Default Risk in Italian Municipalities: A Predictive Model for Early Warning

Gabriele Marras

July 22, 2025

Abstract

According to the Consolidated Law on Local Authorities (TUEL), municipal administrations that are no longer able to meet their financial obligations enter a procedure known as financial distress. The analysis presented here focuses on the third group and aims to identify relationships between a municipality having experienced financial distress and some of the main budgetary indicators. For these indicators, two time periods were considered: one prior to the pandemic (2019) and the most recent available (2022)

1 Introduction

A sound and prudent management of public finances by municipal institutions is a necessary condition for properly administering a territory, ensuring services to citizens, and safeguarding the sustainability of public spending at higher levels of government. According to Article 244 of the Consolidated Law on Local Authorities (TUEL), local governments can experience three types of financial distress: budgetary deficits, pre-distress (or multi-year financial recovery plans), and full financial distress. This study focuses on the latter case, understood as the procedure activated when a local authority is no longer able to perform its functions or provide essential services, and struggles to meet its debt obligations. The analysis draws on a study conducted in 2018 by the Observatory on Local Government Finance and Accounting of the Ministry of the Interior. That study examines the main causes of financial difficulties in Italian local authorities and the scale of the phenomenon, aiming to identify the most significant indicators associated with financial distress. The main sources for the budgetary indicators are the Ministry of the Interior and Istat. For ease of data collection, the latter was used. According to the aforementioned study by the Observatory, the main causes of financial distress in a local authority fall into two categories: internal and external. Internal causes relate to structural budget rigidity, poor management of financial flows, and low capacity for revenue collection and debt repayment. External causes concern the relationship with higher levels of government, with attempts to measure factors such as the degree of dependence on transfers and the level of fiscal autonomy. After cleaning and reorganizing the data, the dataframe was enriched with information on the population of each municipality and its geographical location by region and territorial area (Northwest, South, etc.), both available from ISTAT. Finally, for each of the two years, the main dataset was supplemented with data indicating which municipalities experienced financial default. The database from which the defaults were extracted is available on the website of the Ca' Foscari University Foundation in Venice and includes data from 1989 through 2025.

ISTAT indicators used in the analysis, broken down by category

Category	Indicators
Revenue	Grado di autonomia impositiva, Capacità di riscossione, Trasferimenti correnti / spese correnti, Trasferimenti in c/capitale / spese c/capitale
Expenditure	Rigidità della spesa, Spese interessi passivi / entrate correnti, Spese per rimborso prestiti / entrate correnti, Incidenza redditi da lavoro, Capacità di spesa
Receivables/Payables	Indice accumulo residui attivi, Smaltimento residui attivi, Accumulazione residui passivi, Smaltimento residui passivi
Personnel	Turnover risorse umane
Other	Avanzo/disavanzo su entrate correnti

The unit of observation is the individual Italian municipality, with a total of over 7,900 observations per year. Below is the complete list of Istat indicators that best describe the economic and financial condition of a local authority, although for some of them data are not available from 2015 onward. Some of the collected indicators contained outliers that went far beyond the expected value range, likely due to data collection errors. These were therefore replaced with their mean values. Mean imputation was also applied to all missing values across the variables. Subsequently, two binary variables were created to indicate default (1) and non-default (0). The first one, used to train the models, includes all municipalities that experienced default up to 2019 (599 cases), while the second accounts for those from 2020 onward (91 cases). The year of the pandemic is thus used as the separation criterion.

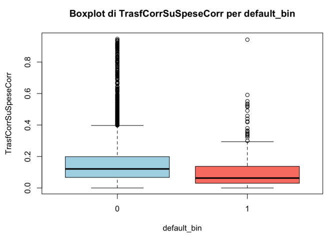


Figure 1

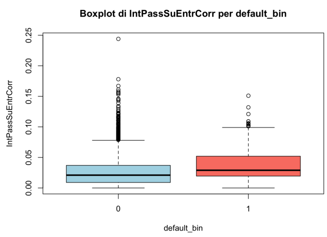


Figure 2

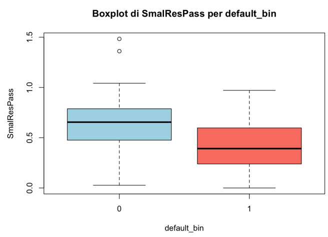


Figure 3

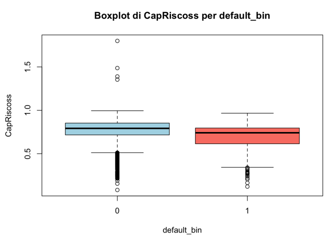


Figure 4

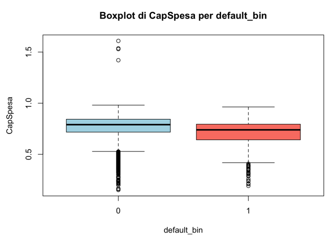


Figure 5

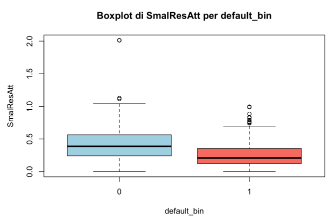
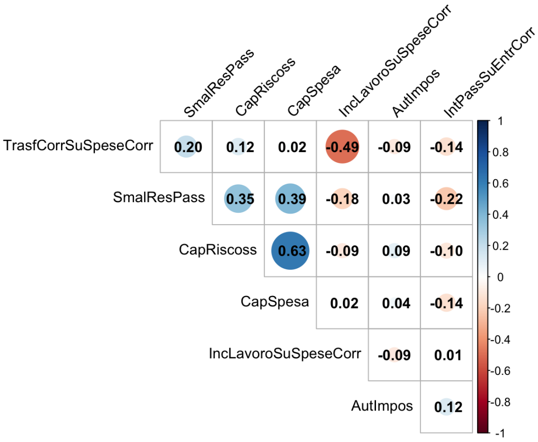


Figure 6

From an initial descriptive analysis of the variables using boxplots, it emerges in particular that the *current transfers over current expenditures* show a higher median and higher quartiles in financially sound municipalities. As for the *interest expenditure over current revenues*, values tend to be higher in municipalities that have declared financial distress at some point—consistent with what one might expect. Other notable differences are found in the *ability to manage outstanding receivables and payables*, with significantly higher values observed in municipalities that have not experienced budgetary crises. Finally, no clear evidence emerges regarding the variables *collection capacity* and *spending capacity*.

Matrix of correlations between variables



From the correlation matrix, the most notable relationships emerge between the variables *collection capacity* and *spending capacity* (around 0.63), indicating that the fiscal margin of a municipal administration tends to be greater when it is more efficient at collecting revenue. Both variables are also correlated with the *index of outstanding payables management* (0.35 and 0.39), highlighting the beneficial effects of the ability to settle debts. On the opposite side, a negative correlation (around -0.49) emerges between *current transfers over current expenditures* and *personnel expenditure incidence* on the same, suggesting a possible trade-off between these two expenditure components. All other associations indicate weak or no linear relationships.

2 Tools and Methods of Analysis

As for the predictive analysis, various classification methodologies were employed to test the ability to predict financial distress. These models were initially trained using financial ratios from 2019 as predictors and the variable containing defaults from 1989 to 2019 as the dependent variable. Subsequently, the models were tested on a dataset reporting defaults from 2020 to the present. This latter variable is included in the dataset containing financial indicators dating back to 2022. The decision to train the models on 2019 data instead of 2022 data is based on two main reasons. First, to ensure a sufficient sample of new defaults to assess the predictive accuracy of the models; second, to use the COVID-19 pandemic year as a temporal separator and a potential exogenous shock to the economy.

The first model employed is logistic regression, due to its parametric nature and versatility in conducting inferential analyses. Specifically, three regressions are performed, with the dependent variable represented by Default_bin. The process starts by including all financial indicators and then progressively eliminating non-significant predictors in the subsequent two steps. In fact, in the second logistic regression model (GLM with reduced variables), the following indicators are excluded: spending capacity, accumulation index of receivables and payables, and the human resources turnover index. In the third model (logistic regression with significant variables), the following are retained: collection capacity, disposal of receivables and payables, interest expenses, and current transfers over current revenues and population.

The second model is a Linear Discriminant Analysis (LDA), under the assumptions that the decision boundary could be approximately linear and that the assumption of normality is reasonably valid. In any case, the LDA model proved more accurate on the training set than a QDA model, which is characterized by a non-linear boundary.

The third statistical method considered is the classification tree, due to its potential explanatory simplicity. However, the resulting classification tree contains six nodes, all belonging to the non-default class. Consequently, the model was discarded. The complete absence of predictive nodes for municipalities in default rendered the model unsuitable for the analytical purpose of the project.

Finally, for completeness of analysis, a non-parametric yet highly flexible model was used: K-Nearest Neighbors (KNN).

Training the models on defaults up to 2019 and testing them on data from 2020 onward constitutes a time-based split of the dataset. However, several data resampling techniques exist. In our case, we opted for K-fold cross-validation with K=10 on two models: the logistic regression model with significant variables and the LDA model. Considering the randomness inherent in sample partitioning, the procedure was repeated three times for each analytical technique.

In conclusion, Principal Component Analysis (PCA) was applied exclusively to municipalities that entered default, as the large number of explanatory variables might benefit from dimensionality reduction. To improve clarity of presentation, municipalities are shown by geographical area (North-West, North-East, Centre, South, and Islands), and also aggregated into two macro-groups (North-Centre and South-Islands).

3 Results of the analysis

The analysis, as described above, was divided into four main stages:

- model estimation on the training set,
- validation on the test set data,
- confirmation of model fitting with resampling,
- principal component analysis to explore the latent structure of the data and identify any spatial patterns.

For each model, accuracy, sensitivity, and specificity indicators were calculated. The model selection was based on a comparison of these metrics, as the objective of our research is not only to assess how accurately the model classifies municipalities, but also how effectively it identifies those in financial distress (sensitivity) and how well it avoids false positives (specificity).

In our applied context, the early identification of municipalities at risk of default is a critical aspect for guiding preventive or supervisory interventions. Therefore, a model with high sensitivity is preferred, as it reduces the risk of overlooking entities that may indeed be heading toward financial difficulties. At the same time, specificity remains important to avoid misclassifying financially healthy municipalities as at risk, which could lead to inefficient resource allocation or unjustified alarm.

The predictive analysis was conducted using the three classification models mentioned earlier, trained on default data up to 2019 and tested on data from 2020 onwards. Below, the performance metrics are reported for the models trained on the training set and tested on the test set.

All models tested, sharing the same cut-off threshold, demonstrated high accuracy on both the training and test sets. Logistic regression showed lower accuracy compared to the other two models, presumably due to the constraint of linearity. LDA, while also assuming linearity, exhibited higher accuracy—likely as a result of its underlying assumptions of

Performance of predictive models on train and test set

Set	Model	Accuracy	Sensitivity	Specificity
Train	Reg. Log. (Full)	0.7852	0.6260	0.7983
	Reg. Log. (Reduced)	0.7838	0.6260	0.7968
	Reg. Log. (Significant)	0.7843	0.6294	0.7970
	LDA	0.9200	0.0367	0.9926
	KNN (k = 6)	0.9293	0.1720	0.9915
Test	Reg. Log. (cut-off – 0.1)	0.7446	0.8132	0.7438
	Reg. Log. (cut-off – 0.05)	0.5209	0.9780	0.5155
	LDA	0.9790	0.0659	0.9898
	KNN (k = 6)	0.9735	0.1319	0.9834

normality and homoscedasticity. KNN, on the other hand, is not a parametric model and adapts well to complex data structures. With regard to sensitivity, the logistic regression model with a cut-off of 0.1 improves the true positive rate at the cost of a slight decrease in specificity. This is particularly useful in our context, as identifying municipalities at risk of default (1) is more important than avoiding a few false positives. A threshold of 0.05 pushes sensitivity even further but compromises overall performance. The LDA and KNN models, although achieving very high accuracy, exhibit poor sensitivity: they primarily identify non-default cases (0). The high accuracy is largely attributable to the class imbalance in the dataset, where the majority of municipalities are not in default—thus, predicting non-default consistently already yields high accuracy. However, neither LDA nor KNN is inherently designed to handle imbalanced datasets without external adjustments. Both the logistic regression model with significant variables and the LDA confirmed the strong performance observed during training. Nonetheless, due to the low interpretability of KNN—and considering the highly practical nature of the research—the final evaluation was conducted using cross-validation, excluding the KNN model. This decision was also supported by the similarity in results between KNN and LDA.

To assess the robustness of the predictive models, we applied a 10-fold cross-validation procedure, repeated three times with different random seeds to ensure the stability of the results. The two best-performing models from the previous analyses were tested: the logistic regression model with significant variables (GLM) and the Linear Discriminant Analysis (LDA). The GLM exhibited an average error rate of 6.39% across all repetitions, while the LDA reported values ranging between 8.02% and 8.06%. These results confirm the strong predictive capacity of both models, with the GLM standing out due to its lower error rate and higher sensitivity.

We chose to use the error rate as the primary metric at this stage because it provides a concise and comparable measure of out-of-sample performance. The repetition of cross-validation with different data partitions further strengthens the reliability of the results obtained.

In light of the metrics obtained at the various stages of the analysis, the logistic regression model with significant variables, using a classification threshold of 0.1, was selected as the best-performing model. This model demonstrated high sensitivity on the test set (81%), good specificity (74%), and stable cross-validation performance, with an average error rate of 6.39%. These results indicate a solid balance between predictive ability and control over classification errors, making it well-suited for preventive screening purposes.

From an interpretive point of view, the model makes it possible to identify some relevant factors in determining the probability of disruption. In particular:

Results of logistic regression model with significant variables - Cut-off 0.1

Variable	Estimate	Std. Error.	Z-value	P-value	Odds ratio	Sign.
(Intercept)	0.7049	0.2441	2.888	0.00388	2.023	**
CapRiscoss	-1.5320	0.3271	-4.683	2.82e-06	0.216	***
SmalResAtt	-1.7800	0.3282	-5.422	5.90e-08	0.169	***
SmalResPass	-2.6480	0.2807	-9.436	<2e-16	0.071	***
IntPassSuEntrCorr	5.4660	1.7840	3.064	0.00218	236.7	**
TrasfCorrSuSpeseCorr	-2.5090	0.4989	-5.030	4.90e-07	0.081	***
popolazione	2.229e-06	9.743e-07	2.288	0.02213	1.000002	*

– *Collection capacity* and the *disposal indices of receivables and payables* have negative and significant coefficients, indicating that efficient management of revenues and legacy flows reduces the risk of default;

– the indicator related to *interest expense* is positively associated with financial distress, consistently with a structural burden on the municipal budget;

– the ratio of *current transfers to current expenditures* also shows a negative effect, suggesting that greater external resources may temporarily offset internal weaknesses;

– finally, *population* is weakly but significantly correlated with the risk of default, suggesting that larger municipalities may be more exposed to financial stress.

The selected model, in addition to demonstrating the best predictive performance, also provides a clear and consistent interpretation of the main risk factors, offering a valuable tool both from an analytical and a policy-making perspective.

In order to explore the latent structure of the data and identify potential territorial patterns, a Principal Component Analysis (PCA) was conducted on the quantitative variables referring to the year 2019. PCA allows for dimensionality reduction while preserving as much variance as possible, by projecting the observations onto the first two principal components. In our case, the first principal component (PC1) explains 17.2% of the total variance, while the second (PC2) accounts for 12.4%; together, the two components capture approximately 29.6% of the information contained in the data. This level of captured variance is relatively limited, suggesting that the data structure is spread across multiple dimensions.

The resulting graphs show the distribution of municipalities according to:

– the classification *Centre-North* vs *South-Islands*;

– the ISTAT partition into five areas (North-West, North-East, Centre, South, Islands).

In the first graph, a certain differentiation between municipalities in the *Centre-North* and *South-Islands* can be observed, with a slight tendency for municipalities in the South-Islands to be positioned toward positive values of PC1. However, the overlap between the two areas is substantial, suggesting the absence of a clear separation. The second graph, based on the five geographical partitions, confirms a high degree of intermixing among regional groups. A greater dispersion of Southern municipalities along PC1 is noticeable, which may reflect broader internal variability in economic and financial indicators.

In conclusion, the principal component analysis does not reveal clearly distinct territorial clusters, but it nonetheless provides a useful synthetic representation for exploratory

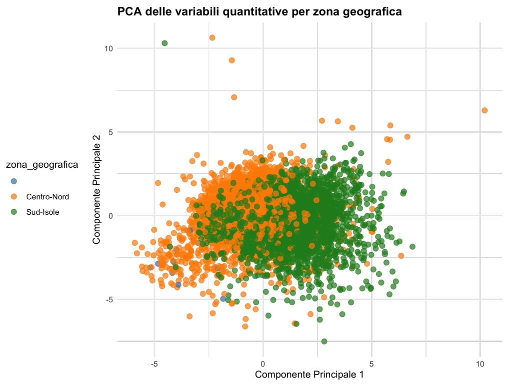


Figure 7

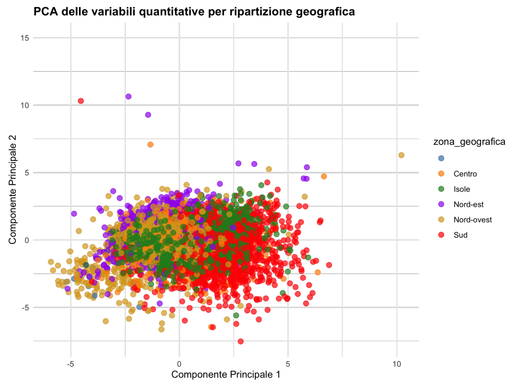


Figure 8

visualization and the identification of complex patterns. The evidence suggests that the variables used are not sufficient, on their own, to discriminate municipalities on a geographical basis.

4 Comparison Between Models and Conclusions

The comparison between models highlighted the following points:

- The GLM with significant variables offered good interpretability and solid results in both the training and test sets, although it showed some sensitivity to the choice of threshold.
- The LDA classifier stood out for its robustness and stability, with a good balance between accuracy and simplicity, ultimately proving to be the best overall model in the K-fold validation.
- The KNN model suffered from greater instability and less satisfactory results, likely due to the high dimensionality and the nature of the data.

The comparison among models revealed very heterogeneous results depending on the metric considered. In our context, the primary objective is the identification of municipalities at potential risk of financial distress, which is why the *sensitivity* metric takes on central importance, even at the expense of perfect accuracy or specificity. The model that demonstrated the best trade-off among these needs is the logistic regression model with significant variables and a cut-off of 0.1.

This model achieved high sensitivity (81%), while maintaining acceptable specificity (74%) and good overall accuracy (74%). These characteristics make it particularly well-suited for screening and prevention purposes, as it reliably identifies the majority of at-risk municipalities, minimizing the risk of omissions. In addition to its predictive effectiveness, this model is also suitable for economic-institutional interpretation thanks to the structure of logistic regression. Through the analysis of estimated coefficients, it is possible to understand which financial indicators have a significant impact on the prediction of distress, thus providing useful information for both policy-making and risk analysis. For this reason, the GLM with significant variables is not only the most effective predictive tool, but also the most informative and transparent for interpreting results. The ability to perform statistical inference on the coefficients and to analyze the odds ratios associated with individual predictors represents an added value compared to less interpretable algorithms such as KNN.