



Dynamic balance sheet simulation and credit default prediction: A stress test model for Colombian firms

Financial Stability Workshop – November 25, 2025

Diego Cuesta - Camilo Gómez



Motivation

- After the 2008 global financial crisis, stress test models became widely used to assess financial institutions' buffers in the presence of severe economic and financial shocks (Dent et al. 2016).
- The unprecedented COVID-19 crisis showed the importance of complementing stress tests in financial institutions with more analytical tools to assess the financial soundness of the corporate sector on a forward-looking basis.
- **Why is this important?**
 - Direct channel: **Firms' role in the credit market.**
 - Indirect channel: **Firms' impact on the broader economy.**
 - **Corporate financial information is less frequent** compared to bank balance sheets, requiring enhanced monitoring tools.

Study purpose

This paper presents a stress test model for Colombian firms, the analytical tool employed by the Financial Stability Department of the Banco de la Republica (Central Bank of Colombia) to perform its financial vulnerability analysis of the Colombian non-financial corporate sector.

The model proposed in this paper has three building blocks.

- 1. Dynamic balance sheet simulation framework based on accounting behavioral rules and econometric analysis.**
 - This block simulates the main accounts of firms' balance sheets conditional on the firms' initial characteristics and a macroeconomic scenario (GDP growth, financial conditions, level of interest rates). [See literature](#)
- 2. Battery of machine learning models (ML) that allow us to predict the credit default probability of firms based on key financial and activity indicators.**
- 3. The third block combines the first two to define firms at credit risk.**
 - Input for banks' stress test models used in the Central Bank of Colombia (see Gamba et al. 2017).

We use data from three sources:

1. **Annual financial information of firms that report their financial statements to the Colombian Superintendence of Companies and Financial Superintendence.** [See descriptives](#)
 - Information from 1999 to 2023 with 517,850 observations, and 66,166 firms.
 - This set of information is the main input for the dynamic balance sheet simulation model.
2. **Colombian credit registry reported by credit institutions to the Colombian Financial Superintendence.**
 - This dataset provides information about firms' delinquency days each year from 2005.
 - We merge the firm's annual data with the Colombian credit registry and used the merged data to train the set of ML models proposed in this paper.

For the whole period, the proportion of firms in default relative to the total sample is **13.4%**, indicating a **highly imbalanced dataset**. [See descriptives](#)

3. **Macroeconomic aggregates produced by the National Statistics Office (DANE).**

Block 1: Dynamic balance sheet simulation framework



Dynamic balance sheet simulation framework

- A regression analysis is first conducted to **link macroeconomic aggregates to key financial indicators**.
- Based on these empirical relationships and using the annual panel dataset described above, we then construct **an accounting-consistent framework to simulate the main financial accounts and P&L items of firms**, following Tressel and Ding (2021)..
- We illustrate dynamic balance sheet simulation results based on the stressed macroeconomic scenario presented in Banco de la República's Financial Stability Report from the second half of 2024.

- **The regression analysis aims to capture the statistical relationship between key macroeconomic and financial aggregates and firms' financial indicators** (input for accounting rules).
 - **Modeled variables:** sales growth ($\Delta \ln \text{Sales}$), financial leverage ($\text{FinancialDebt}/\text{Assets}$).
 - **Macroeconomic variables:** real annual GDP growth, credit boom indicator.

- **Formally, following Tressel and Ding (2022), the OLS dynamic regression has the general form:**

$$Y_{it} = \alpha \cdot Y_{it-1} + \delta \cdot \text{CharFirm}_{it-1} + \Psi_s^{\text{Macro}} \cdot \text{Macro}_t + \Psi^{\text{Fin}} \cdot \text{Fin}_t + d_s + v_{it}, \quad (10)$$

Where: Y_{it} = log change in sales or financial leverage, CharFirm_{it-1} = vector of firms' characteristics (ROA, financial leverage, log of assets, sales-to-assets ratio, log change in sales), Macro_t = set of macroeconomic variables, Fin_t = credit boom indicator, d_s = sector fixed effects, v_{it} = error term.

- **We allow the coefficients of the macroeconomic variables Ψ_s^{Macro} to vary depending on the economic sector to capture heterogeneities in the relationship between sectors and the aggregate economic cycle.**

Finally, the following regression is used to measure the elasticity of costs with respect to sales at the firm-sector level:

$$\Delta \ln \text{Costs}_{it} = \alpha \cdot \Delta \ln \text{Costs}_{it-1} + \delta \cdot \text{CharFirm}_{it-1} + \Psi_s \cdot \Delta \ln \text{Sales}_{it} + v_{it}. \quad (11)$$

Where: CharFirm_{it-1} = vector of firms' characteristics (ROA, financial leverage, log of assets, sales-to-assets ratio, log change in sales), d_s = sector fixed effects, v_{it} = error term.

[See discussion about model assumptions](#)

Dynamic balance sheet simulation framework: Accounting consistent behavioral rules

- Define $h = 0, \dots, H$ as the balance sheet simulation periods, where $h = 0$ refers to the initial point and H to the final point. Index firm by i . Financial costs are modeled with the following equation:

$$\text{FinancialCosts}_{ih} = \text{FinancialDebt}_{ih-1} \cdot [i_{i0}^{eff} + a \cdot \Delta \text{MPR}_h], \quad (1)$$

- a = share of variable interest rate commercial loans ($\approx 80\%$).
 - MPR_h = monetary policy rate in the macroeconomic scenario.
 - i_{i0}^{eff} = initial effective rate of financial debt, measured as financial costs over financial liabilities.
- With equation (1) and given values of sales and costs (input from regression analysis explained later), profits before taxes can be computed and be used in tax calculation according to:

$$\text{Taxes}_{ih} = \text{ProfitsBeforeTaxes}_{ih} \cdot \tau \cdot 1[\text{ProfitsBeforeTaxes}_{ih} > 0],$$

- τ = statutory corporate income tax rate.
- $1[\cdot]$ = indicator function.

Dynamic balance sheet simulation framework: Accounting consistent behavioral rules

- Assuming that dividends are not distributed, profits and equity are defined as:

$$\text{Profits}_{ih} = \text{ProfitsBeforeTaxes}_{ih} - \text{Taxes}_{ih}, \quad (3)$$

$$\text{Equity}_{ih} = \text{Equity}_{ih-1} + \text{Profits}_{ih}. \quad (4)$$

- Following Tressel and Ding (2021), we define the initial net cash of a firm as:

$$\begin{aligned} \text{Cash}_{i0} = & \text{Cash\&Equivalents}_{i0} + \text{Short-TermInvestments}_{i0} + \text{AccountReceivables}_{i0} \\ & - (\text{Short-TermAccruedPayrolls}_{i0} + \text{AccountPayables}_{i0} + \\ & \text{OtherShort-TermNon-FinancialLiabilites}_{i0}). \end{aligned} \quad (5)$$

- With this definition, profits are accumulated in net cash. Moreover, if cash needs arise, financial debt increases. The above is summarized in the following expressions:

$$\text{Cash}_{ih} = \text{Cash}_{ih-1} + \text{Profits}_{ih}, \quad (6)$$

$$\text{FinancialDebt}_{ih} = \text{FinancialDebt}_{ih-1} - \text{Cash}_{ih} \cdot 1[\text{Cash}_{ih} < 0] \quad (7)$$

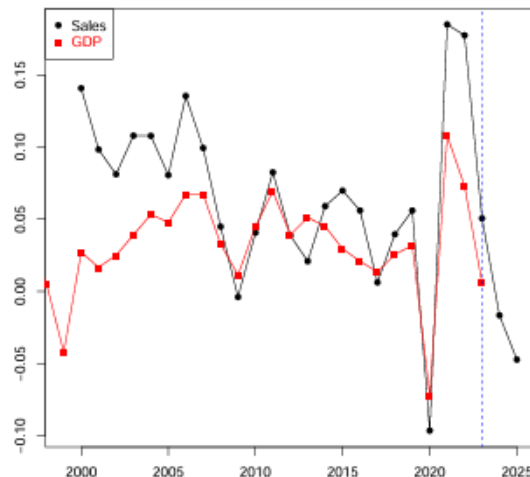
- Finally, liabilities and assets are given by:

$$\text{Liabilities}_{ih} = \text{Liabilities}_{ih-1} + \Delta \text{FinancialDebt}_{ih}, \quad (8)$$

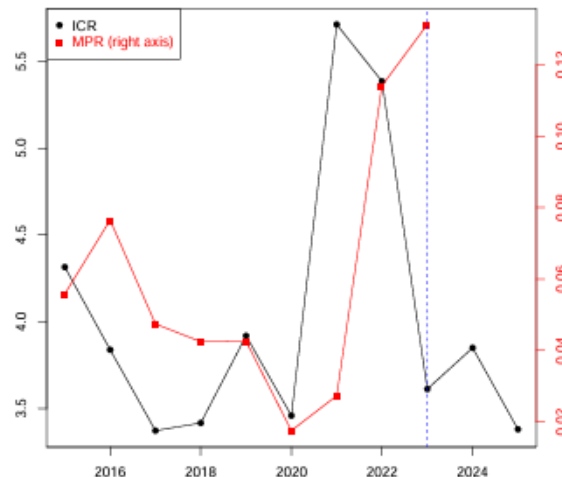
$$\text{Assets}_{ih} = \text{Liabilities}_{ih} + \text{Equity}_{ih}. \quad (9)$$

Dynamic balance sheet simulation framework: Results from FSR 2024-2

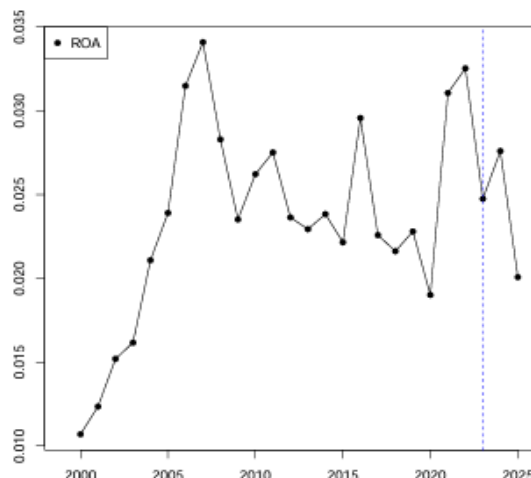
Figure 4: Median simulated values of key financial variables under the stressed macroeconomic scenario



(a) Log change of sales and GDP growth



(b) Interest coverage ratio (ICR)



(c) Return on assets (ROA)



(d) Financial leverage

- **Firms would experience financial pressures during the stress horizon.**
 - The median firm would exhibit sales contractions.
 - In line with this sales drop and the upward pressures on the monetary policy interest rate, the interest coverage ratio (ratio of earnings before taxes to interest expense - ICR) would decrease.
 - ROA and financial leverage would decrease and increase, respectively.
- [See out-of-sample comparison exercise](#) (simulations starting from 2021 through periods 2022 and 2023 are compared with observed data)
 - The model tends to correctly estimate the distribution of operational profits and ROA.
 - However, the model seems to overestimate financial leverage ratios and financial obligations and costs.
 - The proportion of firms with an ICR lower than one also tends to be overestimated.

Block 2: Machine learning models for credit default prediction



Machine learning models for credit default prediction: Classification problem

The objective is to identify whether a firm is in credit default ($Y_{it} = 1$) or not ($Y_{it} = 0$), based on its financial variables, that are a set of lagged firm-level financial indicators (X_{it-1}), the lagged default variable (to capture default persistence) and a set of firm sectoral dummies (S_{it}):

$$\widehat{Y}_{it} = f(X_{it-1}, Y_{it-1}, S_{it})$$

The classification process embedded in f involves two interconnected steps:

1) Estimation of the conditional probability:

$$P(Y_{it} = 1 | X_{it-1}, Y_{it-1}, S_{it})$$

2) Classification of each observation based on its conditional probability and the threshold T :

$$\widehat{Y}_{it} = \begin{cases} 1 & \text{if } P(Y_{it} = 1 | X_{it-1}, Y_{it-1}, S_{it}) \geq T \\ 0 & \text{if } P(Y_{it} = 1 | X_{it-1}, Y_{it-1}, S_{it}) < T \end{cases}$$

Machine learning models for credit default prediction: Confusion matrix

To evaluate the performance of a classification model like the one described above, the confusion matrix is used. This tool compares the model's predictions with the actual observed values in a dataset.

Table 2: Confusion matrix

		Forecast	
		Negatives ($\hat{Y}_{it} = 0$)	Positives ($\hat{Y}_{it} = 1$)
Observed	Negatives ($Y_{it} = 0$)	True negatives (TN)	False positives (FP)
	Positives ($Y_{it} = 1$)	False negatives (FN)	True positives (TP)

Machine learning models for credit default prediction: Performance metrics

Based on the confusion matrix, we can calculate metrics that are usually used in classification models: *Accuracy*, *precision* and *recall*.

In contexts where there is a minority class, *accuracy* may not be a good measure of model performance. This is particularly relevant in our case of credit default, as we face a highly imbalanced class problem (13,4%).

$$\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{Total Predictions}} = \frac{TP + TN}{TP + TN + FN + FP},$$

$$\text{Precision} = \frac{\text{Correct Positive Predictions}}{\text{Total Positive Predictions}} = \frac{TP}{TP + FP},$$

$$\text{Recall} = \frac{\text{Correct Positive Predictions}}{\text{Total Positives}} = \frac{TP}{TP + FN}.$$

We are interested in *precision* and *recall* metrics. In each proposed model, **our primary objective is to maximize the F_2 - score**, which combines precision and recall while giving greater weight to *recall*.

$$F_2 = (1 + 2^2) \frac{\text{Precision} \cdot \text{Recall}}{2^2 \cdot \text{Precision} + \text{Recall}} = \frac{(1 + 2^2) \cdot TP}{(1 + 2^2)TP + 2^2 \cdot FN + FN}.$$

To address the issue of class imbalance, there are several complementary solutions that researchers use. In this research we implement three main strategies:

1. **Sampling methods:** We use the **Synthetic Minority Over-Sampling Technique (SMOTE)** to **oversample the minority class** with the objective of having a balanced **training** sample.
2. **Hyperparameters tuning:** An approach to improve model performance is to choose the **hyperparameters that optimize $F_2 - score$** through a **five-fold cross-validation procedure**.
3. **Classification threshold tuning:** By default, the threshold is set at 0.5; **however, it can be adjusted to enhance the model's performance on the minority class and, consequently, improve the $F_2 - score$** . That tuning is performed using a data set independent of the training and test sets: the evaluation set.

Machine learning models for credit default prediction: Data partitioning strategy

To address the classification problem using different models aimed at maximizing the $F_2 - score$ metric, it is essential to randomly **divide the sample into three mutually exclusive subsets**. This partitioning allows for model evaluation at different stages, ensuring its **generalization ability and reducing the risk of over-fitting** the training data. The original sample is divided as follows:

1. **Training set (75% of the data):** This subset is used to estimate the model parameters after applying the SMOTE technique. With this set, optimal hyperparameters through a 5-fold cross-validation procedure are found with the goal of maximizing the $F_2 - score$ metric.
2. **Evaluation set (10% of the data):** This subset is used to determine the optimal decision threshold that maximizes the $F_2 - score$ metric after model training with the training set.
3. **Test set (15% of the data):** This subset is used to compare the performance of different models based on the selected performance metric using data that was not used during training or evaluation. This ensures a fair comparison across models based on the $F_2 - score$

Machine learning models for credit default prediction: Set of machine learning models

Two family of models are estimated: **logistic regression models** and **classification tree-based methods**.

Table 4: Description of ML models used to predict default

Model	Features used to predict default	Hyperparameters grid	Selected hyperparameters
Logit 1	Lagged default Sector dummies Lagged financial variables from Table 3		
Logit 2	Variables from Logit 1 Squared terms of lagged financial variables from Table 3		
Logit 3	Variables from Logit 2 Interactions between lagged financial variables from Table 3		
Logit Lasso	Variables from Logit 3	$\lambda \in \{0, 0.00001, 0.00002, \dots, 0.0001\} \cup \{0.0001, 0.0003, 0.0005, \dots, 0.001\} \cup \{0.001, 0.002, 0.003, \dots, 1\}$	0.00001
Logit Ridge	Variables from Logit 3	$\lambda \in \{0, 0.00001, 0.00002, \dots, 0.0001\} \cup \{0.0001, 0.0003, 0.0005, \dots, 0.001\} \cup \{0.001, 0.002, 0.003, \dots, 1\} \cup \{1, 1.5, 2, \dots, 100\}$	0.684

Model	Features used to predict default	Hyperparameters grid	Selected hyperparameters
RF	Lagged default Sector dummies Lagged financial variables from Table 3	$\text{min.node.size} \in \{5, 10, 15\}$	$\text{min.node.size} = 10$
XGBoost	Lagged default Sector dummies Lagged financial variables from Table 3	$\text{max_depth} \in \{2, 4, 6\}$ $\text{gamma} \in \{0, 0.025, 0.005, 0.0075, 0.01, 0.015\}$ $\text{min_child_weight} \in \{5, 10\}$	$\text{max_depth} = 4$ $\text{gamma} = 0.01$ $\text{min_child_weight} = 10$

Source: Authors' elaboration based on data from the Colombian Superintendence of Companies and Financial Superintendence.

Machine learning models for credit default prediction: Results

The XGBoost model achieves the best performance according to the F_2 – score metric with the threshold of 0.5 and with tuned threshold

Table 5: Out-of-sample model performance

Model	Threshold - 0.5				Tuned threshold				
	Accuracy	Recall	Precision	F_2 -score	Threshold	Accuracy	Recall	Precision	F_2 -score
Logit 1	0.800	0.612	0.355	0.535	0.480	0.783	0.641	0.337	0.543
Logit 2	0.783	0.635	0.336	0.539	0.460	0.740	0.688	0.297	0.545
Logit 3	0.832	0.531	0.403	0.499	0.360	0.704	0.710	0.270	0.535
Logit Lasso	0.835	0.555	0.413	0.520	0.390	0.733	0.692	0.290	0.542
Logit Ridge	0.715	0.675	0.272	0.521	0.490	0.691	0.700	0.258	0.522
RF	0.739	0.674	0.293	0.535	0.500	0.739	0.674	0.293	0.535
XGBoost	0.778	0.649	0.331	0.544	0.460	0.740	0.694	0.298	0.548

Source: Authors' elaboration based on data from the Colombian Superintendence of Companies and Financial Superintendence.

Machine learning models for credit default prediction: Results of XGBoost model

With the **untuned** threshold, the proportion of TP rate for the XGBoost model is 64.9%, while the TN rate reaches 79.8%. With the **tuned** threshold, the TP rate improves to 69.4%. However, this adjustment also leads to a reduction in the TN rate.

Figure 5: Confusion Matrix of XGBoost

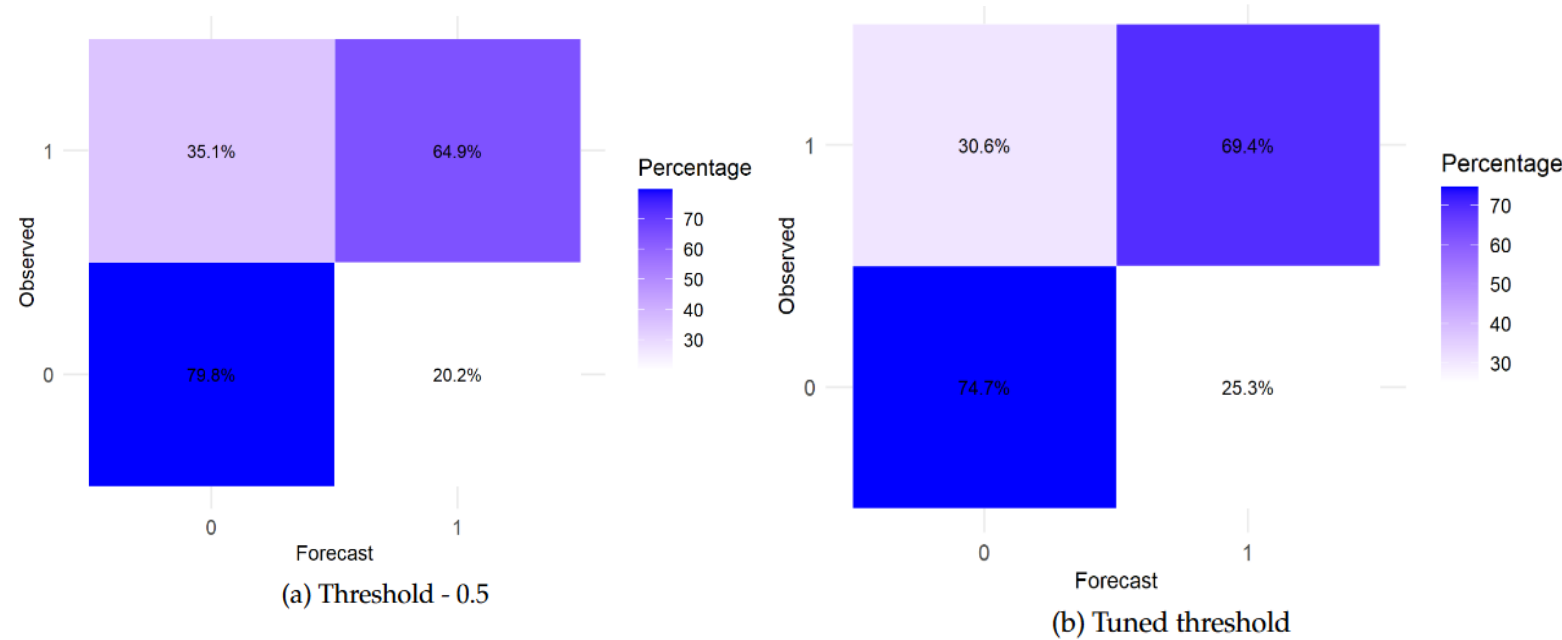
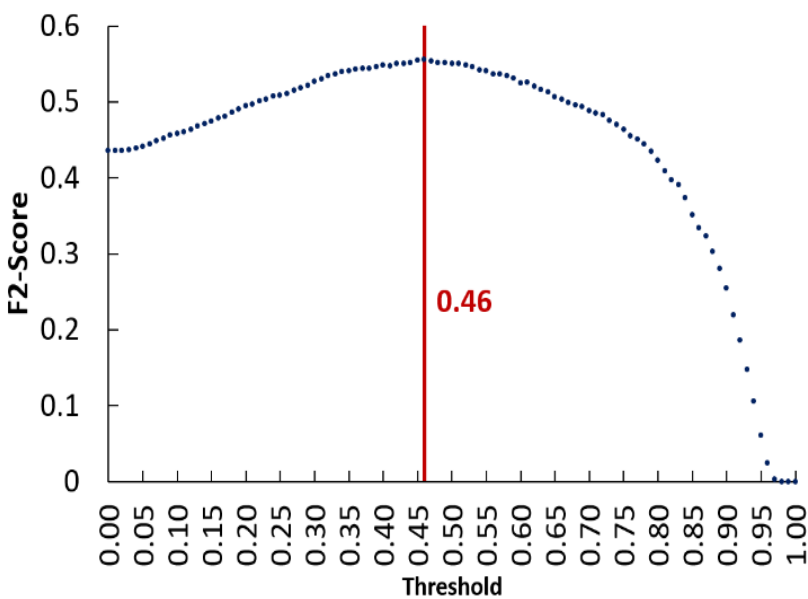


Figure 6: F_2 -score in the evaluation subsample vs. threshold



Block 3: Classification under a stressed scenario



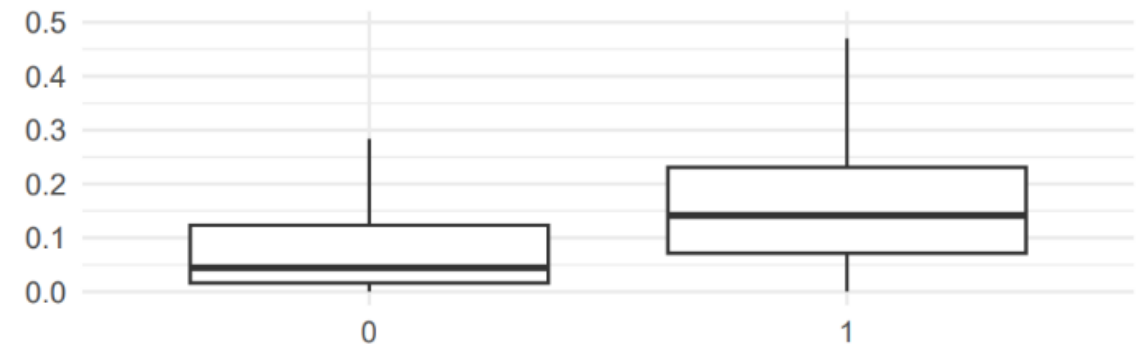
Classification under a stressed scenario

This section integrates the results of block 1 and 2: Specifically, it utilizes the financial indicators of firms simulated under the methodology described in block 1 and applies the ML models presented in block 2 to identify firms that may be classified as in default under the given stress scenario (REF 2024-II).

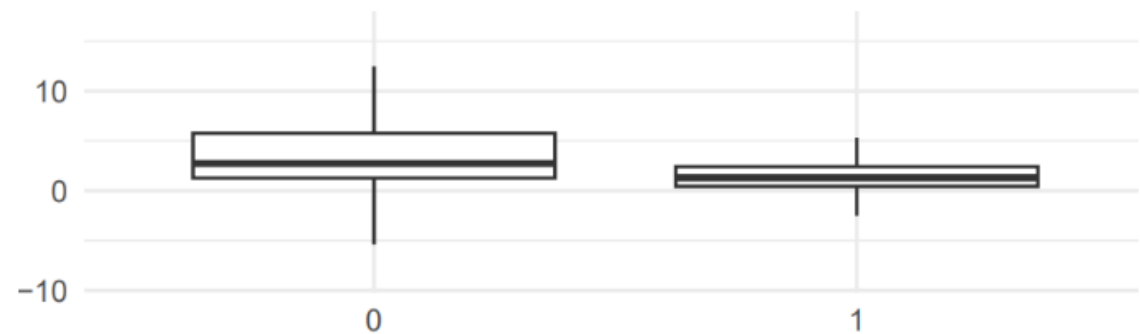
The results presented correspond to those obtained for the year 2025 using the XG-Boost model.

- Firms classified as in **default** generally exhibit **higher financial leverage** (ratio of financial obligations to total assets) indicator in the preceding period.
- Firms classified as in **default** generally exhibit **lower coverage ratios** (ratio of earnings before taxes to interest expense) indicator in the preceding period.

Figure 7: Financial indicators for firms classified in default (1) and not in default (0)



(a) Financial leverage



(b) Interest coverage ratio

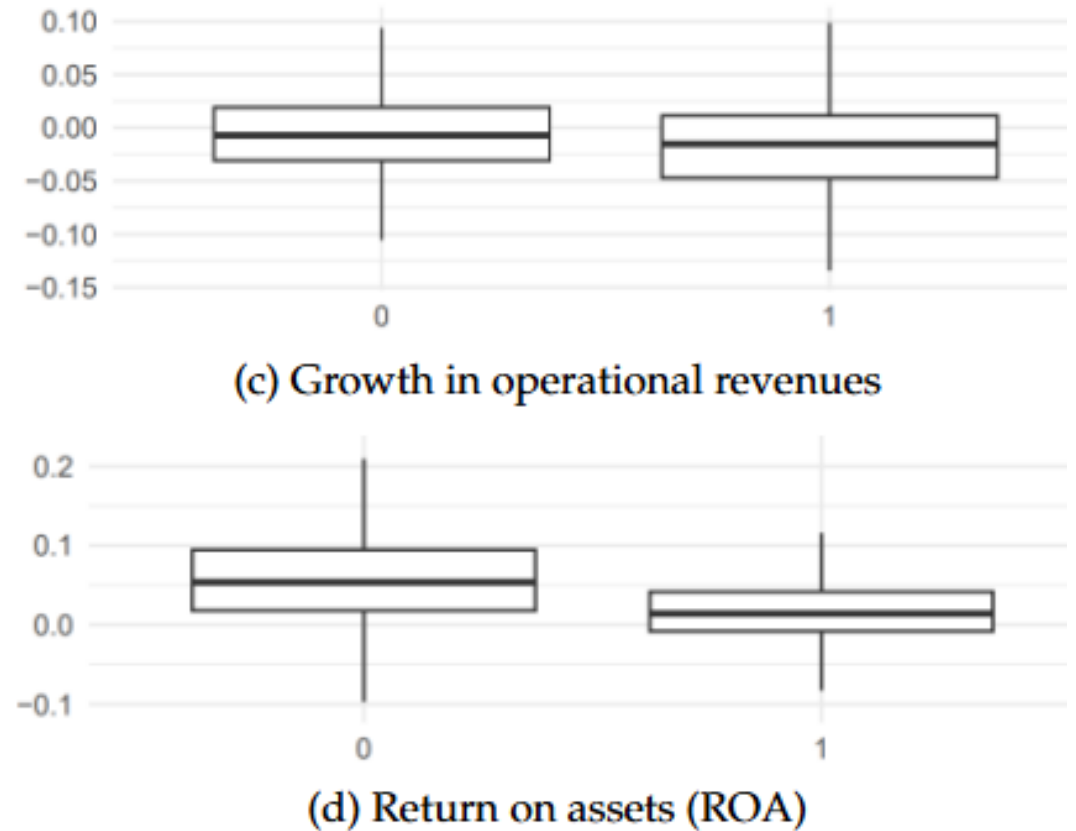
Classification under a stressed scenario

The results presented correspond to those obtained for the year 2025 using the XG-Boost model.

- Firms classified as in **default** generally exhibit **lower growth in operational revenues** in the previous period.
- Firms classified as in **default** generally exhibit **lower ROA** in the preceding period.

These results should be interpreted as illustrative and should be **analyzed holistically, in conjunction with all variables**. Therefore, a firm with a negative or near-zero ROA is not necessarily classified as in default. This is because, **despite potentially having low profitability** due to the nature of its business, **the firm may simultaneously exhibit robust financial leverage or interest coverage indicators**.

Figure 7: Financial indicators for firms classified in default (1) and not in default (0)



Thank you!



Appendix



Related literature

- **Stress testing was first used in engineering to evaluate stability under adverse conditions (Borio, 2014; Dent & Westwood, 2016).**
- **The unprecedented COVID-19 crisis showed the importance of complementing stress tests in financial institutions with more analytical tools to assess the financial soundness of the corporate sector on a forward-looking basis.**
- **Several studies proposed analytical stress testing frameworks for the corporate sector during the pandemic to simulate financial variables in a consistent way.**
 - **Models can depend on exogenous income shocks -e.g., at the sectoral level- (Carletti et al., 2020; Demmou et al., 2021) or can be complemented with firm-level regressions that relate macroeconomic variables to key financial and activity performance variables such as sales growth or leverage (Caceres et al. 2020; Tressel and Ding, 2021).**
- **We closely follow the most comprehensive approach of Tressel and Ding (2021).**
- **This paper is also related to the corporate finance literature deriving the main drivers of firms' default (e.g., Altman, 1968; Bottazzi et al., 2011; Traczynski, 2017; Cathcart et al., 2020, Modina et al., 2023).**
- **For predicting default, we compare simpler models with a set of data-driven, ML approaches that use a wide set of covariates and specifications.**

Table 1: Firm financial variables

Variables	Notes
Sector	
Total Assets	
Cash and equivalents	
Short-term financial liabilities	
Long-term financial liabilities	
Total liabilities	
Equity	
Operating income or sales	
Operating expense or costs	
Other operating income	
Other operating profits or losses	
Profit from operating activities	
Financial income	
Financial costs	Interest expenses before 2015
Profit before taxes	
Taxes	
Total profits	
Trade and other current receivables	
Other current financial assets	Short-term investments in cash definition (equation 5).
Current provisions for employee	Short-Term accrued payrolls in cash definition (equation 5)
Trade and other current payables	
Other current non-financial liabilities	
Other non-current financial assets	
Issued capital	

Notes: Balance and P&L accounts taken from annual financial information of firms.
Source: Authors' elaboration based on data from the Colombian Superintendence of Companies and Financial Superintendence.

Data: Firms default

- The merged dataset used for the set of ML models to predict default one period ahead includes lagged financial information from various indicators. **The dataset employed in the models begins in 2006 to account for the one-period lag.**
- For the whole period, the proportion of firms in default relative to the total sample is **13.4%**, indicating a **highly imbalanced dataset**.

Figure 3: Percentage of firms in default by year



Notes: Yearly number of firms in default (past-due days higher than 30) as a percentage of firms. Source: Authors' elaboration based on data from the Colombian Superintendence of Companies and Financial Superintendence.

Dynamic balance sheet simulation framework: Discussion

- **The dynamic balance sheet simulation model offers a micro-macro consistent tool to evaluate the firms' exposure to the risks identified in a macroeconomic scenario and the most recent exposure of the financial statements to these firms.**
- **Moreover, the accounting framework is flexible enough to evaluate different sensibility scenarios (e.g., sales drop of $x\%$).**
- **However, the model is based on some restrictive accounting behavioral rules and correlations observed in the data -> results must be carefully read.**
 - Strategic behaviors not taken into account: prepaying debt, reducing size, renegotiating debt, etc.
 - When we present the results, we discuss, based on an out-of-sample comparison exercise, how the mentioned assumptions can affect results.

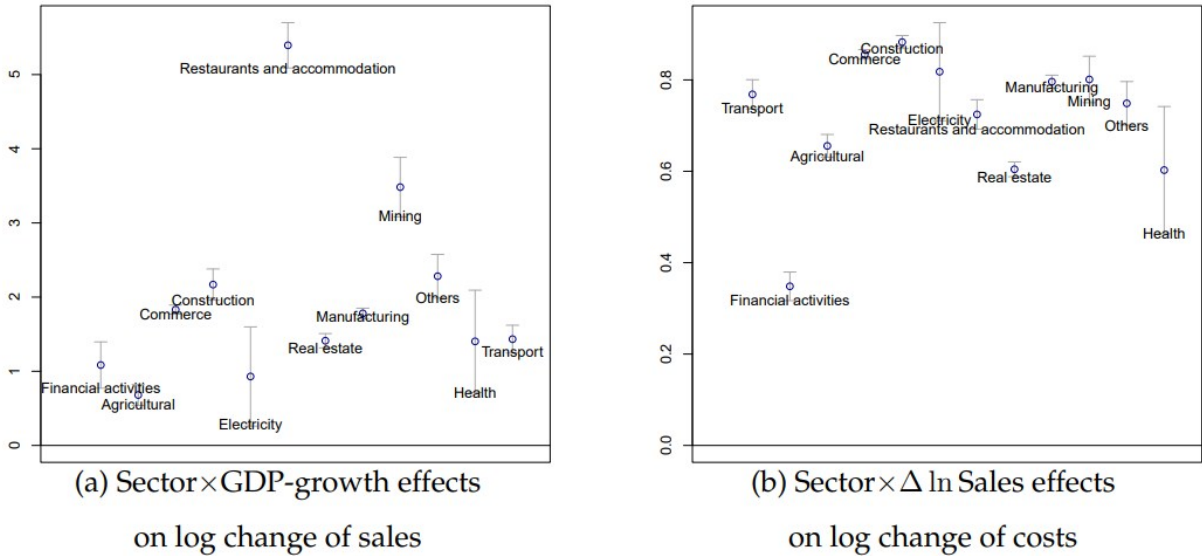
Dynamic balance sheet simulation framework: Regression analysis results. Firms with larger assets, higher profitability, and higher financial leverage and sales-to-assets ratio tend to grow more (Column 1). Sales growth and ROA correlate negatively with financial leverage (Column 2). Sales growth and sales-to-assets rate correlate negatively with the log change of costs (Column 3).

Table A1: Regression analysis

VARIABLES	(1) Log change of sales	(2) Financial leverage	(3) Log change of costs
Lagged dependent variable	-0.07 *** (0.00)	0.73 *** (0.00)	-0.49 *** (0.01)
<i>Lagged firm characteristics</i>			
Log change of sales		-0.003 *** (0.00)	0.38 *** (0.01)
Financial leverage	0.07 *** (0.00)		-0.02 * (0.01)
Sales-to-assets ratio	0.003 *** (0.00)	0.005 *** (0.00)	-0.02 *** (0.00)
ROA	0.08 *** (0.01)	-0.05 *** (0.00)	0.29 *** (0.02)
Log of assets	0.02 *** (0.00)	0.002 *** (0.00)	0.001 *** (0.00)
Credit boom indicator	0.02 *** (0.00)	0.002 *** (0.00)	
Sector fixed effects	Yes	Yes	No
Sector x GDP growth	Yes	No	No
Sector x Log change of sales	No	No	Yes
N	301,689	304,609	299,135
R2	0.03	0.59	0.31

Regression results from equation (10) and (11). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

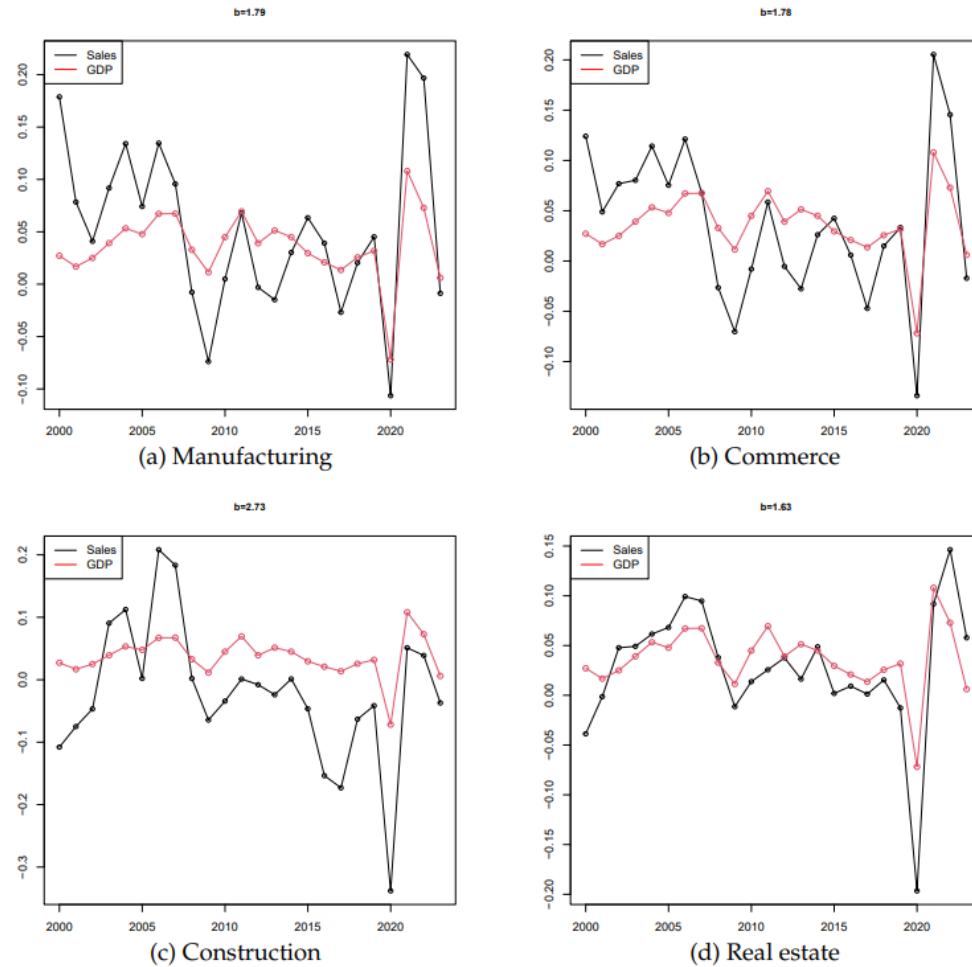
Figure A1: Effects of GDP growth and sales growth accross sectors



Authors’ calculations of Sector x GDP-growth effects (Panel a, Ψ_s^{Macro} in equation 10) and Sector x $\Delta \ln$ Sales effects (Panel b, Ψ_s in equation 11). Robust confidence intervals calculated at the 10% significance level.

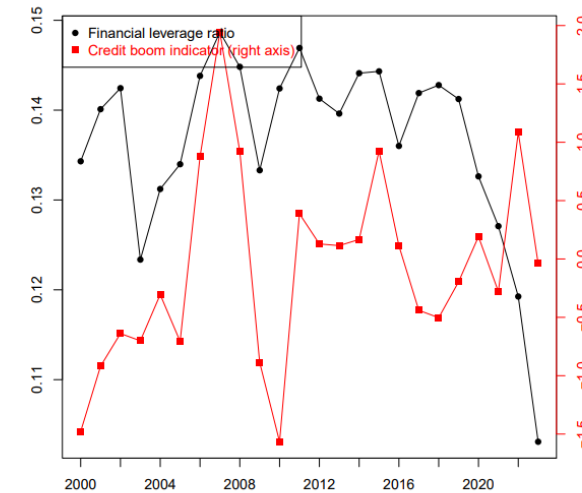
Data: There is a heterogeneous relationship between aggregate GDP growth and average sales growth by sector, as well as a (though weaker) relationship between financial leverage and the credit boom indicator.

Figure 1: Average sales growth for selected sectors and aggregate GDP growth



Notes: Average sales growth for selected sectors and aggregate GDP growth. b is the estimated coefficient, for each sector s , of regression $\Delta \ln \text{Sales}_{st} = a_s + b_s \cdot \text{GDP}_t + v_{st}$, where $\Delta \ln \text{Sales}_{st}$ refers to average log change in sales in sector s and year t , and GDP_t to aggregate GDP growth, and v_{st} to the error term. Source: Authors' elaboration based on data from the Colombian Superintendence of Companies and Financial Superintendence.

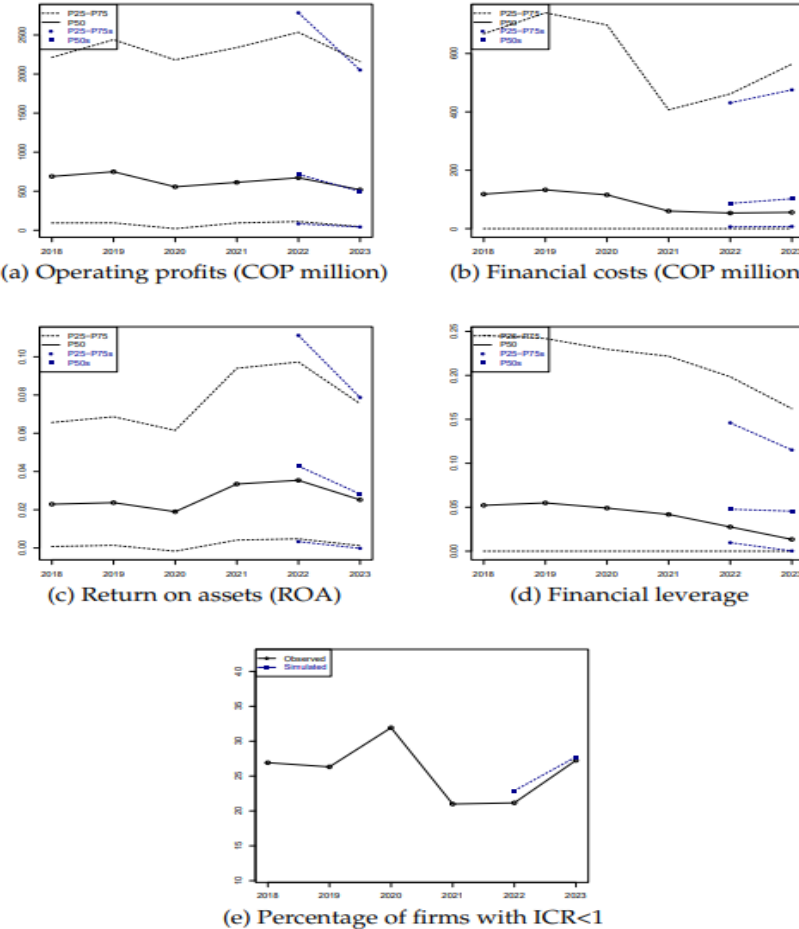
Figure 2: Average financial leverage and aggregate credit boom indicator



Notes: Yearly average of firms' financial leverage and aggregate credit boom indicator. Source: Authors' elaboration based on data from the Colombian Superintendence of Companies and Financial Superintendence.

Dynamic balance sheet simulation framework: Out-of-sample performance of the dynamic balance sheet simulation.

Figure B1: Out-of-sample results of the dynamic balance sheet simulation (2023-2024)



Observed (black lines) and balance sheet simulation (blue lines) of key firms' financial variables based on observed macroeconomic data for 2022 and 2023. Px refers to the x percentile of the corresponding variable.