Bootstrap Estimator Approach to Financial Stability

Wesley Augusto de Freitas Borges[†]

Roseli Basso-Silva[‡]

March 21, 2024

Abstract

This paper proposes the Bootstrap Estimator with Variable Selection (BEVS) procedure to estimate the determinants of the probability of default (PD) in the Brazilian banking system as a case study. 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.

JEL-Codes: C51, C63, E00, G01, G21.

Keywords: Banks, Probability of Default, Bootstrap, Lasso.

[†]Fundo Garantidor de Créditos (FGC) and University of São Paulo - FEARP/USP, Department of Economics, Brazil. The views expressed in this work are those of the author and do not involve the responsibility, should not be reported as and do not necessarily reflect those of the Fundo Garantidor de Créditos – FGC or its members. E-mail: waf.borges@outlook.com.

[‡]University of São Paulo - FEARP/USP, Department of Economics, Brazil. E-mail: roselisilva@usp.br.

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 (PD) across all financial institutions (FIs)[1].

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 (DIA), *Fundo Garantidor de Créditos* (FGC), 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

[1]In this paper, 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).

of appearances.

This work is structured into five sections, beginning with this Introduction. Section 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 describes the data used in our models. Section 4 presents the results and a discussion of our findings, and Section 5 concludes with the final remarks of this paper.

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.

2.1 Probability of Default

In this work, 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[2] (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 1.

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

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

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

[2] 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.

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

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 2.

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

Note that the probability of default is the area under the default barrier, that is, a fraction of total liabilities. Also, note that the negative of d_{2it} can also be used to compute the distance to distress ($D2D$) for a risk neutral environment, which is the distance of the bank's asset value to the distress barrier in $t = 0$, measured in assets value' standard deviations.

Finally, to calculate our aggregate PD for the entire Brazilian financial system, we utilize the following expression given by 3.

$$PD_t = \frac{\sum_{i=1}^{N} IPD_{it} \times \text{Deposits}_{it}}{\sum_{i=1}^{N} \text{Deposits}_{it}}, \quad \forall t \in \{1, \ldots, T\} \tag{3}$$

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 4.

$$\arg\min_{\beta_0,\boldsymbol{\beta}} \left\{ \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 \right\} \qquad s.t. \quad \sum_{j=1}^{p} |\beta_j| \leq t \tag{4}$$

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 5.

$$\hat{\beta}^{lasso} = \arg\min_{\beta_0,\boldsymbol{\beta}} \left\{ \frac{1}{2} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \tag{5}$$

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.

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 6.

$$W(x) = \begin{cases} (1-|d|^3)^3, & \text{for } |d| < 1, \\ 0, & \text{for } |d| \geq 1 \end{cases} \tag{6}$$

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[3]. The weight assigned to a 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]Let W be a weight function with the following properties:

1. $W(x) > 0$ for $|x| < 1$;
2. $W(-x) = W(x)$;
3. $W(x)$ is a non-increasing function for $x \geq 0$;
4. $W(x) = 0$ for $|x| \geq 1$.

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 7.

$$\hat{f}_{avg}(x) = \frac{1}{B} \sum_{b=1}^{B} \hat{f}^b(x) \tag{7}$$

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 8.

$$\hat{f}_{bag}(x) = \frac{1}{B} \sum_{b=1}^{B} \hat{f}^{*b}(x) \tag{8}$$

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.

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[4]. 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.

[4]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[5]. 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

[5]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[6]. 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[7], 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)[8]. 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,

[6]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.

[7]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 4.

[8]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[9]. 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[10]: (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[11]. The metric D^2 represents the fraction of deviance explained[12]. For

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

[10]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.

[11]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.

[12]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 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).

The dataset considers financial conglomerates and independent institutions until December 2014, and the prudential conglomerates and independent institutions before March 2015[13] with the business model category of $b1$, $b2$, $b4$, and $n1$[14], 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

[13]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).

[14]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 (2021).

Merton (1974)'s structural model, we applied the following variables: adjusted total assets[15] 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 utilizing the deposits of individual FIs, yielding our final time series of 60 observations. Table 1 presents the aggregate descriptive statistics for these variables, Figure 1 presents the correlation matrix, and all balance sheet accounts and code variables are shown in Appendix C.

[15]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.

Table 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

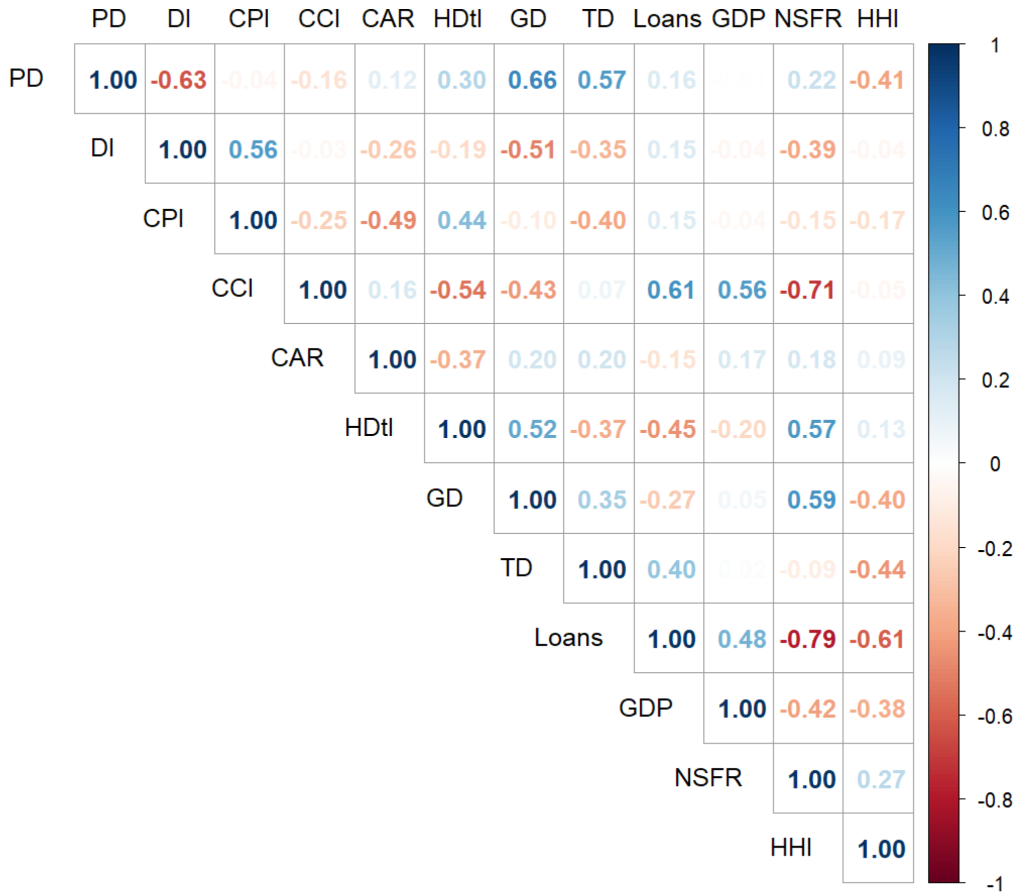
Notes: The sample period runs from 2007:IV-2022:III for the Brazilian financial system. *ATA* = adjusted total assets; *TL* = total liabilities; *TD* = total deposits; *DI* = interest rate (CDI); *AV* = assets volatility; *IPD* = idiosyncratic probability of default; *PD* = weighted probability of default; *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 and *HHI* = Herfindahl-Hirschman index for deposits concentration.

[a] In BRL billion.

[b] In percentage.

[c] In year-over-year (YoY) transformation.

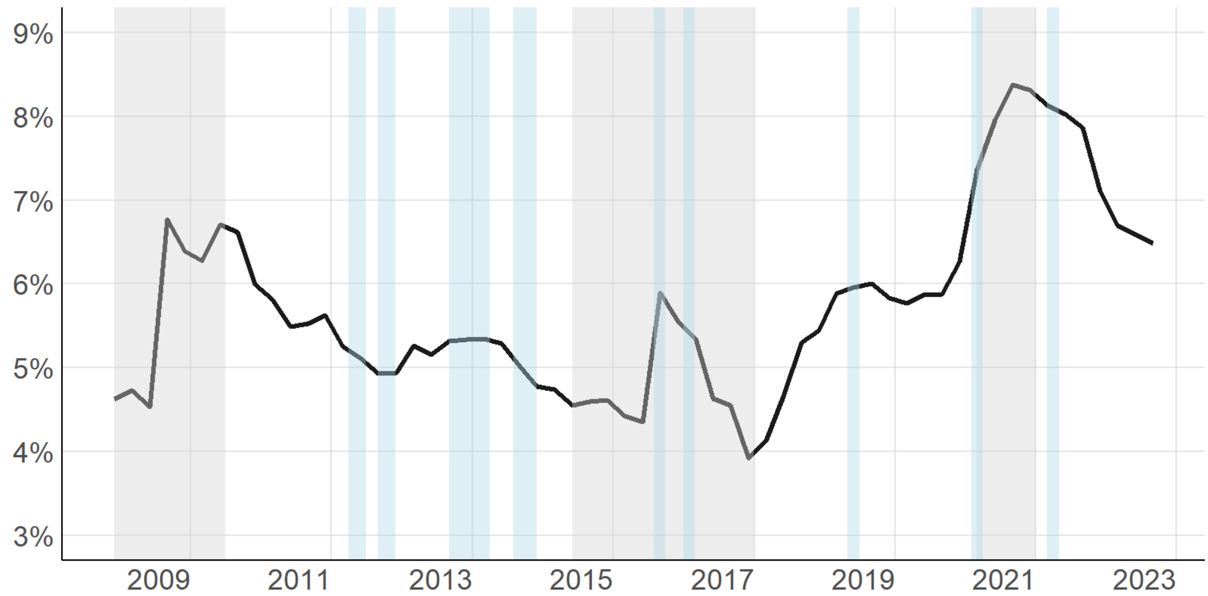
Figure 1: Correlation matrix of the dependent and independent variables.



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 2, the PD is calculated using Equation 3 by weighting the IPD of individual FIs based on their deposits, and the application of the BEVS procedure is outlined in 2.5. Figure 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.

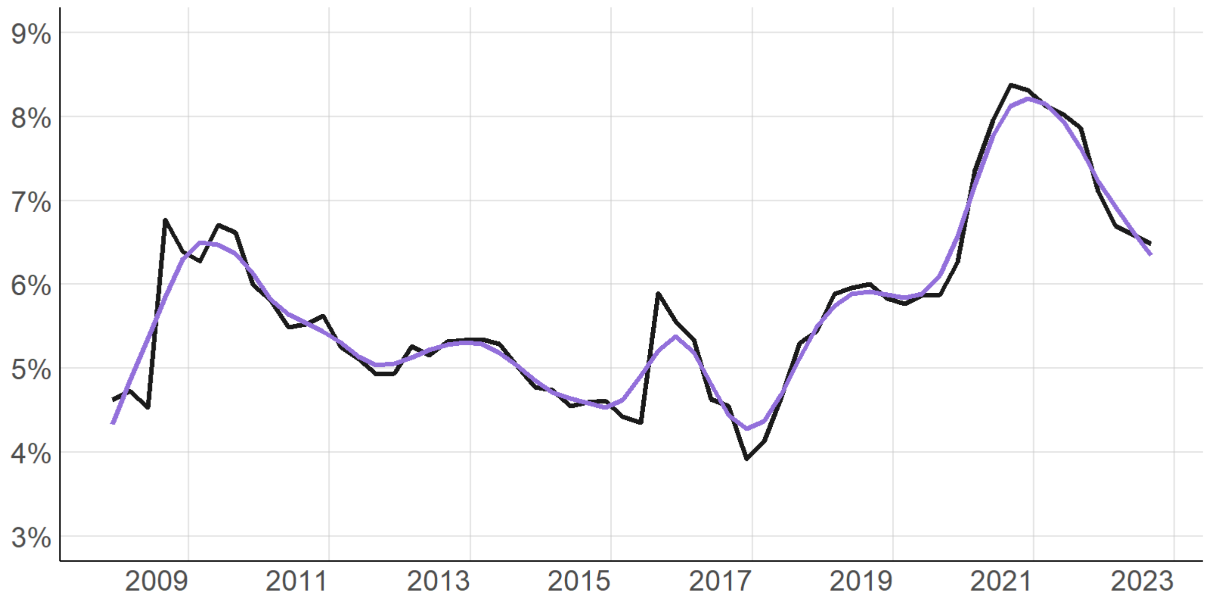
Figure 2: Probability of default of the Brazilian banking system.



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.

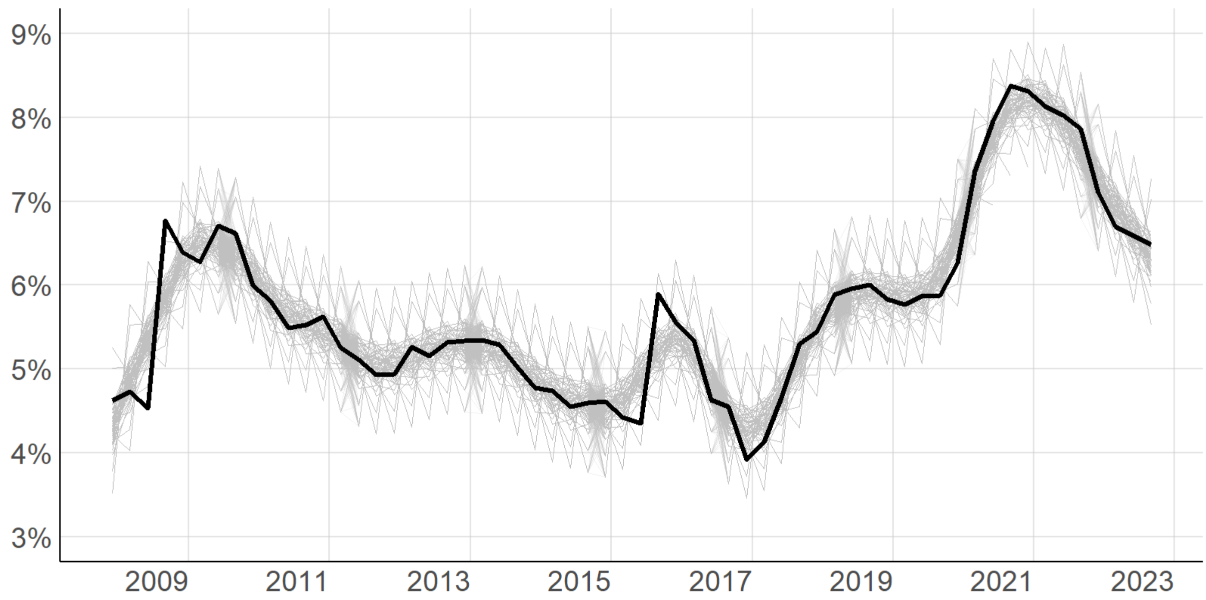
In the initial stage of implementing the BEVS framework, we first apply a Loess fit to the PD series using cross-validation to capture its underlying trends and specifically its low-frequency variations, as illustrated in Figure 3. Subsequently, we employ the circular block bootstrap technique to generate 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 4.

Figure 3: Loess fit of the probability of default.



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

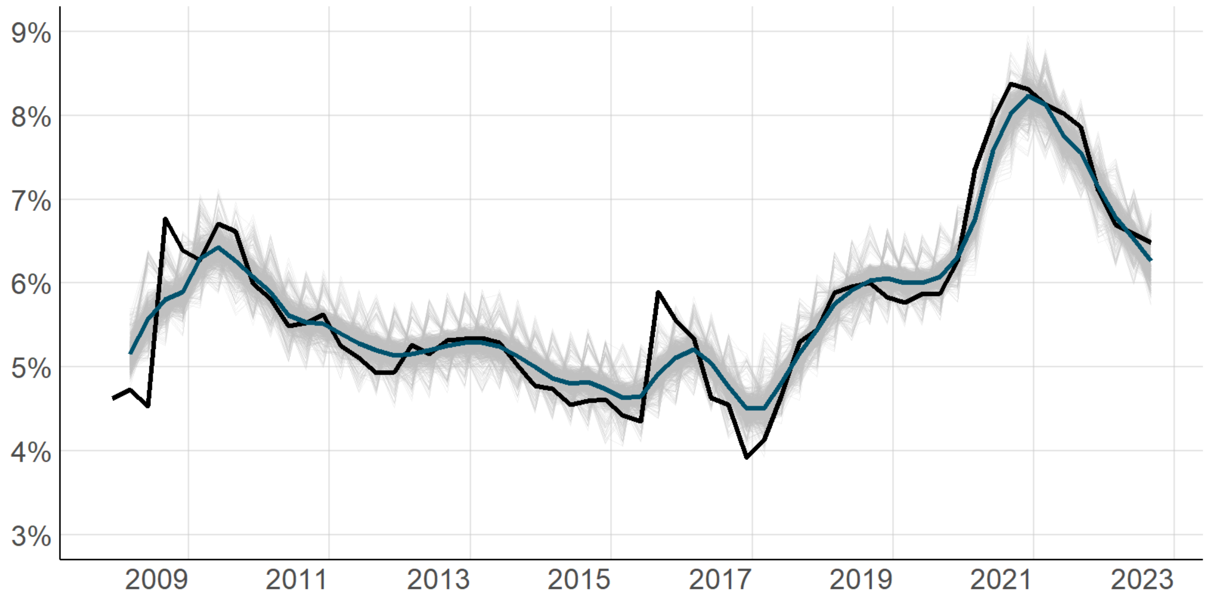
Figure 4: Bootstrapped series of the probability of default.



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

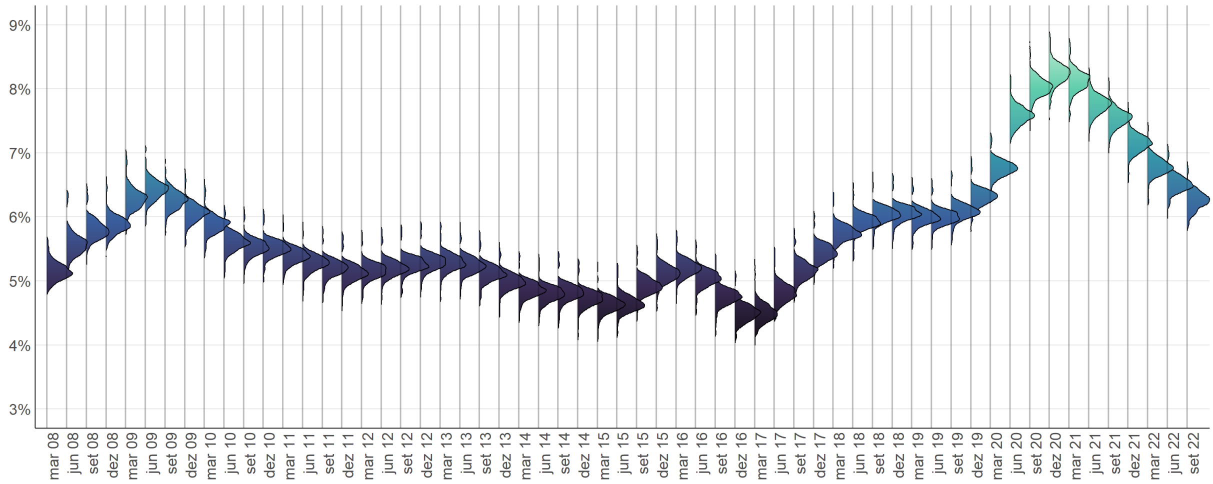
Upon completing the bootstrapping and Loess fitting phases, we construct an ensemble model named the 'BEVS model.' This model aggregates information from all individual Lasso models generated from the 1,000 bootstrapped series by averaging the non-zero coefficients, thus mitigating the uncertainty and risk associated with selecting a single model, as shown in Figure 5. The density distribution of these coefficients, presented in Figure 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 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 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 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. This comparative analysis underscores the performance advantages and statistical robustness achieved by the BEVS approach.

In Table 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 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 reinforcing that 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[16] and appears in 8.3% of all simulations, in accordance with its positive correlation of 22.1% as shown in Figure 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 became a regulatory requirement in Brazil in October 2018 (BCB, 2022), and all values

[16]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. .

before this date are constructed based on a proxy proposed by Takeuti (2020). 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 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. All figures of the dimensionality reduction process can be found in the Appendix 5.

Table 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: 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 2.5.

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 4 and 5, and the results of exercise (ii) are detailed in Tables 9 and 10 in Appendix 5.

In Tables 4 and 5, the BEVS model shows stability when varying the number of bootstrap simulations from 100 to 50,000. Table 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

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 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 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 9 and 10 presented in Appendix 5, 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[17] (Demirel and Willemain, 2002).

4.2 Theoretical Sign Restrictions

In the analysis presented in Section 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

[17]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 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 6.

Table 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 2.5.

The results of this restricted BEVS model, which incorporates the theoretical sign constraints as shown in Table 6, are summarized in Table 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 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 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 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 2.5.

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.

References

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

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

BCB (2022). *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.

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.

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

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

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.

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

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.

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.

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.

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.

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

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.

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.

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

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.

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

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

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

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

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.

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

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.

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.*

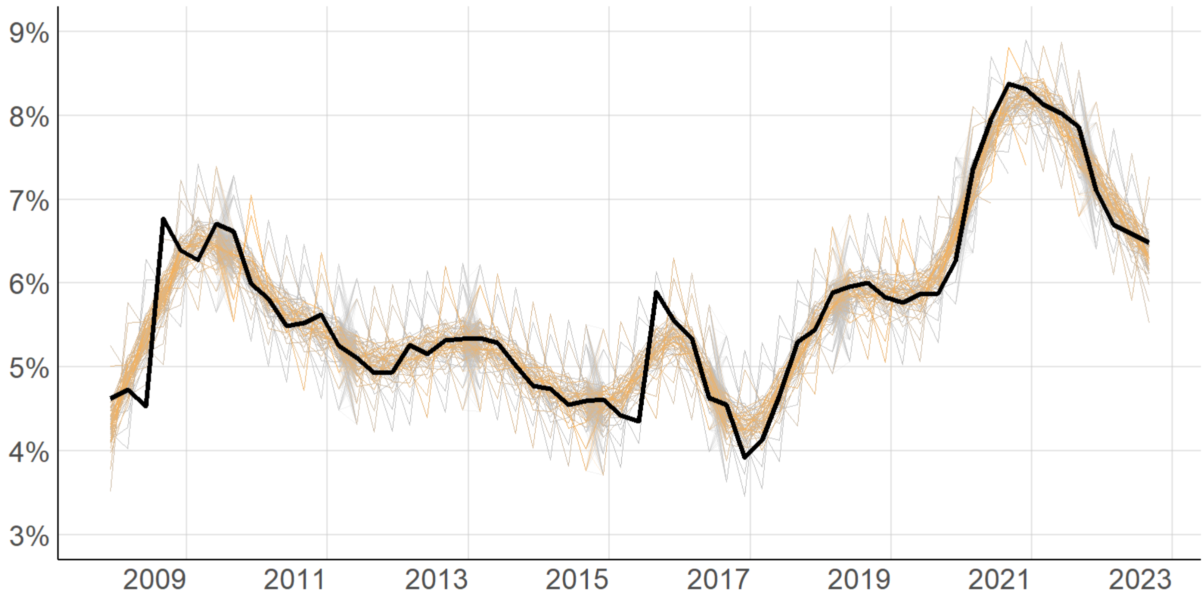
Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological) 58*(1), 267–288.

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.

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.

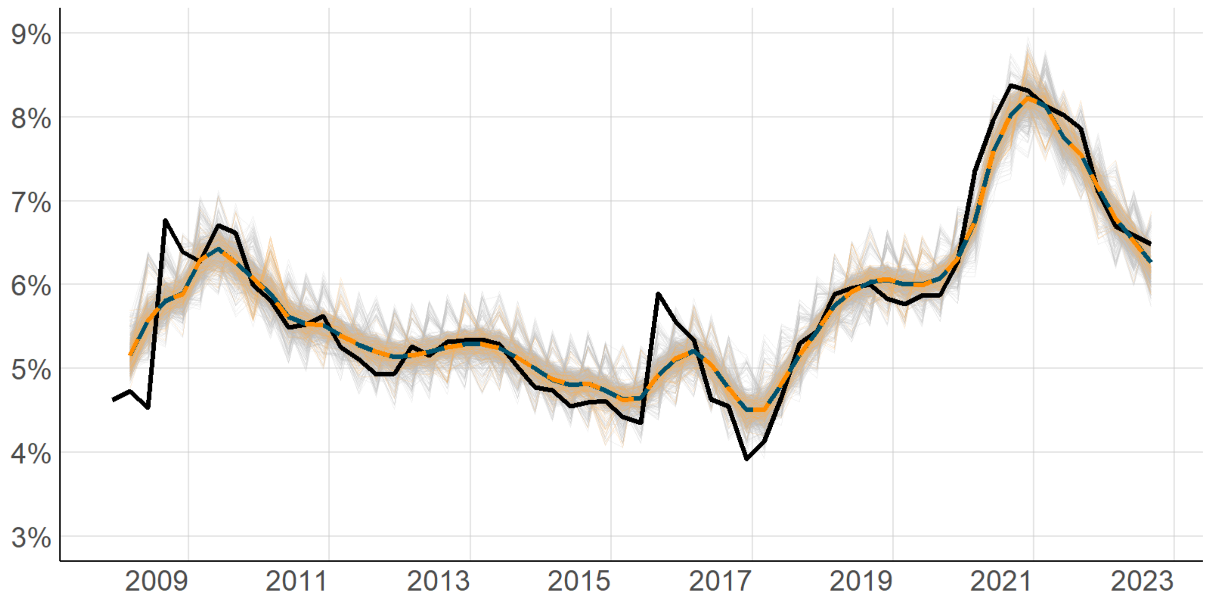
A Dimensionality Reduction Procedure

Figure 7: Bootstrapped series of the probability of default with dimensionality reduction.



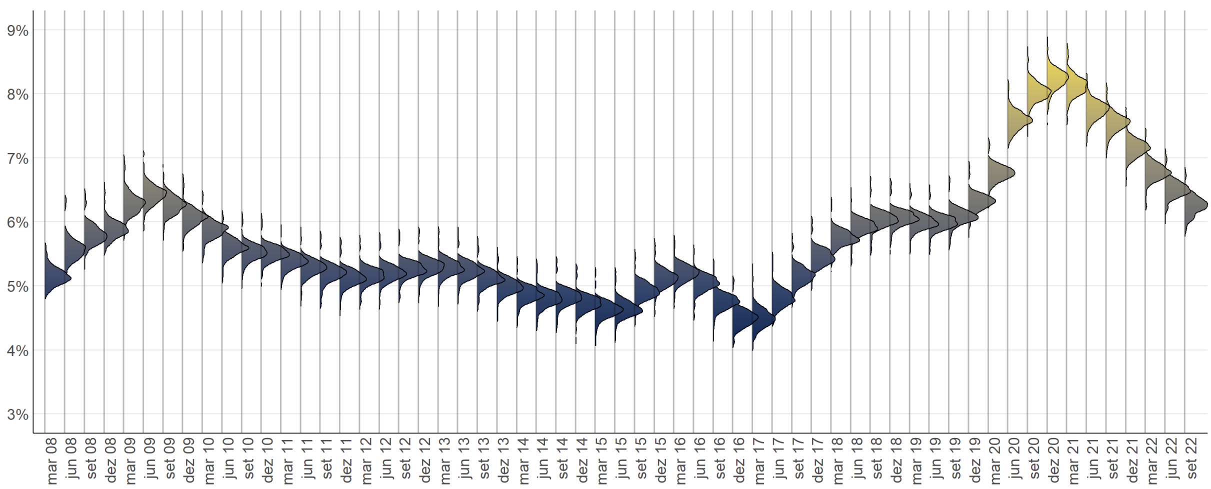
Notes: The lines in gray represent the bootstrap series of the PD in black. The lines in orange represents the removed bootstrapped series in the dimensionality reduction process.

Figure 8: 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 9: 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.

B Robustness Test

Table 9: 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 10: 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 Balance sheets accounts and granular data

Table 11: 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).