictions on variable signs relevant to this case.

As we observed an appearance of 8.3% of NSFR in the model, this aspect of variable importance is addressed in BEVS through the dimensionality reduction procedure, where variables appearing in less than 10% of the simulations are candidates for elimination. Given that NSFR falls under this criterion, all models that incorporate it as an explanatory variable are excluded from the bagging process, and this procedure continues until all variables appear in at least 10% of the remaining ones, leading to a more parsimonious model without compromising the robustness of the results, as shown in Table 3.3. All figures of the dimensionality reduction process can be found in the Appendix C.

[13]See Section 2.2.2.2.3 of Chapter 2 for more details.

Table 3.2: Summary of the benchmark Lasso and the BEVS model with and without dimensionality reduction.

Variable	Average Coefficient	1st Percentile	5th Percentile	50th Percentile	95th Percentile	99th Percentile	Number of Appearances
Panel A: Benchmark Lasso							
Intercept	2.700	-	-	-	-	-	-
PD_{t-1}	0.643	-	-	-	-	-	-
DI	-0.048	-	-	-	-	-	-
CCI	-0.001	-	-	-	-	-	-
CAR	-0.006	-	-	-	-	-	-
HDtI	0.016	-	-	-	-	-	-
TD	0.019	-	-	-	-	-	-
HHI	-4.984	-	-	-	-	-	-
Panel B: BEVS Without Dimensionality Reduction							
Intercept	4.075	2.261	2.744	4.082	5.387	5.915	1,000
PD_{t-1}	0.619	0.412	0.481	0.628	0.719	0.741	1,000
DI	-0.062	-0.108	-0.089	-0.060	-0.043	-0.032	1,000
CPI	0.032	-0.010	0.001	0.030	0.073	0.088	150
CCI	-0.002	-0.007	-0.005	-0.002	-0.000	-0.000	813
CAR	-0.008	-0.046	-0.038	-0.008	0.033	0.058	106
HDtI	0.006	0.000	0.000	0.005	0.017	0.020	360
GD	0.005	0.000	0.000	0.004	0.013	0.022	489
TD	0.012	0.005	0.008	0.012	0.017	0.020	1,000
Loans	0.007	0.000	0.001	0.006	0.020	0.025	287
GDP	-0.012	-0.033	-0.026	-0.011	-0.001	-0.000	772
NSFR	0.337	-1.222	0.000	0.254	1.060	1.712	83
HHI	-9.243	-16.604	-14.369	-9.313	-3.931	-1.362	990
Panel C: BEVS With Dimensionality Reduction							
Intercept	4.110	2.336	2.771	4.105	5.392	5.925	917
PD_{t-1}	0.618	0.410	0.481	0.627	0.718	0.738	917
DI	-0.062	-0.111	-0.088	-0.060	-0.043	-0.031	917
CPI	0.031	-0.011	0.000	0.028	0.073	0.089	138
CCI	-0.002	-0.006	-0.004	-0.002	-0.000	-0.000	764
CAR	-0.008	-0.046	-0.039	-0.008	0.034	0.059	103
HDtI	0.006	0.000	0.000	0.005	0.017	0.020	337
GD	0.005	0.000	0.000	0.004	0.013	0.022	457
TD	0.012	0.005	0.008	0.012	0.017	0.020	917
Loans	0.007	0.000	0.001	0.006	0.020	0.026	274
GDP	-0.012	-0.033	-0.026	-0.011	-0.002	-0.000	711
HHI	-9.203	-16.491	-14.315	-9.310	-3.900	-1.334	907

Notes: *DI* = interest rate (CDI); *CPI* = broad national consumer price index (IPCA) in 12 months; *CCI* = consumer confidence index; *CAR* = capital adequacy ratio; *HDtI* = household debt to income; *GD* = net public debt (federal government and Central Bank in terms of GDP); *TD* = YoY transformation of total deposits; *Loans* = YoY transformation of credit operations outstanding; *GDP* = YoY transformation of gross domestic product at market prices (real growth rate) ; *NSFR* = proxy for the net stable funding ratio and *HHI* = Herfindahl-Hirschman index for deposits concentration. All variables are expressed as a percentage.

Table 3.3: Statistical metrics for the benchmark Lasso and the BEVS model with and without dimensionality reduction.

Test	Benchmark Lasso	Full BEVS	Reduced BEVS
Ljung-Box (t-1)	0.938	0.051	0.050
Ljung-Box (t-2)	0.994	0.120	0.120
Ljung-Box (t-3)	0.747	0.217	0.218
Ljung-Box (t-4)	0.861	0.269	0.269
KS	0.051	0.318	0.306
D^2	0.882	0.894	0.893
AIC	-113.006	-101.684	-101.626
AICc	-110.853	-99.419	-99.376
BIC	-98.346	-87.037	-87.031
MASE	0.825	0.762	0.763
RMSE	0.377	0.323	0.323
R^2	0.882	0.913	0.913
Avg R^2	0.882	0.894	0.893
Number of Final Predictors	8	13	12

Notes: The specifications for each test are addressed in Section 3.2.5.

3.4.1 Robustness Test

In any statistical model that employs bootstrapping techniques, assessing the stability of the results under varying parameters is an important step to ensure robustness. This is especially the case for the BEVS model, which relies on a set of ensemble estimates generated from multiple bootstrap iterations. To this end, we conducted different robustness tests, such as (i) varying the number of bootstrap simulations and (ii) varying the bootstrap block size. This exercise aims to investigate whether the conclusions drawn from the BEVS model remain consistent when altering these parameters. Specifically, we examine how changes in (i) and (ii) influence the distribution of coefficients, the significance of variables, and, ultimately, the model's ability to reliably estimate the PD in the Brazilian banking system. The results of exercise (i) are presented in Tables 3.4 and 3.5, and the results of exercise (ii) are detailed in Tables C.1 and C.2 in Appendix C.

In Tables 3.4 and 3.5, the BEVS model shows stability when varying the number of bootstrap simulations from 100 to 50,000. Table 3.4 indicates minor fluctuations in metrics such as autocorrelation, information criteria, and performance measures, enhancing confidence in the capacity of the model to estimate PD in the Brazilian banking system.

In Table 3.5, similar consistency is observed in the average coefficients and the frequency of the variable appearances. Specifically, variables like the intercept and the autoregressive show almost no variation in their average coefficients or their appearance frequencies across all bootstrap iterations. The use of 1,000 simulations for the BEVS model is shown to be effective for stability, computational efficiency, and interpretability, especially with regard to the number of appearances metric.

Table 3.4: Statistical and performance metrics across different numbers of bootstrap simulations.

Test	Benchmark Simulation	Simulation 1	Simulation 2	Simulation 3	Simulation 4	Simulation 5
Panel A: Without Dimensionality Reduction						
Ljung-Box (t-1)	0.051	0.059	0.053	0.05	0.049	0.049
Ljung-Box (t-2)	0.12	0.129	0.124	0.12	0.119	0.119
Ljung-Box (t-3)	0.217	0.222	0.221	0.216	0.215	0.215
Ljung-Box (t-4)	0.269	0.272	0.27	0.265	0.263	0.263
KS	0.318	0.349	0.377	0.292	0.297	0.304
D^2	0.894	0.898	0.895	0.894	0.894	0.894
AIC	-101.684	-101.872	-101.785	-101.744	-101.701	-101.752
AICc	-99.419	-99.526	-99.495	-99.497	-99.464	-99.52
BIC	-87.037	-86.935	-87.047	-87.157	-87.149	-87.213
MASE	0.762	0.757	0.761	0.763	0.763	0.764
RMSE	0.323	0.32	0.322	0.324	0.324	0.324
R^2	0.913	0.914	0.914	0.913	0.913	0.913
Avg R^2	0.894	0.898	0.895	0.894	0.894	0.894
Number of Final Predictors	13	13	13	13	13	13
Maximum Predictors in Lasso	8	8	8	8	8	8
Dimension Reduction Rate	10%	10%	10%	10%	10%	10%
Number of Bootstrap Blocks	7	7	7	7	7	7
Bootstrap Sample Size	1,000	100	500	5,000	10,000	50,000
Computation Time	1.09 mins	9.87 secs	1.01 mins	9.13 mins	18.04 mins	1.54 hours
Panel B: With Dimensionality Reduction						
Ljung-Box (t-1)	0.05	0.059	0.053	0.049	0.049	0.049
Ljung-Box (t-2)	0.12	0.132	0.125	0.12	0.118	0.118
Ljung-Box (t-3)	0.218	0.229	0.223	0.216	0.214	0.214
Ljung-Box (t-4)	0.269	0.28	0.271	0.263	0.26	0.26
KS	0.306	0.304	0.377	0.29	0.29	0.299
D^2	0.893	0.897	0.895	0.894	0.894	0.894
AIC	-101.626	-101.71	-101.725	-101.709	-101.676	-101.729
AICc	-99.376	-99.363	-99.449	-99.469	-99.444	-99.502
BIC	-87.031	-86.775	-87.033	-87.149	-87.144	-87.212
MASE	0.763	0.758	0.763	0.762	0.763	0.763
RMSE	0.323	0.32	0.322	0.323	0.324	0.324
R^2	0.913	0.914	0.914	0.913	0.913	0.913
Avg R^2	0.893	0.897	0.895	0.894	0.894	0.894
Number of Final Predictors	12	12	12	12	12	12
Maximum Predictors in Lasso	8	8	8	8	8	8
Dimension Reduction Rate	10%	10%	10%	10%	10%	10%
Number of Bootstrap Blocks	7	7	7	7	7	7
Bootstrap Sample Size	1,000	100	500	5,000	10,000	50,000
Computation Time	1.09 mins	9.87 secs	1.01 mins	9.13 mins	18.04 mins	1.54 hours

Notes: Number of bootstrap simulations varies between 100 and 50,000.

Table 3.5: Coefficient and appearance performance across different numbers of bootstrap simulations.

	Benchmark Simulation		Simulation 1		Simulation 2		Simulation 3		Simulation 4		Simulation 5	
Variable	A. C.	%	A. C.	%	A. C.	%	A. C.	%	A. C.	%	A. C.	%
Intercept	4.075	100.0	4.242	100.0	4.124	100.0	4.055	100.0	4.067	100.0	4.059	100.0
PD_{t-1}	0.619	100.0	0.615	100.0	0.618	100.0	0.620	100.0	0.619	100.0	0.620	100.0
DI	-0.062	100.0	-0.064	100.0	-0.063	100.0	-0.062	100.0	-0.062	100.0	-0.061	100.0
CPI	0.032	15.0	0.029	15.0	0.030	16.6	0.031	15.3	0.031	14.9	0.030	14.2
CCI	-0.002	81.3	-0.002	86.0	-0.002	81.4	-0.002	80.5	-0.002	79.9	-0.002	80.0
CAR	-0.008	10.6	0.002	14.0	-0.007	11.2	-0.009	11.1	-0.009	11.1	-0.009	11.3
HDtI	0.006	36.0	0.005	38.0	0.006	35.4	0.006	36.2	0.006	35.5	0.006	35.4
GD	0.005	48.9	0.004	48.0	0.005	48.8	0.005	49.8	0.005	49.6	0.005	49.8
TD	0.012	100.0	0.011	100.0	0.012	100.0	0.012	100.0	0.012	100.0	0.012	100.0
Loans	0.007	28.7	0.008	31.0	0.007	30.4	0.007	26.7	0.007	27.2	0.007	27.2
GDP	-0.012	77.2	-0.012	77.0	-0.012	78.2	-0.012	76.0	-0.012	75.7	-0.012	75.5
NSFR	0.337	8.3	0.583	10.0	0.352	7.8	0.408	7.4	0.427	7.4	0.395	7.6
HHI	-9.243	99.0	-9.834	100.0	-9.396	99.6	-9.182	99.2	-9.223	99.2	-9.205	99.0

Notes: A. C. = Average Coefficient; *%* = Number of Appearances as a percentage of total simulations. Other variables are as previously described.

In Tables C.1 and C.2 presented in Appendix C, we also observe stability and robustness in the BEVS model when varying block sizes. Autocorrelation, KS measures, and other performance indicators show minimal variation, confirming the reliability of the model. Likewise, core variables, such as the intercept and autoregressive terms, remain stable in their average coefficients and appearance frequencies. These findings collectively indicate the robustness of the model and validate our choice of a block size that approximates the square root of the series length, as this balances computational efficiency, desired statistical properties, and interpretive clarity[14] (Demirel and Willemain, 2002).

3.4.2 Theoretical Sign Restrictions

In the analysis presented in Section 3.4, we employed the BEVS procedure without imposing any sign restrictions on the estimated coefficients. This approach was intended to estimate the possible signs of the relationships between the macroeconomic variables

[14]While the square root of the series length, rounded down to the nearest integer, serves as our benchmark for block size selection, alternative criteria could be employed taking into consideration: (i) statistical independence, achieved by minimizing inter-block autocorrelation through appropriate block size; (ii) computational efficiency, balancing the trade-off between block size and processing time; (iii) convergence behavior, assessing the rate at which estimates stabilize with varying block sizes and focusing on minimizing the RMSE to ensure more accurate and reliable estimates.; (iv) domain-specific requirements, guiding block sizes that correspond to inherent temporal structures of the data, such as natural units like months, quarters, or years, or to its seasonality, thereby enhancing the interpretability of the bootstrap estimates (Carlstein, 1986; Hall et al., 1995; Lahiri, 1999; Nordman, 2009).

and the probability of default of the Brazilian banking system in the 1,000 simulated series. Although these results offer a nuanced understanding of variable impacts, it was observed that some of the estimated magnitudes are complex due to the broad range of coefficient values across the quantiles. Thus, to align the model results with economic theory, we introduce sign restrictions as discussed in Section 3.2.5, aiming to enhance interpretability while cautiously limiting their scope to minimize model bias toward specific outcomes. These selected variables and their expected signs are shown in Table 3.6.

Table 3.6: Sign restrictions imposed on variables

Variable Name	Variable Description	Expected Sign
CPI	Broad National Consumer Price Index	Positive
CAR	Capital Adequacy Ratio	Negative
HDtI	Household Debt to Income	Positive
NSFR	Proxy for the Net Stable Funding Ratio	Negative

Notes: In cases where a negative relationship is expected, the coefficient estimates are constrained to the interval $[-\infty, 0]$, ensuring non-positive values. Similarly, for expected positive relationships, coefficients are restricted to $[0, +\infty]$, allowing only non-negative estimates. Where no prior directional expectation exists, coefficients are unrestricted with bounds $[-\infty, +\infty]$, allowing free estimation of optimal values. For more details, see Section 3.2.5.

The results of this restricted BEVS model, which incorporates the theoretical sign constraints as shown in Table 3.6, are summarized in Table 3.7. Note that the application of sign restrictions has refined the model's estimates to be more aligned with economic theory. For instance, in the case of the NSFR variable, the coefficients remain consistently negative in both models with and without dimensionality reduction, enhancing the variable's interpretive clarity. Furthermore, the imposition of sign restrictions led to an increase in the number of appearances for both NSFR and Loans. Specifically, while Loans was already part of the model after the dimensionality reduction and merely bolstered its representation, NSFR, which had fewer initial appearances, achieved the threshold to remain in the model after the reduction process.

However, the CAR, which is generally considered a significant determinant of a bank's probability of default, experienced a decrease in the frequency of its appearances and was ultimately excluded from the reduced BEVS model after the dimensionality reduction process. While this exclusion could be attributed to multicollinearity or model overfitting, the similar statistical metrics between the full and reduced BEVS models, as shown by metrics such as AIC, BIC, and D^2 in Table 3.8, suggest that the full BEVS model may still offer valuable insights into the influence of CAR on the probability of default in the Brazilian banking system.

Table 3.7: Summary of the benchmark Lasso and BEVS models incorporating theoretical sign restrictions.

Variable	Average Coefficient	1st Percentile	5th Percentile	50th Percentile	95th Percentile	99th Percentile	Number of Appearances
Panel A: Benchmark Lasso							
Intercept	3.064	-	-	-	-	-	-
PD_{t-1}	0.631	-	-	-	-	-	-
DI	-0.053	-	-	-	-	-	-
CCI	-0.001	-	-	-	-	-	-
CAR	-0.015	-	-	-	-	-	-
HDtI	0.019	-	-	-	-	-	-
TD	0.020	-	-	-	-	-	-
GDP	-0.005	-	-	-	-	-	-
HHI	-6.537	-	-	-	-	-	-
Panel B: BEVS Without Dimensionality Reduction							
Intercept	4.277	2.485	2.865	4.147	5.967	7.723	1,000
PD_{t-1}	0.582	0.303	0.381	0.607	0.711	0.733	1,000
DI	-0.074	-0.136	-0.118	-0.068	-0.048	-0.042	1,000
CPI	0.042	0.000	0.004	0.042	0.094	0.108	353
CCI	-0.003	-0.009	-0.006	-0.002	-0.000	-0.000	882
CAR	-0.015	-0.040	-0.037	-0.013	-0.000	-0.000	79
HDtI	0.006	0.000	0.000	0.005	0.016	0.019	334
GD	0.012	0.000	0.000	0.011	0.027	0.036	592
TD	0.011	0.003	0.006	0.011	0.017	0.021	1,000
Loans	0.014	0.000	0.001	0.012	0.031	0.039	598
GDP	-0.018	-0.045	-0.038	-0.017	-0.003	-0.000	870
NSFR	-1.339	-3.843	-3.172	-1.239	-0.070	-0.003	107
HHI	-8.445	-15.621	-14.002	-8.626	-2.651	-0.799	955
Panel C: BEVS With Dimensionality Reduction							
Intercept	4.283	2.507	2.847	4.141	6.033	7.720	921
PD_{t-1}	0.575	0.293	0.374	0.598	0.708	0.728	921
DI	-0.075	-0.136	-0.119	-0.070	-0.049	-0.042	921
CPI	0.042	0.000	0.004	0.042	0.094	0.108	345
CCI	-0.003	-0.009	-0.007	-0.003	-0.000	-0.000	805
HDtI	0.006	0.000	0.000	0.005	0.016	0.019	309
GD	0.012	0.000	0.000	0.011	0.028	0.036	568
TD	0.011	0.003	0.006	0.011	0.017	0.021	921
Loans	0.014	0.000	0.001	0.013	0.031	0.039	571
GDP	-0.019	-0.046	-0.039	-0.017	-0.003	-0.000	803
NSFR	-1.324	-3.846	-3.173	-1.236	-0.066	-0.003	106
HHI	-8.426	-15.707	-14.067	-8.565	-2.641	-0.775	876

Notes: DI = interest rate (CDI); *CPI* = broad national consumer price index (IPCA) in 12 months; *CCI* = consumer confidence index; *CAR* = capital adequacy ratio; *HDtI* = household debt to income; *GD* = net public debt (federal government and Central Bank in terms of GDP); *TD* = YoY transformation of total deposits; *Loans* = YoY transformation of credit operations outstanding; *GDP* = YoY transformation of gross domestic product at market prices (real growth rate) ; *NSFR* = proxy for the net stable funding ratio and *HHI* = Herfindahl-Hirschman index for deposits concentration. All variables are expressed as a percentage.

Table 3.8: Statistical metrics for the benchmark Lasso and BEVS models incorporating theoretical sign restrictions.

Test	Benchmark Lasso	Full BEVS	Reduced BEVS
Ljung-Box (t-1)	0.980	0.109	0.107
Ljung-Box (t-2)	1.000	0.173	0.167
Ljung-Box (t-3)	0.727	0.282	0.274
Ljung-Box (t-4)	0.835	0.335	0.329
KS	0.034	0.240	0.232
D^2	0.885	0.900	0.898
AIC	-111.513	-101.043	-100.993
AICc	-108.690	-98.244	-98.199
BIC	-94.758	-84.901	-84.873
MASE	0.807	0.743	0.746
RMSE	0.371	0.319	0.321
R^2	0.885	0.915	0.914
Avg R^2	0.885	0.900	0.898
Number of Final Predictors	9	13	12

Notes: The specifications for each test are addressed in Section 3.2.5.

3.5 Final Remarks

This paper proposes the Bootstrap Estimator with Variable Selection procedure to estimate the determinants of the probability of default of the Brazilian banking system as a case study over the period from December 2007 to September 2022. In this method, we combine techniques such as Lasso regression, Loess smoothing, and bagging, showing that this integrated approach yields improved results compared to those obtained through their individual performance. Our findings indicate that BEVS not only refines the estimate of PD but also offers a comprehensive view of the impact of macroeconomic factors over the study period.

The BEVS model introduces a significant enhancement in time series analysis. It generates a distribution of coefficients, providing a comprehensive view of variables' impacts, and utilizes the number of appearances of each variable as a robust measure of significance. In addition, the ensemble approach improves the detection of marginal effects often overlooked by single-model methods, while simultaneously neutralizing the influence of outliers, enhancing overall model stability. Furthermore, dimensionality reduction in

BEVS leads to a parsimonious, yet effective, model, ensuring efficiency without sacrificing analytical depth. Beyond the Brazilian banking system, the benefits provided by BEVS are applicable to a wide range of time series datasets, making it a versatile tool for various economic and financial applications.

Regarding our results, we contributed to the understanding of how adverse economic conditions influence the PD of the Brazilian banking system, with interest rates being an important element in these dynamics. In addition, we find that the growth of total deposits and loans is associated with higher PD, indicating that the acceleration of these portfolios may reflect a deterioration in the overall risk profile of banks and should be closely monitored by the supervisor.

Future research could extend BEVS analysis to multiple economies, offering a comparative study of the variable selection process and the frequency of number of appearances in diverse macroeconomic environments. Such comparative work could shed light on the unique economic factors that influence the stability of each region's banking system. Furthermore, exploring the interaction and relative impacts of these macroeconomic variables across economies could enhance our understanding of global financial dynamics and inform cross-border risk management strategies.

Concluding Remarks

This thesis consists of three self-contained essays that delve into systemic risk and banking within the Brazilian financial system, examining regulatory impacts, risk contributions of individual banks, and novel approaches to address financial stability. In Chapter 1, we examine the impact of capital regulation on banks' probability of default using the structural model of Merton (1974) and the Z-Score, confirming the importance of regulatory frameworks such as Basel III in enhancing financial stability. In Chapter 2, we evaluate the systemic risk contributions of individual banks and propose a bank run model that accounts for the idiosyncratic probability of default of banks and a systemic risk process through which additional defaults occur via different channels of contagion. Lastly, in Chapter 3, we propose the Bootstrap Estimator with Variable Selection procedure to estimate the determinants of the aggregate probability of default of banks, integrating techniques such as Lasso regression, Loess smoothing, and bagging.

The implications of our findings should be particularly informative for regulators and policymakers concerned with the management of financial stability and systemic risk in emerging markets, especially in terms of understanding the dynamics of systemic risk propagation through different channels of contagion. In practical terms, our models can be used to determine the optimal fund size for a Deposit Insurance Agency to effectively address widespread bank failures. Furthermore, our framework can be customized to incorporate country-specific factors, thus capturing the unique realities and challenges of different economies.

The possibilities for future research based on the findings of this thesis are vast. There are promising avenues in examining how changes in market conditions will affect financial stability, particularly in light of the growing importance of non-bank financial institutions, the insurance market, cryptocurrencies, and digital currencies. Additionally, the use of non-structured data for sentiment analysis and the application of Large Language Models

(LLMs) could provide valuable perspectives related to the management of systemic risk and financial stability. Finally, exploring the effects of climate risk on financial stability could also provide interesting insights as environmental concerns become increasingly integral to financial and regulatory frameworks.

Bibliography

Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge (2023). When Should You Adjust Standard Errors for Clustering? *The Quarterly Journal of Economics 138*(1), 1–35.

Abergel, F., B. K. Chakrabarti, A. Chakraborti, and A. Ghosh (2013). *Econophysics of systemic risk and network dynamics*. Springer.

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2015). Systemic risk and stability in financial networks. *American Economic Review 105*(2), 564–608.

Acharya, V., R. Engle, and M. Richardson (2012). Capital shortfall: A new approach to ranking and regulating systemic risks. *American Economic Review 102*(3), 59–64.

Acharya, V. V., A. N. Berger, and R. A. Roman (2018). Lending implications of u.s. bank stress tests: Costs or benefits? *Journal of Financial Intermediation 34*, 58–90. Assessing Banking Regulation During the Obama Era.

Acharya, V. V., T. F. Cooley, M. P. Richardson, and I. Walter (2010). *Regulating Wall Street: The Dodd-Frank Act and the New Architecture of Global Finance*. John Wiley & Sons.

Acharya, V. V., L. H. Pedersen, T. Philippon, and M. Richardson (2017). Measuring Systemic Risk. *The Review of Financial Studies 30*(1), 2–47.

Adrian, T. and M. K. Brunnermeier (2016). Covar. *American Economic Review 106*(7), 1705–41.

Alexandre, M., K. Michalak, T. C. Silva, and F. A. Rodrigues (2022). Efficiency-stability Trade-off in Financial Systems: a multi-objective optimization approach. Working Papers Series 568, Central Bank of Brazil, Research Department.

Allen, F., E. Carletti, and X. Gu (2019). The Roles of Banks in Financial Systems. In *The Oxford Handbook of Banking*. Oxford University Press.

Allen, F., E. Carletti, and A. Leonello (2011). Deposit insurance and risk taking. *Oxford Review of Economic Policy 27*(3), 464–478.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance 23*(4), 589–609.

Anand, K., B. Craig, and G. von Peter (2015). Filling in the blanks: network structure and interbank contagion. *Quantitative Finance 15*(4), 625–636.

Anand, K., I. van Lelyveld, Ádám Banai, S. Friedrich, R. Garratt, G. Hałaj, J. Fique, I. Hansen, S. M. Jaramillo, H. Lee, J. L. Molina-Borboa, S. Nobili, S. Rajan, D. Salakhova, T. C. Silva, L. Silvestri, and S. R. S. de Souza (2018). The missing links: A global study on uncovering financial network structures from partial data. *Journal of Financial Stability 35*, 107–119. Network models, stress testing and other tools for financial stability monitoring and macroprudential policy design and implementation.

Anderson, H., M. Paddrik, and J. J. Wang (2019). Bank networks and systemic risk: Evidence from the national banking acts. *American Economic Review 109*(9), 3125–61.

Anginer, D. and A. Demirgüç-Kunt (2019). Bank Runs and Moral Hazard: A Review of Deposit Insurance. In *The Oxford Handbook of Banking*. Oxford University Press.

Ardekani, A. M., I. Distinguin, and A. Tarazi (2020). Do banks change their liquidity ratios based on network characteristics? *European Journal of Operational Research 285*(2), 789–803.

B3 (2023). Brasil, Bolsa, Balcão. https://www.b3.com.br/en_us. Accessed January 15, 2023.

Bahaj, S. and F. Malherbe (2020). The forced safety effect: How higher capital requirements can increase bank lending. *The Journal of Finance 75*(6), 3013–3053.

Baker, A. C., D. F. Larcker, and C. C. Wang (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics 144*(2), 370–395.

Baltagi, B. H. (2021). *Econometric Analysis of Panel Data*. Number 978-3-030-53953-5 in Springer Texts in Business and Economics. Springer.

Bardoscia, M., S. Battiston, F. Caccioli, and G. Caldarelli (2015). Debtrank: A microscopic foundation for shock propagation. *PLOS ONE 10*(6), 1–13.

Bassett, W. F. and J. M. Berrospide (2018). The impact of post stress tests capital on bank lending. *Finance and Economics Discussion Series 2018*.

Battiston, S., M. Puliga, R. Kaushik, P. Tasca, and G. Caldarelli (2012). Debtrank: Too central to fail? financial networks, the fed and systemic risk. *Scientific reports 2*(1), 1–6.

BCB (2012). Basel III Accord in Brazil - Public Hearing Notice 40/2012. Accessed September 01, 2023.

BCB (2015). *Relatório de Estabilidade Financeira*, Volume 14. Banco Central do Brasil.

BCB (2016). *Relatório de Estabilidade Financeira*, Volume 15. Banco Central do Brasil.

BCB (2017). *Resolução nº 4.553, de 30 de janeiro de 2017*, Volume 1. Diário Oficial da União.

BCB (2018a). *Resolução nº 4.677, de 31 de julho de 2018*, Volume 1. Diário Oficial da União.

BCB (2018b). *Resolução nº 4.688, de 25 de setembro de 2018*, Volume 1. Diário Oficial da União.

BCB (2021a). *Resolução BCB nº 102, de 7 de junho de 2021*, Volume 1. Diário Oficial da União.

BCB (2021b). *Resolução BCB nº 110, de 1 de julho de 2021*, Volume 1. Diário Oficial da União.

BCB (2022a). *Relatório de Economia Bancária*. Banco Central do Brasil.

BCB (2022b). *Relatório de Estabilidade Financeira*, Volume 21. Banco Central do Brasil.

BCB (2023a). Banco Central do Brasil. IFData. https://www3.bcb.gov.br/ifdata/?lang=1. Accessed January 15, 2023.

BCB (2023b). Banco Central do Brasil. SGS. https://www3.bcb.gov.br/sgspub. Accessed January 15, 2023.

BCBS (2010). *Guidance for national authorities operating the countercyclical capital buffer.* Bank for International Settlements.

BCBS (2011). *Basel III: a global regulatory framework for more resilient banks and banking systems.* Bank for International Settlements.

BCBS (2013). *Basel III: The Liquidity Coverage Ratio and liquidity risk monitoring tools.* Bank for International Settlements.

BCBS (2014). *Basel III: the net stable funding ratio.* Bank for International Settlements.

BCBS (2017). *High-level summary of Basel III reforms.* Bank for International Settlements.

BCBS (2021). *Assessing the impact of Basel III: Evidence from macroeconomic models: literature review and simulations.* Bank for International Settlements.

BCBS (2022). *Evaluation of the impact and efficacy of the Basel III reforms.* Bank for International Settlements.

Beck, T., O. De Jonghe, and G. Schepens (2013). Bank competition and stability: Cross-country heterogeneity. *Journal of Financial Intermediation 22*(2), 218–244.

Begenau, J. (2020). Capital requirements, risk choice, and liquidity provision in a business-cycle model. *Journal of Financial Economics 136*(2), 355–378.

Bellini, T. (2017). Chapter 4 - portfolio credit risk modeling. In T. Bellini (Ed.), *Stress Testing and Risk Integration in Banks*, pp. 123–161. Academic Press.

Benoit, S., J.-E. Colliard, C. Hurlin, and C. Pérignon (2017). Where the Risks Lie: A Survey on Systemic Risk. *Review of Finance 21*(1), 109–152.

Berger, A. N. and C. H. Bouwman (2013). How does capital affect bank performance during financial crises? *Journal of Financial Economics 109*(1), 146–176.

Bergmeir, C., R. Hyndman, and J. Benítez (2016). Bagging exponential smoothing methods using stl decomposition and box–cox transformation. *International Journal of Forecasting 32*(2), 303–312.

Bisias, D., M. Flood, A. W. Lo, and S. Valavanis (2012). A survey of systemic risk analytics. *Annual Review of Financial Economics 4*(1), 255–296.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of political economy 81*(3), 637.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized arch model. *The Review of Economics and Statistics 72*(3), 498–505.

Bonaccorsi di Patti, E., M. Moscatelli, and S. Pietrosanti (2023). The impact of bank regulation on the cost of credit: Evidence from a discontinuity in capital requirements. *Journal of Financial Intermediation 55*, 101040.

Borio, C. (2011). Implementing the macro-prudential approach to financial regulation and supervision. *The financial crisis and the regulation of finance*, 101–117.

Bottazzi, G., A. De Sanctis, and F. Vanni (2020). Non-performing loans and systemic risk in financial networks. *Available at SSRN 3539741*.

Bouwman, C. H. S. (2019). Creation and Regulation of Bank Liquidity. In *The Oxford Handbook of Banking*. Oxford University Press.

Boyd, J. H. and D. E. Runkle (1993). Size and performance of banking firms: Testing the predictions of theory. *Journal of Monetary Economics 31*(1), 47–67.

Breiman, L. (1996). Bagging predictors. *Machine learning 24*(2), 123–140.

Brown, M., S. T. Trautmann, and R. Vlahu (2016). Understanding bank-run contagion. *Management Science 63*(7), 2272–2282.

Brownlees, C. and R. F. Engle (2017). SRISK: A Conditional Capital Shortfall Measure of Systemic Risk. *The Review of Financial Studies 30*(1), 48–79.

Brunnermeier, M. K. and M. Oehmke (2013). Chapter 18 - bubbles, financial crises, and systemic risk. Volume 2 of *Handbook of the Economics of Finance*, pp. 1221–1288. Elsevier.

Cabrales, A., P. Gottardi, and F. Vega-Redondo (2017). Risk Sharing and Contagion in Networks. *The Review of Financial Studies 30*(9), 3086–3127.

Caccioli, F., P. Barucca, and T. Kobayashi (2018). Network models of financial systemic risk: a review. *Journal of Computational Social Science 1*, 81–114.

Calabrese, R. and P. Giudici (2015). Estimating bank default with generalised extreme value regression models. *Journal of the Operational Research society 66*(11), 1783–1792.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics 225*(2), 200–230. Themed Issue: Treatment Effect 1.

Cardoso, V. R. d. S., L. A. Campos, J. A. Dantas, and O. R. d. Medeiros (2019). Factors associated with the structural liquidity of banks in Brazil. *Revista Contabilidade & Finanças 30*, 252–267.

Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics 14*(3), 1171–1179.

Castro Miranda, R. C., S. R. Stancato de Souza, T. C. Silva, and B. M. Tabak (2014). Connectivity and systemic risk in the Brazilian national payments system. *Journal of Complex Networks 2*(4), 585–613.

Chiaramonte, L. and B. Casu (2013). The determinants of bank cds spreads: evidence from the financial crisis. *The European Journal of Finance 19*(9), 861–887.

Chiaramonte, L., B. Casu, and R. Bottiglia (2013). *The Assessment of the Net Stable Funding Ratio (NSFR) Value. Evidence from the Financial Crisis*, pp. 83–94. London: Palgrave Macmillan UK.

Chiaramonte, L., E. Croci, and F. Poli (2015). Should we trust the Z-score? Evidence from the European Banking Industry. *Global Finance Journal 28*, 111–131.

Christiano, L. J., M. S. Eichenbaum, and M. Trabandt (2018). On DSGE models. *Journal of Economic Perspectives 32*(3), 113–40.

Cinelli, C. and T. C. Silva (2022). *NetworkRiskMeasures: Risk Measures for (Financial) Networks*. R package version 0.1.4.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association 74*(368), 829–836.

Cleveland, W. S. and S. J. Devlin (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association 83*(403), 596–610.

Coccorese, P. and L. Santucci (2019). The role of downward assets volatility in assessing the book-value distance to default. *Journal of Financial Economic Policy 11*, 485–504.

CODACE (2023). Comunicado do comitê de datação de ciclos econômicos. Instituto Brasileiro de Economia - FGV. Accessed October 08, 2023.

Cont, R., A. Moussa, and E. B. Santos (2013). *Network Structure and Systemic Risk in Banking Systems*, pp. 327–368. Cambridge University Press.

Couaillier, C. (2021). What are banks' actual capital targets? Working Paper Series 2618, European Central Bank.

Crosbie, P. and J. Bohn (2003). Modeling default risk. In *Modeling Methodology*, pp. 5–31. Moody's KMV Company.

Crouhy, M., D. Galai, and R. Mark (2000). A comparative analysis of current credit risk models. *Journal of Banking & Finance 24*(1), 59–117.

da Rosa München, D. (2022). The effect of financial distress on capital structure: The case of brazilian banks. *The Quarterly Review of Economics and Finance 86*, 296–304.

Dantas, J. A., O. R. d. Medeiros, and L. R. Capelleito (2012). Determinantes do spread bancário ex post no mercado brasileiro. *RAM. Revista de Administração Mackenzie 13*, 48–74.

Davydov, D., S. Vähämaa, and S. Yasar (2021). Bank liquidity creation and systemic risk. *Journal of Banking & Finance 123*, 106031.

De Bandt, O. and P. Hartmann (2019). Systemic Risk in Banking after the Great Financial Crisis. In *The Oxford Handbook of Banking*. Oxford University Press.

de Chaisemartin, C. and X. D'Haultfœuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review 110*(9), 2964–96.

De Lisa, R., S. Zedda, F. Vallascas, F. Campolongo, and M. Marchesi (2011). Modelling deposit insurance scheme losses in a basel 2 framework. *Journal of Financial Services Research 40*(3), 123–141.

De Nicolò, G. (2019). The costs and benefits of bank capital requirements. *Global Finance Journal 39*, 21–25. Prudential Financial Regulation, Insolvency Resolution and Safety Net Design.

Demirel, O. F. and T. R. Willemain (2002). Generation of simulation input scenarios using bootstrap methods. *The Journal of the Operational Research Society 53*(1), 69–78.

Diamond, D. and A. Kashyap (2016). Liquidity requirements, liquidity choice, and financial stability. Volume 2 of *Handbook of Macroeconomics*, pp. 2263–2303. Elsevier.

Diamond, D. W. and P. H. Dybvig (1983). Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy 91*(3), 401–419.

Ebrahimi Kahou, M. and A. Lehar (2017). Macroprudential policy: A review. *Journal of Financial Stability 29*, 92–105.

Eisenberg, L. and T. H. Noe (2001). Systemic risk in financial systems. *Management Science 47*(2), 236–249.

Engle, R. (2002). Dynamic conditional correlation. *Journal of Business & Economic Statistics 20*(3), 339–350.

Engle, R. (2009). *Anticipating Correlations: A New Paradigm for Risk Management*. Princeton University Press.

Engle, R. (2018). Systemic risk 10 years later. *Annual Review of Financial Economics 10*(1), 125–152.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica 50*(4), 987–1007.

Engle, R. F. (2016). Dynamic Conditional Beta. *Journal of Financial Econometrics 14*(4), 643–667.

Fernández-Aguado, P. G., E. T. Martínez, R. M. Ruíz, and A. P. Ureña (2022). Evaluation of european deposit insurance scheme funding based on risk analysis. *International Review of Economics & Finance 78*, 234–247.

Ferrara, G., S. Langfield, Z. Liu, and T. Ota (2019). Systemic illiquidity in the interbank network. *Quantitative Finance 19*(11), 1779–1795.

FGC (2023). Fundo Garantidor de Créditos. https://www.fgc.org.br/?lang=en-us. Accessed January 15, 2023.

Fiche, M., G. Souza, F. Basílio, and R. Ellery (2017). Determinantes do retorno financeiro dos bancos no brasil: uma análise acerca do spread ex-post. *Cadernos de Finanças Públicas 17*(1).

Fidrmuc, J. and R. Lind (2020). Macroeconomic impact of Basel III: Evidence from a meta-analysis. *Journal of Banking & Finance 112*, 105359. Challenges to global financial stability: interconnections,credit risk, business cycle and the role of market participants.

Fraisse, H., M. Lé, and D. Thesmar (2020). The real effects of bank capital requirements. *Management Science 66*(1), 5–23.

Freixas, X. and B. M. Parigi (2019). Lender of Last Resort: A New Role for the Old Instrument. In *The Oxford Handbook of Banking*. Oxford University Press.

FSB (2018). Principles on Bail-in Execution. pp. 1–26.

FSB (2021). Bail-in Execution Practices Paper. pp. 1–60.

Galati, G. and R. Moessner (2013). Macroprudential policy – a literature review. *Journal of Economic Surveys 27*(5), 846–878.

Gertler, M., N. Kiyotaki, and A. Prestipino (2019). A Macroeconomic Model with Financial Panics. *The Review of Economic Studies 87*(1), 240–288.

Giordana, G. A. and I. Schumacher (2017). An Empirical Study on the Impact of Basel III Standards on Banks' Default Risk: The Case of Luxembourg. *Journal of Risk and Financial Management 10*(2).

Glasserman, P. and H. P. Young (2016). Contagion in financial networks. *Journal of Economic Literature 54*(3), 779–831.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance 48*(5), 1779–1801.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics 225*(2), 254–277. Themed Issue: Treatment Effect 1.

Gorton, G. B. (2010). *Slapped by the invisible hand: The panic of 2007*. Oxford University Press.

Greenwood, R., A. Landier, and D. Thesmar (2015). Vulnerable banks. *Journal of Financial Economics 115*(3), 471–485.

Gropp, R., T. Mosk, S. Ongena, and C. Wix (2018). Banks Response to Higher Capital Requirements: Evidence from a Quasi-Natural Experiment. *The Review of Financial Studies 32*(1), 266–299.

Guerra, S. M., T. C. Silva, B. M. Tabak, R. A. de Souza Penaloza, and R. C. de Castro Miranda (2016). Systemic risk measures. *Physica A: Statistical Mechanics and its Applications 442*, 329–342.

Gupton, G., C. Finger, and M. Bhatia (2007). CreditMetrics – Technical Document. *The RiskMetrics Group*.

Hagendorff, J. (2019). Corporate Governance and Culture in Banking. In *The Oxford Handbook of Banking*. Oxford University Press.

Hall, P., J. L. Horowitz, and B.-Y. Jing (1995). On blocking rules for the bootstrap with dependent data. *Biometrika 82*(3), 561–574.

Hannoun, H. (2010). Towards a global financial stability framework. *45th SEACEN Governors' Conference*.

Hartigan, J. A. and M. A. Wong (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics) 28*(1), 100–108.

Hastie, T., R. Tibshirani, and J. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Volume 2. Springer.

Hastie, T., R. Tibshirani, and M. Wainwright (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations*. CRC press.

Huang, X., I. Vodenska, S. Havlin, and H. E. Stanley (2013). Cascading failures in bi-partite graphs: Model for systemic risk propagation. *Scientific Reports 3*.

Hurd, T. R. (2016). *Contagion!: Systemic Risk in Financial Networks*. Springer.

IBGE (2023). Instituto Brasileiro de Geografia e Estatística. SIDRA: Banco de Tabelas Estatísticas. https://sidra.ibge.gov.br/home/pnadct/brasil. Accessed January 15, 2023.

Imai, K. and I. S. Kim (2021). On the use of two-way fixed effects regression models for causal inference with panel data. *Political Analysis 29*(3), 405–415.

Jackson, M. O. and A. Pernoud (2021). Systemic risk in financial networks: A survey. *Annual Review of Economics 13*(1), 171–202.

Jancey, R. (1966). Multidimensional group analysis. *Australian Journal of Botany 14*(1), 127–130.

Jayadev, M. (2013). Basel III implementation: Issues and challenges for Indian banks. *IIMB Management Review 25*(2), 115–130.

Jobst, M. A. A. and M. D. F. Gray (2013). *Systemic contingent claims analysis: Estimating market-implied systemic risk*. International Monetary Fund.

Jones, L., R. Alsakka, O. ap Gwilym, and N. Mantovan (2022). The impact of regulatory reforms on european bank behaviour: A dynamic structural estimation. *European Economic Review 150*, 104280.

Karolyi, G. A., J. Sedunov, and A. G. Taboada (2023). Cross-Border Bank Flows and Systemic Risk*. *Review of Finance 27*(5), 1563–1614.

Kato, P. and J. Hagendorff (2010). Distance to default, subordinated debt, and distress indicators in the banking industry*. *Accounting & Finance 50*(4), 853–870.

Kaufman, L. and P. J. Rousseeuw (1990). *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons.

Kerlin, J. (2017). *The Role of Deposit Guarantee Schemes as a Financial Safety Net in the European Union*. Palgrave Macmillan Cham.

Kiss, H. J., I. Rodriguez-Lara, and A. Rosa-Garcia (2018). Panic bank runs. *Economics Letters 162*, 146–149.

Lahiri, S. N. (1999). Theoretical comparisons of block bootstrap methods. *The Annals of Statistics 27*(1), 386–404.

Le, T. N. L., M. A. Nasir, and T. L. D. Huynh (2020). Capital requirements and banks performance under Basel-III: A comparative analysis of Australian and British banks. *The Quarterly Review of Economics and Finance*.

Lehar, A. (2005). Measuring systemic risk: A risk management approach. *Journal of Banking & Finance 29*(10), 2577–2603.

Lepetit, L. and F. Strobel (2013). Bank insolvency risk and time-varying z-score measures. *Journal of International Financial Markets, Institutions and Money 25*, 73–87.

Li, L., Y. Kang, and F. Li (2023). Bayesian forecast combination using time-varying features. *International Journal of Forecasting 39*(3), 1287–1302.

Liberman, M., K. Barbosa, and J. Pires (2018). Falência bancária e capital regulatório: Evidência para o brasil. *Revista Brasileira de Economia 72*, 80 – 116.

Lloyd, S. (1957). Least squares quantization in PCM. *Published in 1982 in IEEE transactions on information theory 28*(2), 129–137.

MacQueen, J. (1967). Classification and analysis of multivariate observations. In *5th Berkeley Symp. Math. Statist. Probability*, pp. 281–297.

Magalhães-Timotioa, J. G., J. P. A. Eçab, and G. A. L. Filhoc (2018). Relação entre indicadores contábeis e o spread ex-post dos bancos brasileiros. *Revista de Administração, Contabilidade e Economia da Fundace 9*(2).

Malovaná, S. and D. Ehrenbergerová (2022). The effect of higher capital requirements on bank lending: the capital surplus matters. *Empirica 49*, 793–832.

Mare, D. S., F. Moreira, and R. Rossi (2017). Nonstationary z-score measures. *European Journal of Operational Research 260*(1), 348–358.

Maredza, A. (2016). Do capital requirements affect cost of intermediation? evidence from a panel of south african banks. *The Journal of Developing Areas 50*(3), 35–51.

Martin, A. (2006). Liquidity provision vs. deposit insurance: preventing bank panics without moral hazard. *Economic Theory 28*(1), 197–211.

Mateev, M., T. Nasr, and A. Sahyouni (2022). Capital regulation, market power and bank risk-taking in the mena region: New evidence for islamic and conventional banks. *The Quarterly Review of Economics and Finance 86*, 134–155.

Matt, R. T. and L. B. d. Andrade (2019). Estimativa do risco de concentração individual baseada em modelos de simulação de perdas em operações de crédito. *Revista do BNDES 26*(51), 101–142.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance 29*(2), 449–470.

Nguyen, T. C. (2021). Economic policy uncertainty and bank stability: Does bank regulation and supervision matter in major european economies? *Journal of International Financial Markets, Institutions and Money 74*, 101387.

Nier, E., J. Yang, T. Yorulmazer, and A. Alentorn (2007). Network models and financial stability. *Journal of Economic Dynamics and Control 31*(6), 2033–2060. Tenth Workshop on Economic Heterogeneous Interacting Agents.

Nordman, D. J. (2009). A note on the stationary bootstrap's variance. *The Annals of Statistics 37*(1), 359 – 370.

Noss, J. and P. Toffano (2016). Estimating the impact of changes in aggregate bank capital requirements on lending and growth during an upswing. *Journal of Banking & Finance 62*, 15–27.

Ognjenovic, D. (2017). *Deposit Insurance Schemes: Funding, Policy and Operational Challenges*. Palgrave Macmillan Cham.

O'Keefe, J. and A. Ufier (2017). Determining the Target Deposit Insurance Fund: Practical Approaches for Data-Poor Deposit insurers. In J. Nolte and I. Khan (Eds.), *Deposit Insurance Systems: Addressing emerging challenges in Funding, Investment, Risk-based Contributions & Stress Testing*. The World Bank Group.

Paltalidis, N., D. Gounopoulos, R. Kizys, and Y. Koutelidakis (2015). Transmission channels of systemic risk and contagion in the european financial network. *Journal of Banking & Finance 61*, S36–S52. Global Trends in Banking, Regulations, and Financial Markets.

Pariès, M. D., P. Karadi, C. Kok, and K. Nikolov (2022). The impact of capital requirements on the macroeconomy: Lessons from four macroeconomic models of the euro area. *International Journal of Central Banking 18*(5), 1–50.

Parrado-Martínez, P., P. Gómez-Fernández-Aguado, and A. Partal-Ureña (2019). Factors influencing the european bank's probability of default: An application of symbol methodology. *Journal of International Financial Markets, Institutions and Money 61*, 223–240.

Petersen, M. A. (2008). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *The Review of Financial Studies 22*(1), 435–480.

Petropoulos, F., R. J. Hyndman, and C. Bergmeir (2018). Exploring the sources of uncertainty: Why does bagging for time series forecasting work? *European Journal of Operational Research 268*(2), 545–554.

Pichler, A., S. Poledna, and S. Thurner (2021). Systemic risk-efficient asset allocations: Minimization of systemic risk as a network optimization problem. *Journal of Financial Stability 52*, 100809. Network models and stress testing for financial stability: the conference.

Poledna, S., S. Martínez-Jaramillo, F. Caccioli, and S. Thurner (2021). Quantification of systemic risk from overlapping portfolios in the financial system. *Journal of Financial*

Stability 52, 100808. Network models and stress testing for financial stability: the conference.

Politis, D. and J. Romano (1992). A Circular Block-Resampling Procedure for Stationary Data. *Exploring the Limits of Bootstrap 270*, 426.

R Core Team (2022). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Radev, D. (2022). *Measuring Systemic Risk: A Probabilistic Perspective*, Volume 409. Springer Nature.

Ramirez, J. (2017). *Handbook of Basel III Capital: Enhancing bank capital in practice.* John Wiley & Sons.

Robatto, R. (2019). Systemic banking panics, liquidity risk, and monetary policy. *Review of Economic Dynamics 34*, 20–42.

Rosa, P. S. and I. R. Gartner (2018). Financial distress in brazilian banks: an early warning model. *Revista Contabilidade & Finanças 29*(77), 312–331.

Shim, J. (2013). Bank capital buffer and portfolio risk: The influence of business cycle and revenue diversification. *Journal of Banking & Finance 37*(3), 761–772.

Silva, T. C., S. R. S. de Souza, and B. M. Tabak (2016). Network structure analysis of the brazilian interbank market. *Emerging Markets Review 26*, 130–152.

Silva, T. C., B. M. Tabak, and S. M. Guerra (2017). Why do vulnerability cycles matter in financial networks? *Physica A: Statistical Mechanics and its Applications 471*, 592–606.

Silva, W., H. Kimura, and V. A. Sobreiro (2017). An analysis of the literature on systemic financial risk: A survey. *Journal of Financial Stability 28*, 91–114.

Souza, S. R., T. C. Silva, B. M. Tabak, and S. M. Guerra (2016). Evaluating systemic risk using bank default probabilities in financial networks. *Journal of Economic Dynamics and Control 66*, 54–75.

Souza, S. R., B. M. Tabak, T. C. Silva, and S. M. Guerra (2015). Insolvency and contagion in the brazilian interbank market. *Physica A: Statistical Mechanics and its Applications 431*, 140–151.

Stewart, R. and M. Chowdhury (2021). Banking sector distress and economic growth resilience: Asymmetric effects. *The Journal of Economic Asymmetries 24*, e00218.

Sun, L., Y.-H. Huang, and T.-B. Ger (2018). Two-way cluster-robust standard errors—a methodological note on what has been done and what has not been done in accounting and finance research. *Theoretical Economics Letters 08*, 1639–1655.

Takeuti, L. J. (2020). O Impacto do Novo Requerimento de Liquidez (NSFR) no Spread e na Rentabilidade das Instituições Financeiras: Um Estudo Empírico no Período de 2001 a 2018. *Universidade de Brasília*.

Tarullo, D. K. (2019). Financial regulation: Still unsettled a decade after the crisis. *The Journal of Economic Perspectives 33*(1), 61–80.

Thakor, A. V. (2014). Bank capital and financial stability: An economic trade-off or a faustian bargain? *Annual Review of Financial Economics 6*(1), 185–223.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological) 58*(1), 267–288.

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73*(3), 273–282.

Tibshirani, R., G. Walther, and T. Hastie (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 63*(2), 411–423.

Tran, S., D. Nguyen, and L. Nguyen (2022). Concentration, capital, and bank stability in emerging and developing countries. *Borsa Istanbul Review 22*(6), 1251–1259.

Uluc, A. and T. Wieladek (2018). Capital requirements, monetary policy and risk shifting in the mortgage market. *Journal of Financial Intermediation 35*, 3–16. Banking and regulation: the next frontier.

Valahzaghard, M. and M. Bahrami (2013). Prediction of default probability in banking industry using CAMELS index: A case study of Iranian banks. *Management Science Letters 3*(4), 1113–1118.

Van Der Weide, M. E. and J. Y. Zhang (2019). 707Bank Capital Requirements after the Financial Crisis. In *The Oxford Handbook of Banking*. Oxford University Press.

Vazquez, F. and P. Federico (2015). Bank funding structures and risk: Evidence from the global financial crisis. *Journal of Banking & Finance 61*, 1–14.

Venkat, S. and S. Baird (2016). *Liquidity Risk Management: A Practitioner's Perspective*. John Wiley & Sons.

Wang, X., R. J. Hyndman, F. Li, and Y. Kang (2023). Forecast combinations: An over 50-year review. *International Journal of Forecasting 39*(4), 1518–1547.

Wen, B., S. Bi, M. Yuan, and J. Hao (2023). Financial constraint, cross-sectoral spillover and systemic risk in china. *International Review of Economics & Finance 84*, 1–11.

Wieser, C. (2022). *Quantification of Structural Liquidity Risk in Banks*. Springer Gabler Wiesbaden.

World Bank (2020). Bank Regulation and Supervision a Decade after the Global Financial Crisis.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control 18*(5), 931–955.

Appendix A

Chapter 1

A.1 Capital Requirements Transition

Table A.1: Transition schedule of capital requirements in Brazil.

	2013	2014	2015	2016	2017	2018	2019
Principal Capital - Basel III	3.500%	4.000%	4.500%	4.500%	4.500%	4.500%	4.500%
Principal Capital - Brazil	**4.500%**	**4.500%**	**4.500%**	**4.500%**	**4.500%**	**4.500%**	**4.500%**
Tier I Capital - Basel III	4.500%	5.500%	6.000%	6.000%	6.000%	6.000%	6.000%
Tier I Capital - Brazil	**5.500%**	**5.500%**	**6.000%**	**6.000%**	**6.000%**	**6.000%**	**6.000%**
Total Capital Ratio (TCR) - Basel III	8.000%	8.000%	8.000%	8.000%	8.000%	8.000%	8.000%
TCR - Brazil	**11.000%**	**11.000%**	**11.000%**	**9.875%**	**9.250%**	**8.625%**	**8.000%**
Additional Fixed Capital - Basel III	-	-	-	0.625%	1.250%	1.875%	2.500%
Additional Fixed Capital - Brazil	**-**	**-**	**-**	**0.625%**	**1.250%**	**1.875%**	**2.500%**
Counter-cyclical Additional Capital - Basel III	-	-	-	Up to 0.625%	Up to 1.250%	Up to 1.875%	Up to 2.5%
Counter-cyclical Additional Capital - Brazil	**-**	**Up to 0.625%**	**Up to 1.250%**	**Up to 1.875%**	**Up to 2.500%**	**Up to 2.500%**	**Up to 2.5%**
TCR + Fixed Additional Capital - Basel III	8.000%	8.000%	8.000%	8.625%	9.250%	9.875%	10.500%
TCR + Fixed Additional Capital - Brazil	**11.000%**	**11.000%**	**11.000%**	**10.500%**	**10.500%**	**10.500%**	**10.500%**
TCR + Max Counter-cyclical and Fixed Additional Capital - Basel III	8.000%	8.000%	8.000%	9.250%	10.500%	11.750%	13.000%
TCR + Max Counter-cyclical and Fixed Additional Capital - Brazil	**11.000%**	**11.625%**	**12.250%**	**12.375%**	**13.000%**	**13.000%**	**13.000%**

Notes: This table shows the transition schedule of capital requirements in Brazil according to Basel III and specific Brazilian regulations. Adapted from BCB (2012).

A.2 Balance sheets accounts and granular data

Table A.2: Balance sheets accounts.

Variable	Composition	Description
Adjusted Total Assets	(+)[10000007] (+)[20000004] (+)[49908008]	Current Assets and Long Term Receivables Fixed Assets Creditor for Advanced Residual Value
Equity	(+)[60000002] (+)[70000009] (+)[80000006]	Equity Gross Revenues Gross Expenses
Total Liabilities	(+)[40000008] (+)[50000005] (+)[60000002] (+)[70000009] (+)[80000006]	Current and Long Term Liabilities Deferred Income Equity Gross Revenues Gross Expenses
Loan, Lease and Other Credit Operations by Risk Level	(+)[31000000]	Loan, Lease and Other Credit Operations by Risk Level
Gross Interest Income (a)	(+)[71100001] (+)[71200004] (+)[71300007] (+)[71400000] (+)[71500003] (+)[71910002] (+)[71915007] (+)[71920009] (+)[71925004] (+)[71940003] (+)[71945008] (+)[71947006] (+)[71950000] (+)[71955005] (+)[71960007] (+)[71965002] (+)[71980001] (+)[71986005] (+)[71990053] (+)[71990101] (+)[71990125] (+)[71990156] (+)[71990204] (+)[71990266] (+)[81400007] (+)[81500000] (+)[81830055] (+)[81830103] (+)[81830127] (+)[81830158] (+)[81830206] (+)[81830268] (+)[81915004] (+)[81940000] (+)[81945005] (+)[81950007]	Total from Credit, Leases, Securities, Financial Derivative, BCB Mandatory Reserve Income and Foreign Exchange Net Income
Interest Expenses (b)	(+)[71300007] (+)[71990307] (+)[71990352] (+)[71990400] (+)[71990503] (+)[71990606] (+)[81100008] (+)[81200001] (+)[81300004] (+)[81400007] (+)[81830309] (+)[81830354] (+)[81830402] (+)[81830505] (+)[81830550] (+)[81830608] (+)[81912007] (+)[81960004] (+)[81980008] (+)[81986002]	Total Expenses of Funding, Borrowing and Onlending, Lease, Foreign Exchange Net Income and Net Loan Loss Provisions
Net Interest Income (c)	c = a + b	Sum of Gross Interest Income and Interest Expenses
Net Income	(+)[70000009] (+)[80000006] (-)[81956001]	Total of Net Operating Income and Net Non-operating Income, Income Tax Expenses and Profit Sharing

Notes: The numbers in the composition column for the dependent variables correspond to the Cosif balance sheet information, which is the accounting framework for all financial institutions in the Brazilian financial market. For more information, see BCB (2023a).

Table A.3: Effects of current and lagged CAR on FI's PD and Z-Score.

	Probability of Default				
	(1)	(2)	(3)	(4)	(5)
Capital Adequacy Ratio	-8.3943^{**}	-3.7089^{*}	-4.5305^{**}	-5.0312^{***}	-5.2941^{**}
	(3.5424)	(2.1385)	(2.0131)	(1.9495)	(2.0862)
Capital Adequacy Ratio$_{t-1}$		-4.0481^{*}	-3.1416^{**}	-3.0509^{**}	-2.1769^{**}
		(2.2651)	(1.2530)	(1.4530)	(0.9592)
Capital Adequacy Ratio$_{t-2}$			0.3966	-0.1216	-1.0109
			(1.9407)	(1.2005)	(1.4040)
Capital Adequacy Ratio$_{t-3}$				0.9082	2.1461
				(2.3340)	(2.3368)
Capital Adequacy Ratio$_{t-4}$					-1.4237
					(2.0425)
Observations	9,475	9,216	8,944	8,667	8,393
R^2	0.0064	0.0057	0.0052	0.0058	0.0066
Adjusted R^2	-0.0291	-0.0309	-0.0325	-0.0328	-0.0326
F Statistic	58.8755^{***}	25.2641^{***}	15.0741^{***}	12.0965^{***}	10.7337^{***}

Notes: This table presents the two-way fixed effects estimates of the FI's capital adequacy ratio on their PD and Z-Score. Both dependent and independent variables are in the natural log. Robust standard errors double-clustered are in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

A.3 Robustness Test

Table A.4: Effects of CAR and performance on FI's PD and Z-Score.

	PD			Z-Score		
	(1)	(2)	(3)	(4)	(5)	(6)
ROA	-1.18^{***}			0.07^{***}		
	(0.35)			(0.02)		
ROE		-0.00^{***}			0.00^{***}	
		(0.00)			(0.00)	
Spread			-0.87^{***}			0.06^{***}
			(0.33)			(0.02)
CAR	-3.48^{***}	-3.39^{***}	-3.16^{***}	0.37^{***}	0.40^{***}	0.35^{***}
	(0.77)	(0.65)	(0.75)	(0.04)	(0.04)	(0.05)
CAR_{t-1}	-0.25	-0.66	-0.85	0.07^{*}	0.11^{***}	0.13^{***}
	(0.78)	(0.60)	(0.75)	(0.04)	(0.04)	(0.04)
Observations	7,717	9,355	7,800	7,709	9,295	7,761
R^2	0.05	0.04	0.05	0.16	0.15	0.16
Adjusted R^2	0.01	0.00	0.01	0.12	0.12	0.12
F Statistic	140.34^{***}	122.82^{***}	133.44^{***}	463.85^{***}	547.41^{***}	455.43^{***}

Notes: This table presents the two-way fixed effects estimates of the FI's capital requirement and performance on their PD and Z-Score. *CAR* = capital adequacy ratio (tier I and II); *ROA* = return on assets and *ROE* = return on equity. All variables are in the natural log except *PD* and *ROE*. Robust standard errors double-clustered are in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

Appendix B

Chapter 2

B.1 Balance sheets accounts and granular data

Table B.1: Balance sheets accounts.

Variable	Composition	Description
Adjusted Total Assets	(+)[10000007] (+)[20000004] (+)[49908008]	Current Assets and Long Term Receivables Fixed Assets Creditor for Advanced Residual Value
Equity	(+)[60000002] (+)[70000009] (+)[80000006]	Equity Gross Revenues Gross Expenses
Total Liabilities	(+)[40000008] (+)[50000005] (+)[60000002] (+)[70000009] (+)[80000006]	Current and Long Term Liabilities Deferred Income Equity Gross Revenues Gross Expenses
Loan, Lease and Other Credit Operations by Risk Level	(+)[31000000]	Loan, Lease and Other Credit Operations by Risk Level
Interbank Investments	(+)[12000005]	Liquidity Interbank Investments
Interbank Deposits	(+)[41300006]	Interbank Deposits
Demand Deposits	(+)[41100000]	Demand Deposits
Saving Deposits	(+)[41200003]	Saving Deposits
Time Deposits	(+)[41500002]	Time Deposits

Notes: The numbers in the composition column for the dependent variables correspond to the Cosif balance sheet information, which is the accounting framework for all financial institutions in the Brazilian financial market. For more information, see BCB (2023a).

B.2 Additional Analysis

Table B.2: Effects of capital adequacy ratio and loans on FI's PD.

	Probability of Default
Capital Adequacy Ratio	-9.2193^{***}
	(3.2585)
Loans	2.4771^{*}
	(1.4439)
Observations	8,780
R^2	0.0148
Adjusted R^2	-0.0222
F Statistic	63.7530^{***}

Notes: This table presents the two-way fixed effects estimates of the FI's capital adequacy ratio and loan operations by risk level on their PD. The equations uses quarterly data from December 2000 to September 2022 for 244 Brazilian financial institutions, resulting in public 9.653 observations provided by the Central Bank of Brazil (BCB, 2023a). All variables are in the natural log. Robust standard errors double-clustered are in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

Table B.3: Effect of loans on FI's total deposits.

	Total Deposits
Loans	0.3727***
	(0.0665)
Observations	7,665
R^2	0.0962
Adjusted R^2	0.0602
F Statistic	784.5530***

Notes: This table presents the two-way fixed effects estimates of the FI's loan operations by risk level on their total deposits. The equations uses quarterly data from December 2000 to September 2022 for 244 Brazilian financial institutions, resulting in public 9.653 observations provided by the Central Bank of Brazil (BCB, 2023a). All variables are in the natural log. Robust standard errors double-clustered are in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

B.3 Codes

Listing B.1: Adapted minimum density method in R language.

```

min_dens_improved_fs = function (rowsums, colsums, c = 1, lambda = 1, k = 100, alpha = 1/sum(rowsums), delta = 1/sum(rowsums), theta = 1, remove.prob = 0.01, max.it = 100000, abs.tol = 0.001, verbose = TRUE, target_pr_max=25,round_matrix=FALSE, matrix_moments = FALSE)
{
  emp_results = data.table()
  a = rowsums
  l = colsums
  if (lambda > 1 | lambda < 0)
    stop("lambda must be between 0 and 1")
  n = length(a)
  X = matrix(0, n, n)
  mindex <- matrix(1:length(X), n, n)
  mu = 1:length(X)
  mu = mu[mu != diag(mindex)]
  v = numeric(0)
  ad = a - rowSums(X)
  ld = l - colSums(X)
  probs = Q(ad, ld, n)
  if (verbose)
    cat("Starting Minimum Density estimation.\n\n")
  for (t in 1:max.it) {
    if (t > k)
      lambda <- 1
    if ((runif(1) < remove.prob && t > 1 && length(v) > 0) ||
        sum(probs[mu]) == 0) {
      ij = sample(v, 1)
      index <- which(mindex == ij, arr.ind = T)
      i <- index[1]
      j <- index[2]
      ad[i] = ad[i] + X[ij]
      ld[j] = ld[j] + X[ij]
      X[ij] = 0
      mu = c(mu, ij)
```

```
      v = v[v != ij]
    }
    else {
      ci <- sample.int(length(mu), 1, prob = probs[mu])
      ij = mu[ci]
      index <- which(mindex == ij, arr.ind = T)
      i <- index[1]
      j <- index[2]
      Xnew = X
      Xnew[ij] <- lambda * min(ad[i], ld[j])
      adnew = ad
      adnew[i] = adnew[i] - Xnew[ij]
      ldnew = ld
      ldnew[j] = ldnew[j] - Xnew[ij]
      dif = V(Xnew, adnew, ldnew, c = c, alpha = alpha, delta = delta) -
    V(X, ad, ld, c = c, alpha = alpha, delta = delta)
      comp1 = dif > 0
      comp2 = exp(theta * dif) > runif(1)
      if (comp1 || comp2) {
        X = Xnew
        ad = adnew
        ld = ldnew
        v = c(v, ij)
        mu = mu[mu != ij]
      }
    }
    probs = Q(ad, ld, n)

    ### Adding modifications ###

    if (matrix_moments) {
      for (i in 1:nrow(X)) {
        aux5 = X[i,]
        aux8 = (aux5/PR_T1[i])*100
        aux8 = rep_inf_na(aux8) ; aux8 = rep_na_0(aux8)
        aux11 = ifelse(aux8>target_pr_max, (PR_T1[i]*target_pr_max/100),
    aux5)
        if (PR_T1[i]==0) {
          aux11 = rep(0,length(aux5))
```

```
      }
      X[i,] = aux11
    }

    if (round_matrix) {
      X = round(X,0)    }
    matrix_temp = graph_from_adjacency_matrix(X, weighted = T)
    matrix_den =  edge_density(matrix_temp)
    matrix_assort = assortativity_degree(matrix_temp)
    matrix_degree = mean(igraph::degree(matrix_temp))

  } else {
    matrix_den =  0
    matrix_assort = 0
    matrix_degree = 0
  }

  ### Finishing the modifications ###

  if (verbose)
    cat("- Iteration number: ", t, " -- total alocated: ", round(100 *
  (sum(a - ad)/sum(a)), 6), " %", " | Dens: ",round(matrix_den,4)," |
 Assort: ",round(matrix_assort,4)," | Grau:",round(matrix_degree,4), "
  \n",sep = "")

  ### Adding modifications ###
  stacked_results_aux = data.table(Iteration = t,Total_Allocated=round
 (100*(sum(a - ad)/sum(a)),6),Density=matrix_den,Assort=matrix_assort,
 Degree=matrix_degree,PR_Target = target_pr_max)
  stacked_results = rbind(stacked_results,stacked_results_aux)

  ### Finishing the modifications ###

  if (sum(abs(ad) - 0) < abs.tol)
    break
}
if (verbose) {
  if (t >= max.it)
    cat("\nMaximum number of iterations reached! Change the max.it
```

```r
    parameter or other settings.\n")
104     cat("\nMinimum Density estimation finished.", "\n * Total Number of Iterations: ", t, "\n * Total Alocated: ", round(100 * (sum(a - ad)/sum(a)), 6), " % \n", sep = "")
105   }
106   rownames(X) <- colnames(X) <- names(rowsums)
107   return(list(Matrix = X,Resultados_Empilhados = emp_results))
108 }
109
110 V = function(z, ad, ld, c = 1, alpha = 1, delta = 1)
111   -c*sum(z > 0) - sum((alpha^2)*ad) - sum((delta^2)*ld)
112
113 Q = function(ad, ld, n){
114   Q = rep.int(ad,  n)/rep(ld, each = n)
115   index = (Q < 1 | is.na(Q)) # Q < 1/Q
116   Q[index]  = (1/Q)[index]
117   Q[is.na(Q) | is.infinite(Q)] = 0
118   return(Q) }
```

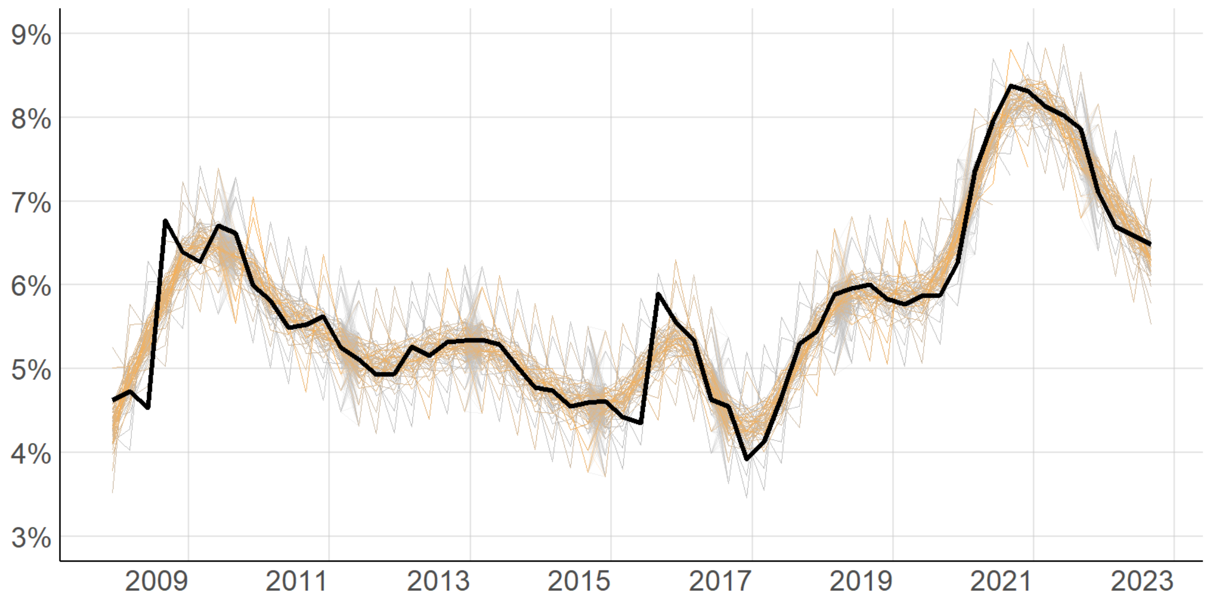
Notes: The original R Statistical Software (R Core Team, 2022) code is publicly available in the `NetworkRiskMeasures` package published by Cinelli and Silva (2022). The authors implemented the Minimum Density method based on Anand et al. (2015). All significant changes to the original code begin at line 60.

Appendix C

Chapter 3

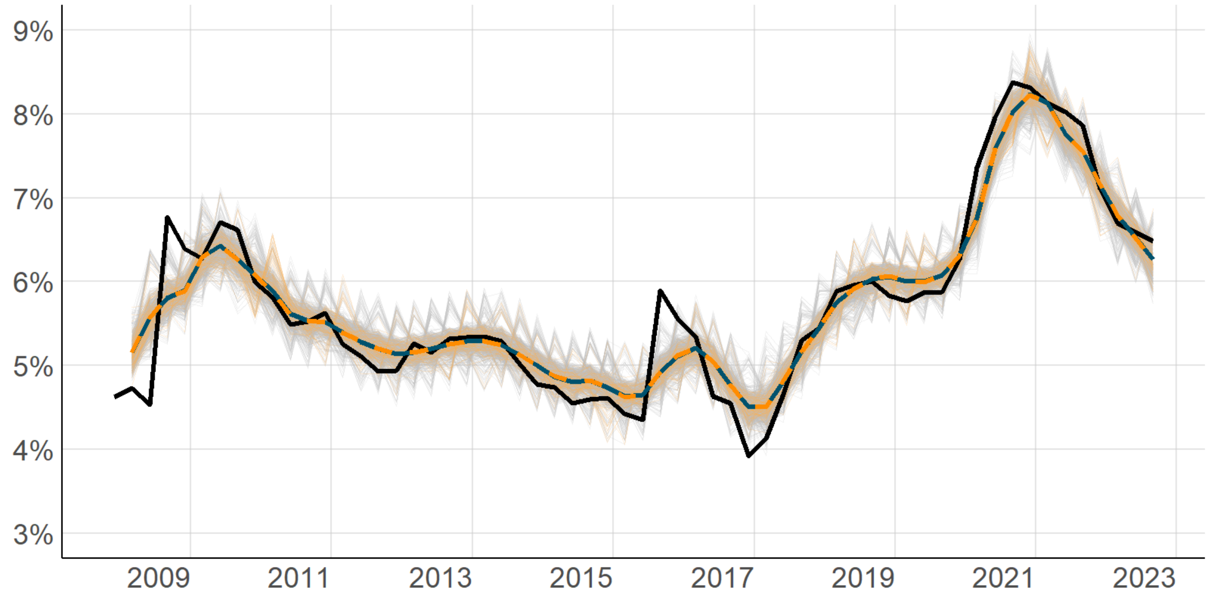
C.1 Dimensionality Reduction Procedure

Figure C.1: Bootstrapped series of the probability of default with dimensionality reduction.



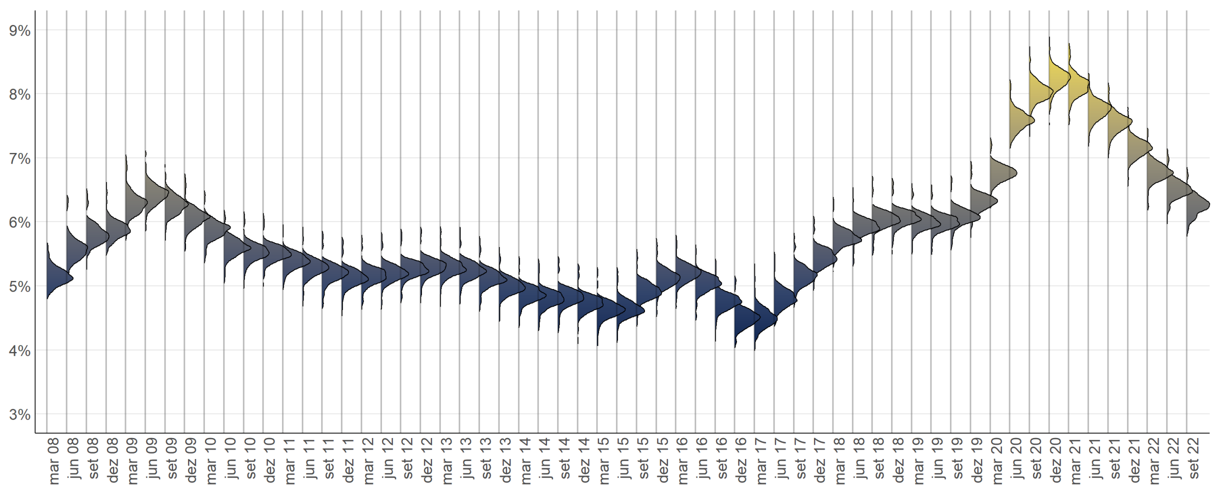
Notes: The lines in gray represent the bootstrap series of the PD in black. The lines in orange represents the removed bootstrapped series in the dimensionality reduction process.

Figure C.2: BEVS and Lasso fit of the probability of default with dimensionality reduction.



Notes: The line in blue represents the BEVS fit of the PD (in black), which is the bagged model of all Lasso fit in the bootstrapped series (in gray) of the PD. The solid lines in orange represents removed models in the dimensionality reduction process. The dashed line in orange represents the BEVS fit after the dimensionality reduction.

Figure C.3: Density level of BEVS model with dimensionality reduction.



Notes: Each point in time shows the distribution of the BEVS estimation for that particular point after the dimensionality reduction process. The color of the distribution is related to the absolute value of the PD, in which the darker the color, the lower is the related value.

C.2 Robustness Test

Table C.1: Statistical and performance metrics across different numbers of bootstrap blocks.

Test	Benchmark Simulation	Simulation 1	Simulation 2	Simulation 3	Simulation 4	Simulation 5
Panel A: Without Dimensionality Reduction						
Ljung-Box (t-1)	0.051	0.051	0.048	0.046	0.046	0.043
Ljung-Box (t-2)	0.12	0.12	0.117	0.113	0.112	0.108
Ljung-Box (t-3)	0.217	0.215	0.211	0.206	0.205	0.198
Ljung-Box (t-4)	0.269	0.261	0.258	0.257	0.259	0.25
KS	0.318	0.3	0.405	0.36	0.32	0.315
D^2	0.894	0.898	0.893	0.894	0.894	0.894
AIC	-101.684	-102.243	-101.677	-101.619	-101.495	-101.576
AICc	-99.419	-100.041	-99.425	-99.394	-99.243	-99.349
BIC	-87.037	-87.813	-87.068	-87.102	-86.89	-87.052
MASE	0.762	0.76	0.768	0.764	0.763	0.768
RMSE	0.323	0.321	0.324	0.323	0.324	0.325
R^2	0.913	0.914	0.912	0.913	0.913	0.912
Avg R^2	0.894	0.898	0.893	0.894	0.894	0.894
Number of Final Predictors	13	13	13	13	13	13
Maximum Predictors in Lasso	8	8	8	8	8	8
Dimension Reduction Rate	10%	10%	10%	10%	10%	10%
Number of Bootstrap Blocks	7	2	4	8	10	12
Bootstrap Sample Size	1,000	1,000	1,000	1,000	1,000	1,000
Computation Time	1.67 mins	1.4 mins	1.37 mins	1.37 mins	1.18 mins	1.62 mins
Panel B: With Dimensionality Reduction						
Ljung-Box (t-1)	0.05	0.049	0.047	0.045	0.046	0.044
Ljung-Box (t-2)	0.12	0.118	0.116	0.111	0.113	0.11
Ljung-Box (t-3)	0.218	0.212	0.21	0.204	0.207	0.201
Ljung-Box (t-4)	0.269	0.256	0.256	0.251	0.257	0.249
KS	0.306	0.271	0.395	0.367	0.308	0.307
D^2	0.893	0.897	0.892	0.893	0.894	0.894
AIC	-101.626	-102.254	-101.638	-101.615	-101.545	-101.576
AICc	-99.376	-100.06	-99.391	-99.396	-99.298	-99.352
BIC	-87.031	-87.86	-87.052	-87.122	-86.962	-87.065
MASE	0.763	0.759	0.767	0.764	0.762	0.767
RMSE	0.323	0.321	0.324	0.323	0.324	0.325
R^2	0.913	0.914	0.912	0.913	0.913	0.912
Avg R^2	0.893	0.897	0.892	0.893	0.894	0.894
Number of Final Predictors	12	12	12	12	12	12
Maximum Predictors in Lasso	8	8	8	8	8	8
Dimension Reduction Rate	10%	10%	10%	10%	10%	10%
Number of Bootstrap Blocks	7	2	4	8	10	12
Bootstrap Sample Size	1,000	1,000	1,000	1,000	1,000	1,000
Computation Time	1.67 mins	1.4 mins	1.37 mins	1.37 mins	1.18 mins	1.62 mins

Notes: Number of bootstrap blocks varies between 2 to 12.

Table C.2: Coefficient and appearance performance across different numbers of bootstrap blocks.

	Benchmark Simulation		Simulation 1		Simulation 2		Simulation 3		Simulation 4		Simulation 5	
Variable	A. C.	%	A. C.	%	A. C.	%	A. C.	%	A. C.	%	A. C.	%
Intercept	4.075	100.0	4.023	100.0	4.055	100.0	4.030	100.0	4.022	100.0	4.043	100.0
PD_{t-1}	0.619	100.0	0.627	100.0	0.619	100.0	0.620	100.0	0.620	100.0	0.621	100.0
DI	-0.062	100.0	-0.060	100.0	-0.061	100.0	-0.061	100.0	-0.061	100.0	-0.061	100.0
CPI	0.032	15.0	0.026	11.2	0.026	15.1	0.028	13.7	0.029	15.7	0.027	14.5
CCI	-0.002	81.3	-0.002	82.6	-0.002	80.0	-0.002	79.7	-0.002	79.8	-0.002	79.8
CAR	-0.008	10.6	-0.008	11.8	-0.010	10.9	-0.011	10.8	-0.008	11.5	-0.012	12.1
HDtI	0.006	36.0	0.006	35.5	0.006	37.2	0.007	35.8	0.006	33.2	0.006	36.2
GD	0.005	48.9	0.005	46.7	0.005	50.8	0.006	49.9	0.005	50.2	0.005	52.3
TD	0.012	100.0	0.012	100.0	0.012	100.0	0.012	99.9	0.012	100.0	0.012	100.0
Loans	0.007	28.7	0.007	28.7	0.007	29.6	0.008	29.5	0.007	29.6	0.007	25.7
GDP	-0.012	77.2	-0.013	70.8	-0.013	72.7	-0.013	73.4	-0.012	75.9	-0.012	71.5
NSFR	0.337	8.3	0.476	7.8	0.337	7.2	0.414	7.3	0.495	8.1	0.399	7.7
HHI	-9.243	99.0	-9.119	99.5	-9.107	99.7	-9.110	98.8	-9.091	99.0	-9.070	99.3

Notes: A. C. = Average Coefficient; % = Number of Appearances as a percentage of total simulations. Other variables are as previously described.

C.3 Balance sheets accounts and granular data

Table C.3: Balance sheets accounts and macroeconomic variables.

Variable	Composition	Description
Dependent Variable		
Adjusted Total Assets	(+)[10000007]	Current Assets and Long Term Receivables
	(+)[20000004]	Fixed Assets
	(+)[49908008]	Creditor for Advanced Residual Value
Total Liabilities	(+)[40000008]	Current and Long Term Liabilities
	(+)[50000005]	Deferred Income
	(+)[60000002]	Equity
	(+)[70000009]	Gross Revenues
	(+)[80000006]	Gross Expenses
Demand Deposits	(+)[41100000]	Demand Deposits
Saving Deposits	(+)[41200003]	Saving Deposits
Time Deposits	(+)[41500002]	Time Deposits
Independent Variable		
Interest Rate	4389	Interest rate CDI in annual terms (basis 252)
Broad National Consumer Price Index	13522	Broad National Consumer Price Index (IPCA) in 12 months
Consumer Confidence Index	4393	Consumer confidence index
Capital Adequacy Ratio	21424	Regulatory Capital to Risk-Weighted Assets
Household Debt to Income	29037	Household debt to income (Households gross disposable national income)
Net Public Debt	4503	Net public debt (% GDP) - Total Federal Government and Central Bank
Total Deposits	27790 27805 1835	Demand + Time + Savings deposits (end-of-period balance)
Credit Operations Outstanding	20539	Total Credit operations outstanding
Gross Domestic Product at Market Prices	6561	Gross Domestic Product at Market Prices - Quarterly Rate (compared to the same period of the previous year) - Table 5932

Notes: The numbers in the composition column for the dependent variables correspond to the Cosif balance sheet information, which is the accounting framework for all financial institutions in the Brazilian financial market. The numbers for the independent variables in this same column are based on the codes of macroeconomic variables from BCB (2023b), except for GDP, which comes from IBGE (2023). For additional information on Cosif balance sheet data, refer to BCB (2023a).