doi.org/10.26398/asaproc.0073

Modelling default risk in Italian SMEs: A spatial probit approach

Emanuele Amoruso[a], Emma Bruno[b], Rosalia Castellano[b], Gennaro Punzo[a]

[a] Department of Economic and Legal Studies, University of Naples Parthenope, Naples, Italy.
[b] Department of Management and Quantitative Studies, University of Naples Parthenope, Naples, Italy.

1. Introduction

Default risk analysis has increasingly become crucial for banks and rating agencies (Altman et al., 2020), as they require analytical tools to effectively evaluate loan approvals and creditworthiness, especially for small and medium-sized enterprises (SMEs) (Modina et al. 2023; Ciampi et al., 2021). Unlike large companies, which undergo financial statement audits, lending to SMEs is often riskier due to their less transparent financial disclosures and non-audited financial statements (Calabrese et al., 2019). Moreover, small firms face greater challenges in accessing credit and obtaining financial resources and encounter stricter borrowing requirements from financial institutions compared to larger firms (Wang and Guedes, 2024).

In the Italian economic context, SMEs form the backbone of the business landscape, representing nearly all Italian companies. SMEs employ approximately 76.5% of the workforce and contribute around 65% to the creation of added value at factor cost (Istat, 2023). Analysing their financial stability is essential for understanding overall economic health and making informed loan approvals by financial institutions. Therefore, constant and rapid monitoring of both internal risk factors and firms' financial performance is needed to develop predictive models for insolvency or acquisition likelihoods (Altman et al., 2023).

Alongside internal factors, it is reasonable to assume that a firm's location provides additional information, particularly in assessing the default risk of SMEs, which are presumed to be more influenced by local economic conditions (Fernandes and Artes, 2015). Recent research (Calabrese, 2023; Medina-Olivares et al., 2022; Chen and Lee, 2020) underscores the necessity of considering interdependences among businesses in credit scoring models (Calabrese et al., 2019), as the default probability of firms can be mutually affected by the performance of neighbouring businesses.

This research aims to evaluate how spatial dependencies alongside firms' financial indicators contribute to assessing default risk using spatial probit models on 2022 firm-level data of Italian SMEs. These models account for endogenous and exogenous interaction effects by integrating the spatial lag in the target variable and predictors. This approach intersects with spatial finance, an emerging field that integrates spatial data into financial theory and practice (Caldecott et al., 2022). Spatial finance leads the forefront of future financial analysis by spatially modelling information markets, risk assessment, valuation techniques, and investment opportunities, paving the way for innovative financial decision-making and risk management approaches.

2. Methods and data

The data were sourced from the AIDA database, which includes economic and financial performance indicators, as well as demographic and commercial information for all active and bankrupt joint-stock Italian companies with a history spanning up to ten years. The AIDA database is georeferenced, providing location-based information for each company. Focusing on SMEs, a sample of 9,292 firms with fewer than 250 employees and annual turnover or balance sheet totals not exceeding 50 million euros or 43 million euros, respectively, were extracted in the latest

available year, 2022. It serves as a representative sample for evaluating credit ratings, assessing creditworthiness, and constructing predictive models for insolvency or acquisition likelihoods.

To develop a supervised learning model, it is crucial to establish a target variable for classifying firms as solvent or insolvent. This variable was based on the legal status of SMEs, categorising them as: active in insolvency status, bankrupt, ceased, closed due to bankruptcy, and active but with annual total debts exceeding 500,000 euros. These firms were coded as '1', while all other active solvent firms were coded as '0'. The choice of covariates was informed by relevant literature (Altman et al., 2020), while accounting for constraints related to data availability and missing data issues: i) EBITDA to sales ratio; ii) Banks liabilities to turnover ratio; iii) Debts to EBITDA ratio; iv) Debts to equity ratio; v) Liquidity index; vi) Debt ratio; vii) Short-term investment index; viii) Employee costs; and ix) ROI.

To simultaneously explore the predictors of default probability among SMEs and the role of spatial proximity between firms, two probit spatial models were proposed (Lacombe and LeSage, 2018; McMillen, 1992): the probit spatial autoregressive (PSAR) model, including spatial lag only in the target variable:

$$\Phi^{-1}\big(P_r(Y_i = 1)\big) = \rho WY + X\beta + \epsilon \quad with \;\; \epsilon \sim N(0, \sigma^2 I) \quad (1)$$

and the probit spatial Durbin model (PSDM), which additionally includes spatial lag among predictors:

$$\Phi^{-1}\big(P_r(Y_i = 1)\big) = \rho WY + X\beta + \theta WX + \epsilon \quad with \;\; \epsilon \sim N(0, \sigma^2 I) \qquad (2)$$

where P_r denotes probability, Y is the $n\ x\ 1$ vector of the binary target variable with a Bernoulli distribution $Y \sim Be\big(\Phi(X^T\beta)\big)$, X is the $p\ x\ n$ matrix of covariates, β is the $p\ x\ 1$ vector of unknown parameters, Φ denotes the cumulative distribution function of the standard normal distribution, ρ is the autoregressive parameter, W is the spatial weight matrix, and ε is the error term; θ in equation 2 is the $p\ x\ 1$ vector of unknown parameters of lagged covariates.

The parameter estimation for both probit spatial models was based on maximum likelihood. A *k*-nearest neighbours (*knn*) spatial weight matrix was used with different neighbour counts (5, 10, 15, 20, and 25) to determine the model's optimal specification. The choice of *k* was informed by classifier accuracy indicators, such as sensitivity and specificity, as well as the model's ability to capture spatial autocorrelation.

3. Results and discussion

This section discusses the results of the probit model without spatial effects, the probit spatial autoregressive (PSAR), and the probit spatial Durbin model (PSDM) for 2022 (Table 1). Probit spatial models analyse the causal relationship between firms' status (insolvent or solvent) and economic and financial predictors selected via the LASSO method. LASSO shrinks the estimated regression coefficients and sets those with low explanatory power to zero, retaining only those that ensure the best model fit, thereby avoiding multicollinearity. The model's complexity is determined by both the original predictors selected by LASSO and the spatially lagged predictors, including the spatial lag of the target variable.

Both the PSAR and PSDM with *k*=15 show the best performance compared to other *k* values. Accuracy further improves with PSDM, showing a 24.5% increase compared to the probit model without spatial effects. Sensitivity (the ability to correctly predict defaulting firms – improved by 15.8%, while specificity (the ability to correctly predict solvent firms) improved by 48.1%. Moreover, PSDM specification has nearly eliminated spatial autocorrelation in the residuals, achieving an 87.7% reduction and resulting in negligible Moran's I residuals. Further supporting evidence comes from the ROC curves, where the PSDM specification shows an area of 0.984 compared to 0.816 for the probit model without spatial effects.

Table 2. Probit without spatial effects, PSAR, and PSDM estimates for k=15

	Probit without spatial effects	**PSAR**	**PSDM**
Intercept	-0.286***	-1.843***	-1.157***
Rho	–	0.739***	0.686***
Liquidity index	-0.178***	-0.474***	-0.579***
Banks liabilities to turnover ratio	0.012***	0.011***	0.047***
Employee costs	0.002***	0.005***	0.006***
Debts to EBITDA ratio	0.034***	0.101***	0.109***
W*Liquidity index	–	–	-0.423***
W* Banks liabilities to turnover ratio	–	–	-0.088***
W*Employee costs	–	–	0.001***
W*Debts to EBITDA ratio	–	–	0.124***
Accuracy	0.746	0.907	0.929
Sensitivity	0.819	0.932	0.948
Specificity	0.603	0.858	0.893
Moran' s I residuals	0.381***	0.094	0.047

***significant at 1%; **significant at 5%, *significant at 10%

The positive spatial autocorrelation coefficient shows spatial dependency among neighbouring firms. Such dependency arises because firms located near each other may share common risks or be influenced by local factors, such as regional economic policies, market conditions, or shared infrastructure. Consequently, these shared influences likely lead to similar economic and financial characteristics, resulting in comparable default risks.

The results highlight several factors influencing firms' default risk and reveal how the economic and financial performance of both firms and their neighbours can affect the probability of default. Specifically, a higher ratio of banks' liabilities to turnover increases default risk, indicating that firms with a greater exposure to bank debts are more vulnerable to default. However, this effect is partially mediated by the ratio of banks' liabilities on turnover of neighbouring firms, meaning that firms' default risk may be mitigated by their operational context and the debt situation of nearby businesses. Higher employee costs increase default risk, suggesting that firms burdened with higher workforce-related expenses may encounter heightened financial strain. In contrast to the spatial effect of banks' liabilities to turnover, high employee costs among neighbouring firms may intensify default risk. A higher debt-to-EBITDA ratio indicates higher default risk, reflecting firms with substantial debt relative to their earnings facing increased vulnerability to default. The effect of neighbouring firms' debt-to-EBITDA ratio reinforces the positive impact on default risk, suggesting systemic financial challenges across adjacent firms. On the other side, a higher liquidity index decreases default risk, indicating that firms with stronger liquidity status are less likely to default due to their greater ability to meet financial obligations. Similarly, the liquidity status of neighbouring firms reinforces the dampening effect on default risk.

4. Conclusion

The results drawn from this research highlight both the data modelling outcomes and the practical implications. The study demonstrates that both the estimated spatial probit models outperform traditional probit models without spatial effects, providing more efficient results and

proving to be effective tools for assessing default risk. The findings underscore the importance of considering spatial dependencies in credit risk assessment to enable financial institutions to make more informed lending, as well as implications for policy and corporate strategy in managing default risk among SMEs. Factors such as banks' liabilities to turnover, employee costs, debt-to-EBITDA ratios, and liquidity indexes play crucial roles in shaping default probabilities among neighbouring firms.

Effective regional economic policies should focus on fostering financial networks, targeted monitoring, and intervention strategies tailored to economic diversification. Such measures can enhance resilience against economic downturns and mitigate systemic risks across neighbouring firms. Strategies centred on collaborative risk management, optimising labour costs, and careful liquidity management are crucial for SMEs. These approaches not only strengthen firms' financial health but also contribute to collective resilience within local business ecosystems. Financial institutions are encouraged to integrate spatial context into risk assessments and offer diversified financial solutions tailored to varying risk profiles among SMEs.

References

Altman, E.I., Balzano, M., Giannozzi, A., Srhoj, S. (2023). Revisiting SME default predictors: The omega score. *Journal of Small Business Management,* **61**(6), pp. 2383-2417.

Altman, E.I., Esentato, M., Sabato, G. (2020). Assessing the credit worthiness of Italian SMEs and mini-bond issuers. *Global Finance Journal*, **43**, 100450.

Calabrese, R. (2023). Contagion effects of UK small business failures: A spatial hierarchical autoregressive model for binary data. *European Journal of Operational Research*, ***305***(2), pp. 989-997.

Calabrese, R., Andreeva, G., Ansell, J. (2019). "Birds of a feather" fail together: Exploring the nature of dependency in SME defaults. *Risk Analysis*, **39**(1), pp. 71-84.

Caldecott, B., McCarten, M., Christiaen, C., Hickey, C. (2022). Spatial finance: Practical and theoretical contributions to financial analysis. *Journal of Sustainable Finance & Investment*, pp. 1-17.

Chen, T. H., Lee, C. C. (2020). Spatial analysis of liquidity risk in China. *The North American Journal of Economics and Finance*, **54**, 101047.

Ciampi, F., Giannozzi, A., Marzi, G., Altman, E.I. (2021). Rethinking SME default prediction: A systematic literature review and future perspectives. *Scientometrics,* **126**, pp. 2141-2188.

Fernandes, G.B., Artes, R. (2016). Spatial dependence in credit risk and its improvement in credit scoring. *European Journal of Operational Research,* **249**(2), pp. 517-524.

Istat (2023). *Annuario Statistico Italiano*. Istat.

Lacombe, D. J., LeSage, J. P. (2018). Use and interpretation of spatial autoregressive probit models. *The Annals of Regional Science*, **60**, pp. 1-24.

McMillen, D.P. (1992). Probit with spatial autocorrelation. *Journal of Regional Science*, **32**(3), pp. 335-348.

Medina-Olivares, V., Calabrese, R., Dong, Y., Shi, B. (2022). Spatial dependence in microfinance credit default. *International Journal of Forecasting*, ***38***(3), pp. 1071-1085.

Modina, M., Pietrovito, F., Gallucci, C., Formisano, V. (2023). Predicting SMEs' default risk: Evidence from bank-firm relationship data. *The Quarterly Review of Economics and Finance,* **89**, pp. 254-268.

Wang, W., Guedes, M.J. (2024). Firm failure prediction for small and medium-sized enterprises and new ventures. *Review of Managerial Science,* pp. 1-34.