Modina, Michele; Zedda, Stefano (2023) **Diagnosing default syndromes: early symptoms of entrepreneurial venture insolvency**, *JOURNAL OF SMALL BUSINESS AND ENTERPRISE DEVELOPMENT*, v. 30, n.1

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1462-6004.htm

Diagnosing default syndromes: early symptoms of entrepreneurial venture insolvency

Michele Modina
Department of Economics, Universita degli Studi del Molise, Campobasso, Italy, and
Stefano Zedda
Department of Economics and Business Sciences, University of Cagliari, Cagliari, Italy

Received 14 February 2022
Revised 11 April 2022
Accepted 22 April 2022

Abstract

Purpose – In this study, a panel of 74,128 Italian SMEs was analyzed to verify whether any syndromes could be identified and defined through financial ratios. Defining relevant syndromes (i.e. the set of correlated signs and symptoms often associated with a particular disorder) can be of importance for assessing which specific intervention can solve a firm's difficulties.

Design/methodology/approach – To identify the main syndromes involved in company defaults, firstly, financial data on defaulted firms for each of the main economic sectors were examined through a cluster analysis; the results obtained for each sector were then compared to verify whether syndromes recur across sectors. Finally, the effects of each syndrome were compared with possible default causes, as described by previous literature.

Findings – Results show that a significant share of corporate insolvencies is characterized by a set of recurrent signs and symptoms so that the main syndromes can be identified. The results also show that these syndromes recur across sectors, even if specific values characterize each sector.

Research limitations/implications – The approach adopted in this study sets a new direction for the analysis of default risk, as the study shows that certain key syndromes can be defined and described, and the study suggests that different problems can induce different risk patterns. Further analyses of other samples could confirm whether the same syndromes recur over countries and over time.

Originality/value – This is the first study aimed at identifying and describing the syndromes affecting SMEs, conducted by means of balance-sheet ratios.

Keywords Default syndromes, SMEs, Cluster analysis, Default risk, Business failure

Paper type Research paper

1. Introduction

Within the business sciences, little attention has been devoted to identifying syndromes (sets of associated symptoms) that characterize firms' weaknesses and defaults. Researchers have mainly focused their attention on assessment of firms' probability of defaulting (Altman, 1968), while analysis of the paths leading to firm defaults has either been developed with reference to specific case studies, to the characteristics and attitudes of entrepreneurs (Khelil, 2016; Mayr *et al.*, 2021) or to firm life cycles (Habib and Hasan, 2017).

The word "syndrome" is typically used in medical circles, where diseases are described through sets of associated symptoms. The *Oxford Dictionary* defines a syndrome as "a group of symptoms that consistently occur together, or a condition characterized by a set of associated symptoms" [1].

Although defining recurring symptoms may seem an easy task, identifying the differences among typical values for firms of different sizes, from different business sectors, or living different life cycle phases is not straightforward. Previous studies have conducted extensive investigations of methodological tools and technical approaches in the credit-risk assessment of SMEs. Yet, no studies have sought to describe, through financial ratios and cluster analyses, the syndromes affecting firms and leading them to default. This

Journal of Small Business and Enterprise Development
© Emerald Publishing Limited
1462-6004
DOI 10.1108/JSBED-02-2022-0088

paper provides novel evidence of the possibility of detecting default syndromes in SMEs, obtained through cluster analysis of a large sample of 74,128 non-financial Italian SMEs (for the period 2011–2014) financed by 123 local banks.

This paper contributes to three fields of literature. The first concerns the business failure research stream. The main factors relating to small-firm bankruptcies are the personal characteristics of the entrepreneur and internal and external firm-specific factors (Mayr *et al.*, 2017; Strotmann, 2007; Laitinen and Gin Chong, 1999). In SMEs, internal factors and a lack of resources (especially financial resources) are the main cause of business failure (Mayr *et al.*, 2021). Examining the typology of four different bankruptcy processes, Ooghe and De Prijcker (2008) found similarities in the evolution of financial performance ratios during the years preceding bankruptcy. However, these financial ratios show unequal predictive power on different firms' categories (unsuccessful start-up companies, ambitious companies with spectacular growth trajectories, apathetic established companies, etc.). Efforts to define and describe syndromes are important for evaluating the typical ways the SMEs report their specific weaknesses and gaining a better understanding of the causes of their failure. Our paper contributes to this field by proposing a way for overcoming the pervious literature limits, showing that syndromes can be identified and described by financial ratios, and bridging the quantitative and qualitative approach on this field.

The second field refers to the specificities of the credit risk assessment of SMEs. Various forms of corporate credit-risk assessment models have been examined in the literature, but such analyses have, for the most part, failed to take into account the characteristics of smaller firms. SMEs have specific peculiarities that set them apart from larger firms on which the existing literature on default prediction modeling has mainly been based (Peel and Peel, 1989; Norden and Weber, 2010). Several studies show that SME lending suffers from more severe agency problems, exhibits a higher default risk and that bank-firm relationship are more informationally opaque (Burgstahler *et al.*, 2006). Few papers have explicitly analyzed the quality and intensity of the credit relationship between local banks and small and medium-sized enterprises (Behr and Guttler, 2007). In this paper, instead, a dataset made up of over 70,000 Italian SMEs (operating with more than 100 local banks) was analyzed to describe the syndromes that lead firms to default, suggesting a new point of view to assess the creditworthiness of borrowing companies. By proposing a diagnosis based on financial indices, our approach allows the bank to carry out periodic automated screening of SMEs which, by focusing attention on specific syndromes (sources of risk), favor a more effective bank-company relationship.

The third field of contribution refers to literature on evaluating the probability to default (PD). Starting from the seminal paper of Altman (1968, 2002) and Bellovary *et al.* (2007) highlighted the amount of academic and professional work devoted to modeling and predicting borrowers' (probability to) default. As reported by Sigrist and Hirnschall (2019), initially, the statistical methodology used for predicting defaults quite often involved linear classification models, such as linear discriminant analysis (LDA) or logistic regression (Altman, 1968; Shumway, 2001; Altman and Sabato, 2007; Bauer and Agarwal, 2014; Tian *et al.*, 2015). Following the evolution of statistical techniques, more recent papers have also used generalized additive models, neural networks, classification trees, ensemble models (Jones *et al.*, 2017; Cheng *et al.*, 2018; Jones and Wang, 2019), and standard boosting algorithms (Xia *et al.*, 2017). Among the few studies using cluster analysis to predict borrowers' creditworthiness, Lim and Sohn (2007) employed a model based on a neural network that dynamically determines high-risk borrowers' credibility by setting up separate classifiers for each cluster at different points in time.

This paper complements previous findings in this field by changing the point of view and aiming to verify how firms typically default, i.e. to describe the syndromes that lead firms to default, rather than measuring the probability of default, and suggesting that each firm can be characterized by its risk of suffering from each syndrome, instead of a unspecific probability to default. The results showed that the main syndromes affecting SMEs are quite similar across sectors, even if the values of variables are different. More specifically, the main syndromes were not found to be characterized by the specific values of particular variables but by the difference between these values and standard values for healthy firms from the same sector.

The remainder of this paper is organized as follows: Section 2 presents a brief overview of relevant literature, Section 3 presents the methodology and data, Section 4 describes the empirical strategy, Section 5 discusses the results, and Section 6 concludes.

2. Literature review and key hypotheses

The contribution of SMEs to modern economies is very significant. In the European Union (EU), they make up 99.8% of all enterprises in the EU-28 non-financial sectors; they employ two-thirds of the working population, and they generate 56.4% of the added value (European Commission, Executive Agency for Small and Medium-sized Enterprises, 2020). SMEs are the backbone of economic development and adapt to new trends, such as sustainable growth and innovation (Canto-Cuevas *et al.*, 2019). Despite their role, access to credit or other financial sources is not easy for many SMEs. Although the last decade has seen the development of new financing alternatives (including non-banking ones – among others, crowdfunding (Eldrige *et al.*, 2021) and peer-to-peer financing mechanisms (Liu *et al.*, 2020), bank credit continues to be the main financial source for SMEs (Beck *et al.*, 2018, 2011; Franquesa and Vera, 2021), being fundamental for their economic activity and survival.

Financial resources provided through the banking channel are highly variable within the business cycle, creating significant effects on SMEs (Jorion and Zhang, 2009) and leading to an increase in firm failures, many of which were involuntary during the crisis period (McGuiness *et al.*, 2018). When uncertainty increases, information asymmetries between borrowers and lenders also increase, heightening lenders' exposure to credit risk (Delli Gatti *et al.*, 2003; Sette and Gobbi, 2015). Consequently, banks' propensity to lend declines, making it more difficult for SMEs to access bank credit (Calabrese *et al.*, 2016).

Financial restrictions during financial crises and economic downturns were exacerbated by the introduction of new regulations, starting with the first Basel Committee capital accord in 1988 and its subsequent revisions (i.e. Basel II/III), which required banks to implement an internal assessment system aimed at measuring and managing their credit-risk exposure (Duarte *et al.*, 2018) and covering such (and other) risks through a minimum capital requirement. In fact, as financial crises typically make the probability of default significantly rise, banks' typical reaction is to reduce credit risk by narrowing their credit lines, creating more difficulties for borrowers, and inducing procyclical effects, in particular to SMEs.

Given the importance of bank credit for SMEs and the difficulties in accessing it (especially in times of uncertainty), predicting the probability of company insolvency is a critical activity when it comes to banks managing credit risk, and it can have a marked effect on the capacity of the banking channel. Although credit scoring and rating models have evolved over time, their focus has remained on individual borrowers' creditworthiness (i.e. PD estimates), and the literature mainly focuses on larger firms.

Credit scoring and rating models must be tuned to their specificities when addressing default risk assessment for SMEs. SMEs rely more on bank debt (Beck and Demirguc-Kunt, 2006) and less on capital markets (Berger and Udell, 2002), are more informationally opaque (Burgstahler *et al.*, 2006), and suffer from more severe agency problems and financial constraints (Beck *et al.*, 2018; Stiglitz and Weiss, 1981), which increase during periods of crisis (Ryan *et al.*, 2014; Ghosal and Ye, 2019).

Different factors affect SMEs' capacity to repay their debt. The empirical literature on corporate credit risk relies on different information sources to estimate borrowers' probability of default. A substantial line of research has focused on accounting-based approaches. Such methodologies include all the statistical techniques which (based on quantitative information about each borrower, i.e. financial ratios and balance sheets) report a numerical score reflecting borrower credit quality and give an indication of the probability of a borrower defaulting (Beaver, 1966; Altman, 1968; Edminster, 1972; Blum, 1974; Ohlson, 1980; Grice and Ingram, 2001; Pindado *et al.*, 2008; Louzada *et al.*, 2016). Veganzones and Severin (2021) have provided a review of recent developments in this field.

In the most common setting, data on a sample of borrowers, in default and not in default, are used to build and adapt the model. The dataset consists of financial ratios and behavioral characteristics that allow an estimation to be made of the counterparty's risk level. The model result is expressed as a risk score, usually associated with a PD estimate and a risk-rating class (Andriosopulos *et al.*, 2019).

Over the years, credit-risk data have become much more comprehensive, integrating information from other sources, e.g. bank-firm relationship information (Norden and Weber, 2010; Fiordelisi *et al.*, 2014), enhancing the predictive power of such models (Altman *et al.*, 2010) and investigating determinants of the quality of banking choices (Brighi *et al.*, 2019).

Even with these model improvements, such assessments are typically used to help banks decide whether to finance a firm or not or whether to claw back credit allowances or not. Credit scoring models cannot support an interactive process within the bank-firm relationship aimed at reducing firms' default risk by treating problems specific to particular firms. The first step in developing this kind of interactive process is a diagnosis, i.e. evaluating which syndrome characterizes a firm, as would be the case with a medical doctor planning treatment for a patient, having determined the nature of the disease that the patient is suffering from. In this way, and given the importance of SMEs for the economic system, it would be of fundamental importance to detect and describe the main syndromes affecting SMEs, which may well cause them to default. In statistical terms, this means finding groups of observations characterized by a set of associated symptoms. After this classification and description process, it will then be possible to implement more efficient credit models, taking into account specific syndromes.

Previous studies (Laitinen *et al.*, 2014) have analyzed different groupings of firm defaults, but the use of formal bankruptcy as a trigger, a limited number of firms and financial ratios being considered, along with the inclusion of different sectors and different countries, seriously limit the capability of such analyses to identify the main characteristics of each group (limited to groups' dynamic evolution). Their results mainly show that some defaults happen in a sudden way, others are preceded by negative signs in previous years, and others indicate a slow decline.

As reported in the introduction, a cluster analysis is the standard tool for finding groups of observations characterized by similar results. This estimation proved to be more complicated than expected for the following reasons: (1) within the bank-firm relationship, the definition of default refers to irregular refunding of debts, which is not reported in balance sheets or annual reports and can be obtained from internal bank data; (2) to define syndromes, a large number of defaults are needed which can be ascribed to actual groups; (3) the characteristics of syndromes can differ across business sectors.

Cluster analysis enables cases to be grouped with similar values in terms of the variables considered. Clustering facilitates recognition of borrowers' heterogeneity, improving the overall efficiency of the analysis (Berry, 2000). Regarding defaulted firms, theoretically, three different results may emerge. The first possibility is that all defaulting firms are characterized by similar symptoms so that just one main group is identified; the second possibility is that some main groups are found, each one characterized by certain specific characteristics; the third possibility is that each firm path to default is specific, and no main groups can be identified. These possibilities can be represented by the following (the first) hypothesis:

H1. Firms' path to default, as described by financial ratios, can be classified into key groups.

If the process is able to classify firm defaults into main groups, then the second fundamental step is to describe the groups, i.e. the syndromes affecting firms and leading them to default, to investigate better how and why these firms have defaulted.

Existing empirical works have ascribed little value to the differences and similarities among firms in different economic sectors, which might make it possible to improve assessments of the probability to default, even though it is well known that the underlying business structures and the characteristics of firms are different among sectors, this inducing different significance and reference values for financial ratios (see, e.g. Brealey *et al.*, 2020). A test of the significance of distinctions between sectors is therefore needed in order to set the correct methodology. This can be represented by the second hypothesis:

H2. Each business sector is characterized by different balance-sheet equilibriums and different significance of financial ratios.

Previous studies of risk-prediction models in the banking industry have focused on the development and adaptation of estimation models (Crook *et al.*, 2007; Ravi Kumar and Ravi, 2007), trying to refine the algorithm by working on three aspects: the type of credit score data, the classification algorithms employed and the indicators used to assess these algorithms (Baesens *et al.*, 2003).

In fact, after identifying the syndromes affecting each business sector, an intersectoral comparison will enable exploration of the recurrence of similar syndromes across sectors to verify whether the same pathologies affecting different business sectors show different but recognizable symptoms. If this is true, then the number of syndromes (and thus the number of different approaches needed to treat unhealthy firms) can be considerably simplified, and the actual implementation of this approach will become simpler and more effective. This gives rise to the third hypothesis:

H3. The main syndromes leading firms to default are recurrent across business sectors.

3. Method and data

3.1 Methodology

The analysis was conducted by focusing on local bank-business relationship specificities and combining firms' financial variables with confidential information on the lending relationship. For identifying the main syndromes leading SMEs to default, data from defaulted firms for each key sector of activity were examined through a cluster analysis; the results obtained for each sector were then compared to verify whether the syndromes recurred across sectors (see Figure 1).

Regarding the selection of variables, the analysis started from a list of 60 balance-sheet financial ratios (see Table A1) originating from firms' annual financial statements (including the composition of assets and liabilities, the level and quality of debt, and economic performances). Lending information was used to form a list of 18 bank variables (see Table A2), reflecting bank account movements, long-term and short-term loans, the lending relationship (e.g. whether the customer was a shareholder of the cooperative bank), and credit default events.

As the number of variables was too large, and not all of them were actually significant, a set of logistic regressions was performed, for determining which variables reported some default-predicting power in any sector, and limit the subsequent analyses to the actually significant variables.

As a second step, to detect the main syndromes leading firms to default, a cluster analysis of the defaulted firms was performed separately for each selected sector. The analysis has been conducted on the significant balance-sheet variables, for the values reported one year before default, and filtering for the cases where the core activity sales were the primary

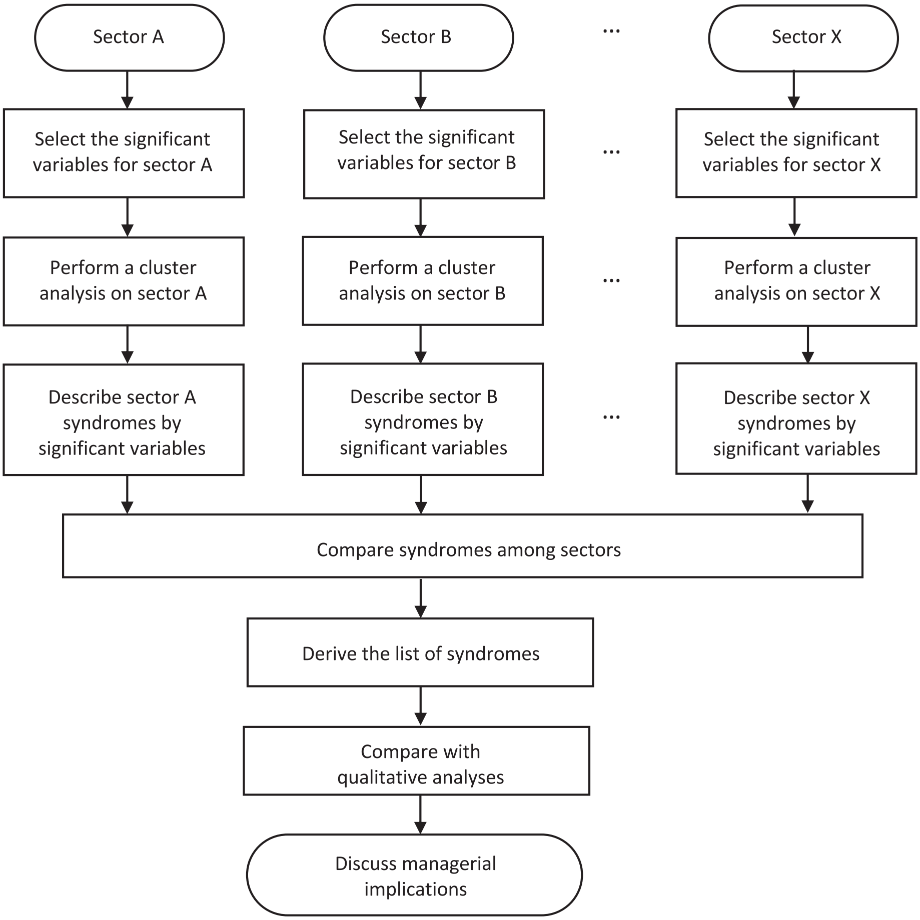

Figure 1.
Research plan flow chart

income source. On the one hand, this separation by sector allowed the different balance-sheet equilibriums reported in the previous results to be neutralized across sectors; on the other hand, it resulted in a parallel analysis of different samples, which allowed for cross-validation of results (Huang *et al.*, 2012).

Due to the need for significant clustering, the analysis was restricted to the five sectors with the highest numbers of defaulted firms (those with available measures for all significant variables), namely: wholesale, excluding cars and motorcycles (Ateco code 46); manufacture of metal products, excluding machinery and equipment (Ateco code 25); construction of buildings (Ateco code 41); retail, excluding cars and motorcycles (Ateco code 47); and specialized construction works (Ateco code 43).

3.2 Data

As reported above, the use of cluster analyses to define default syndromes requires two main features. The first one concerns the default definition. A default event in banking terms is defined as the irregular fulfillment of bank debts. This definition cuts off both commercial databases and institutional balance-sheet repositories since, for them, the trigger event is only formal bankruptcy, which can occur years after the first credit default. So, the possibility of performing any effective creditworthiness analysis is limited to internal banking data. Such data play a crucial role in assessing the firms' creditworthiness but often cannot be obtained from banks for confidentiality reasons. As a consequence, these datasets are often not available for scientific analyses, and when they are available, they are likely to come from a single bank interested in some in-depth analysis of its borrowers to suit the bank's management needs, or the datasets may refer to previous years, as, in this case, the loss of competitive advantage arising from public availability of data analyses is no more dangerous for the bank.

The second problem limiting the possible use of cluster analyses is the number of default cases, given that, in order to group defaults into significant groups, a large number of default cases are needed.

In this study, both problems were overcome by using a proprietary dataset combining public firm-level financial information from balance sheets with private bank-firm lending information for a sample of 74,128 balance sheets from Italian SMEs operating with 123 local banks. As reported above, to limit the possible loss of competitive advantage arising from public availability of results, the dataset time span was limited by the data source to the period from 2011 to 2014.

The considered setting included firms operating in different industries, classified according to the two-digit Ateco industry classification. In particular, the largest share of firms in the considered sample belonged to the following six macro-industries: agriculture, commerce, transport and hotels, manufacturing, building and services. Public administration and financial firms were excluded due to their specificities.

With reference to the sample significance, it is worth pointing out that even though the sample only contained Italian firms, the diversification of each region's economic structure meant that the sample had an interesting level of diversity and had the potential to be a proxy for a wide list of European regions. In fact, the per capita GDP of the considered Italian regions in 2014 ranged from 16,200 to 39,900 euro, which spans the average value range for 14 European countries for the same year [2]. Furthermore, the economic structure of regions is very diversified, so the chi-squared test for the association between regions and economic activities is significant at 99%.

The data source was Centrale Rischi Finanziari (CRIF), an Italian credit rating agency. Firms were anonymized and identified by a unique code produced by CRIF. Individual firms could be associated with one or more cooperative banks included in our sample (and measured) and other relationships with banks not included (and not measured).

The availability of such a comprehensive dataset and the inclusion of bank variables (including bank default events) was thus one of the key points for this analysis, even if data were limited in terms of the time span (2011–2014) and only related to Italian firms.

The analysis of the specific paths leading firms to default was conducted by the firm activity sector for the most frequent sectors. Table 1 shows the sample summary statistics

The underlying implicit assumption of the distinction among sectors is that different assets induce different balance-sheet equilibriums, which is especially true of small and medium-sized enterprises (SMEs), which appear as a heterogeneous whole that varies in size, organizational model, ownership structure, and propensity for growth and innovation (Karoui *et al.*, 2017). The presence of multiple characteristics is reflected in the formation of different economic and financial behaviors. Differences in ownership and organization and the company's life cycle can offer some explanatory elements (e.g. start-ups have different financing needs than mature, family-run companies). Other factors reside in the different geographical locations and various economic sectors in which SMEs operate, giving rise to more or less complicated financial behavior.

Companies whose main activity was not clearly identified (i.e. when the "other revenues/ sales" ratio was greater than 1) were excluded to maintain consistency with sectors.

Table 1.
Code and frequencies of the most frequent activity categories

AtecoCode	Description	Freq.	Percent	No default	Default
46	Wholesale – excluding cars and motorcycles	9,258	15.24	8,842	416
25	Manufacture of metal products – excl. machinery and equipment	4,311	7.09	4,112	199
41	Construction of buildings	4,309	7.09	3,751	558
47	Retail – excluding cars and motorcycles	3,785	6.23	3,554	231
43	Specialized construction works	3,698	6.09	3,434	264
68	Real estate brokerage	3,144	5.17	2,883	261
45	Retail – cars and motorcycles	2,244	3.69	2,142	102
28	Manufacture of machinery and equipment	2,055	3.38	1,972	83

4. Results

Starting from the sample described in the previous section, a set logistic regression was performed for selecting the default-predicting variables for each sector. Table 2 shows the best results for each sector. It is important to highlight that the number of actual observations is lower than what is reported in Table 1 because the analysis only included those companies for which all the values for the considered variables were available.

Table 2 summarizes the results of the logistic regressions.

The values reported in Table 2 show that the estimated coefficients and significance were different for different sectors, confirming the differences highlighted above. For example, the return on investment (ROI) for sector 46 had a significance of 99% and a coefficient of −0.028; sector 25 had a significance of 99% but a coefficient of −0.043; and sector 41 had no significance and a coefficient of 0.000.

Concerning the second hypothesis set in Section 2, the results show that each considered business sector was characterized by different balance-sheet equilibriums and significance of financial ratios, so the validity of hypothesis H2 was confirmed.

These differences in the default-predicting variables among sectors suggest that even the default process can follow specific paths within each sector.

The list of significant variables for subsequent clustering was selected based on the previous logistic regressions. Tables 3 and 4 show the list of selected variables.

For a simpler reading of the results, the average value of the selected variables was also computed for the non-defaulted firms in each considered sector. The values shown in Table 5 confirm that the values of the variables were different across sectors (for example, added value on production value, inventories on total assets and inventory duration), showing how a different balance-sheet equilibrium characterizes each sector.

In the subsequent step, a clustering procedure was performed [3]. For each sector, the clustering procedure was set for yielding 15 clusters, reported in Table 6. As such output can often produce several clusters containing just one or two cases (therefore not defining an actual group), attention was focused on the five clusters which reported the highest number of cases.

For a clearer identification of each cluster's specificity, the average value for each considered variable was computed and compared with the average value and standard error of non-defaulted firms. Tables A3–A7 in Appendix show the results for each sector, while the captions at the bottom of each Table describe each cluster's characteristics.

Table 2. Coefficients for the significant variables on the logistic regressions, by sector

Ateco code	46	25	41	47	43	68	45	28
Number of balance sheets	3532	1913	911	1135	1299	322	756	889
Pseudo R^2	0.2661	0.3818	0.2967	0.3052	0.3072	0.3350	0.3092	0.3789
Continuous overdraft	0.012^{**}	0.022^{***}	0.007	0.015	0.012^{**}	0.010	0.010^{*}	0.053^{**}
Shareholder dummy	-0.732^{***}	-1.127^{***}	-0.212	-0.493	-0.834^{***}	-0.583	-0.453	-0.245
Credit line usage	0.025^{***}	0.023^{***}	0.028^{***}	0.046^{***}	0.027^{***}	0.032^{**}	0.030^{**}	0.023^{**}
Violation share	-0.025^{***}	-0.026^{***}	-0.011	-0.032^{*}	-0.027^{***}	0.028	-0.030^{**}	-0.019^{*}
Violation months	0.054	0.184^{***}	0.136^{***}	0.025	0.101^{*}	0.129	0.041	-0.042
Credit limit violation flag	0.615^{**}	-0.167	0.368	0.799^{***}	0.275	-1.376^{*}	0.499	0.949
Blank cheques	0.189^{***}	-0.117	0.179	0.215	0.345^{***}	2.578^{*}	-0.368	0.691
Degree of indebtedness	0.080^{***}	0.050	0.072	0.028	0.016	-0.118^{***}	-0.007	-0.026
Added value on production value	0.025^{**}	-0.022^{*}	0.000	0.004	0.018^{*}	-0.005	-0.062^{**}	-0.019
Net assets coverage	-0.005	0.372^{*}	95.549^{**}	-0.060	0.258	0.176	0.265	0.073
Bank debt on total liabilities	0.008	0.020^{**}	0.022^{***}	-0.013	-0.005	0.022^{**}	-0.002	0.048^{***}
ROA	-0.051^{***}	-0.076^{*}	-0.022	-0.025	0.018	-0.079	-0.103^{***}	-0.140^{**}
ROI	-0.028^{***}	-0.043^{***}	0.000	-0.012	-0.015^{*}	0.021	-0.022	-0.005
Tangible assets on Total assets	-0.016^{**}	-0.014^{*}	-0.021^{**}	0.000	0.001	-0.011	-0.020	-0.021
Inventories on Total assets	-0.018	-0.009	-0.025^{***}	-0.007	-0.001	-0.037^{**}	-0.004	-0.020
Invested Capital turnover	0.001	-0.001	-0.008^{*}	-0.001	0.008^{**}	-0.006	-0.008^{**}	-0.010
Financial autonomy	0.048	-0.341^{*}	-95.482^{**}	0.064	-0.327^{*}	-0.283	-0.290	-0.073
Shareholders' equity on equity and inventories	0.006	0.009	-0.017	-0.001	0.000	-0.039^{**}	0.019	-0.002
Payables to suppliers on shareholders' equity	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000^{***}
Inventory duration	0.002^{***}	0.006^{**}	0.000^{***}	0.000	0.000^{***}	0.000	0.000	0.001^{**}
Debt burden index	0.001^{*}	-0.003	-0.007^{*}	0.003	0.001	-0.001	-0.002	-0.014
Constant	-12.585^{***}	-8.194^{**}	-10.372^{**}	-8.762^{**}	-6.555^{***}	9.795^{**}	-1.871	-2.206

Note(s): ***signals the coefficient significance at 99%, **at 95% and * at 90%

Table 3.
Balance sheet variables

Variable name	Computation
Degree of indebtedness	Total debt/(total liabilities + equity)
Added value on production value	Added value/sales
Net assets coverage	(Equity+ long term liabilities)/net assets
Bank debt on total liabilities	Bank debt/total liabilities
Return on Assets	EBIT/total assets
Return on Investments	EBIT/invested capital
Tangible assets on Total assets	Tangible assets/total assets
Inventories on Total assets	Inventories/total assets
Invested Capital turnover	Sales/invested capital
Financial autonomy	Equity/total liabilities + equity
Shareholders' equity on equity and inventories	Shareholders' equity/equity + inventories
Payables to suppliers on shareholders' equity	Payables to suppliers/shareholders' equity
Inventory duration	(Inventory/sales) *360
Debt burden index	Financial costs/EBITDA

Table 4.
Bank variables

Variable name	Computation
Continuous overdraft	Number of days of continuous overdraft
Shareholder dummy	Dummy for the lending bank shareholders
Credit line usage	Credit line usage/credit line limit
Violation share	Violation of credit limit/credit line limit
Violation months	Number of months of credit limit violation
Credit limit violation flag	Credit limit violation over the year (flag)
Blank cheques	Number of blank cheques

Table 5.
Average value of the non-defaulted firms for the significant variables, by sector

Ateco code	46	25	41	47	43	68	45	28
Continuous overdraft	1.01	1.0	2.5	0.8	1.5	2.7	1.8	0.4
Shareholder dummy	0.43	0.5	0.4	0.4	0.3	0.3	0.4	0.4
Credit line usage	46.24	41.8	65.3	55.1	48.7	59.1	64.2	37.4
Violation share	8.19	8.9	2.5	2.0	4.6	1.6	13.5	6.1
Violation months	1.10	1.3	1.6	1.4	1.5	1.6	1.4	1.1
Credit limit violation flag	0.3	0.4	0.5	0.4	0.4	0.4	0.4	0.3
Blank cheques	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0
Degree of indebtedness	74.0	68.6	77.8	73.3	73.1	67.8	73.7	68.5
Added value on production value	23.0	45.7	59.7	27.3	47.1	65.2	25.4	40.2
Net assets coverage	20.0	21.2	18.2	19.1	17.5	30.1	19.0	22.8
Bank debt on total liabilities	32.6	32.1	36.0	30.4	30.6	39.9	35.8	27.2
ROA	1.6	2.3	1.0	0.7	2.4	0.3	0.9	2.1
ROI	16.3	20.5	28.0	18.1	24.3	11.7	19.6	20.5
Tangible assets on Total assets	19.3	31.4	17.2	27.5	23.6	42.4	25.1	22.4
Inventories on Total assets	22.1	14.2	49.8	36.3	16.9	33.5	34.6	22.7
Invested Capital turnover	147.5	104.3	69.8	153.4	107.8	27.5	170.8	107.7
Financial autonomy	20.8	22.2	18.3	20.2	18.6	30.4	20.1	23.5
Shareholders' equity on equity and inventories	47.6	58.0	31.5	36.9	53.4	54.9	38.0	48.8
Payables to suppliers on shareholders' equity	598.8	379.8	374.9	498.8	442.9	171.1	415.2	338.0
Inventory duration	81.3	57.4	6115.8	275.5	128.3	2476.0	99.1	96.6
Debt burden index	20.5	13.8	20.5	18.7	14.0	37.6	20.9	13.4

Table 6. Clusters for each sector

Sector	46 (Wholesale – excluding cars and motorcycles)			25 (Manufacture of metal products – excl. machinery and equipment)			41 (Construction of buildings)			47 (Retail – excluding cars and motorcycles)			43 (Specialized construction works)		
Clusters	Freq.	%	Cum.	Freq.	%	Cum.	Freq.	%	Cum.	Freq.	%	Cum.	Freq.	%	Cum.
1	*43*	*32.82*	32.82	*47*	*51.65*	51.65	*6*	*5.61*	5.61	4	7.02	7.02	*19*	*20.21*	20.21
2	*8*	*6.11*	38.93	*5*	*5.49*	57.14	*4*	*3.74*	9.35	1	1.75	8.77	*35*	*37.23*	57.45
3	*22*	*16.79*	55.73	1	1.10	58.24	*56*	*52.34*	61.68	*5*	*8.77*	17.54	*4*	*4.26*	61.7
4	1	0.76	56.49	2	2.20	60.44	*15*	*14.02*	75.7	*6*	*10.53*	28.07	*15*	*15.96*	77.66
5	*31*	*23.66*	80.15	4	4.40	64.84	*7*	*6.54*	82.24	3	5.26	33.33	*7*	*7.45*	85.11
6	1	0.76	80.92	*10*	*10.99*	75.82	3	2.80	85.05	*7*	*12.28*	45.61	1	1.06	86.17
7	1	0.76	81.68	1	1.10	76.92	*4*	*3.74*	88.79	*5*	*8.77*	54.39	1	1.06	87.23
8	1	0.76	82.44	*7*	*7.69*	84.62	*4*	*3.74*	92.52	1	1.75	56.14	2	2.13	89.36
9	2	1.53	83.97	2	2.20	86.81	1	0.93	93.46	2	3.51	59.65	2	2.13	91.49
10	5	3.82	87.79	*5*	*5.49*	92.31	2	1.87	95.33	*12*	*21.05*	80.7	1	1.06	92.55
11	*7*	*5.34*	93.13	1	1.10	93.41	1	0.93	96.26	3	5.26	85.96	2	2.13	94.68
12	3	2.29	95.42	2	2.20	95.6	1	0.93	97.2	1	1.75	87.72	1	1.06	95.74
13	2	1.53	96.95	1	1.10	96.7	1	0.93	98.13	*5*	*8.77*	96.49	2	2.13	97.87
14	3	2.29	99.24	2	2.20	98.9	1	0.93	99.07	1	1.75	98.25	1	1.06	98.94
15	1	0.76	100.00	1	1.10	100	1	0.93	100	1	1.75	100.00	1	1.06	100.00
Total	131	100.00		91	100.00		107	100.00		57	100.00		57	100.00	

Note(s): ^The selected clusters for each sector are the ones highlighted in italic

Default syndrome	Variables significantly higher than average	Variables significantly lower than average
Underperformer with high bank debt	Bank debt on total liabilities Credit line usage	ROA ROI
Undercapitalized	Degree of indebtedness	Net assets coverage Shareholders' equity on equity and inventories Financial autonomy
Negatively affected by inventories	Inventories on Total assets Inventory duration	Shareholders' equity on equity and inventories Invested Capital turnover
Supply dependent with low capitalization	Payables to suppliers on shareholders' equity Degree of indebtedness ROI	Net assets coverage Shareholders' equity on equity and inventories Financial autonomy ROA
Slow mover	Tangible fixed assets Financial autonomy Net assets coverage	Invested capital turnover Turnover of inventories ROI

Table 7.
Default syndromes and their characterizing variables

Table 7 shows the main syndromes and the most significant variables characterizing each syndrome. For example, the default syndrome labeled "underperformer with a high level of bank debt" is mainly characterized by higher financial exposures (i.e. degree of bank debt and tension in the use of credit lines), coupled with low returns (ROA and ROI).

This description gives significant hints in examining the key weaknesses determining each syndrome, and investigating how and why the firms affected by those syndromes would often default.

5. Discussion

As shown in the previous section, the results showed that SMEs could be affected by specific syndromes, signaled by typical symptoms, in the run-up to a default.

With reference to the first hypothesis set in Section 2, these results confirm that the paths leading firms to default are not signaling just one group, nor singularly differentiated, but can be grouped into significant groups, each one characterized by some differences from healthy firms in terms of significant variables, confirming the correctness of hypothesis H1.

This suggests that the ability to diagnose and treat the particular syndrome that a company is affected by (as early as possible) can allow for an effective treatment to reduce the insolvency risk. For example, any SME showing high bank debt and low returns will easily struggle to service the debt. The widening of credit lines can be a possible solution in the short term, but it must be coupled with important interventions to improve the cash flow generation, otherwise the firm will be weakened further over the medium term. From this point of view, the results of this study make a significant contribution. The identification of the main syndromes that lead to the insolvency of SMEs allows the identification of the company's weaknesses and directs more quickly the interventions necessary for the recovery.

The analysis by sectors highlighted that the main syndromes affecting SMEs are quite similar across sectors, even if specific equilibriums and values characterize each sector. This means that the syndromes (undercapitalization, underperformance with or without high financial indebtedness, supply dependence and inventories' weight) affect different sectors, but different balance sheet equilibria characterize each sector.

The results in Table 8 show five cluster groups, coming from different sectors, depicting specific syndromes affecting defaulted SMEs, different for each sector but recurrent over sectors, confirming the correctness of hypothesis H3 set in Section 2.

Sector	46	25	41	47	43	Syndrome
Group I	1	1	1, 5			Underperformer with high bank debt
Group II	5	6	3	4, 10	2	Undercapitalized
Group III	2, 3		4		3	Negatively affected by inventories
Group IV	11	8, 10	7, 8	13	4, 5	Supply dependent
Group V		2	2	6	1	Slow mover

Table 8. Clusters and recurrence of default syndromes over sectors

The financial symptoms of each group and related clusters are described below. For a better understanding of the management implications of this study, the results observed in each cluster are analyzed by relating them to the main causes of failure, as found in the literature (Kucher *et al.*, 2020; Mayr *et al.*, 2017, 2021), and identifying the possible interventions to mitigate the probability of default.

Group I: The first group comprises four clusters related to three sectors (46:1, 25:1, 41:1 and 41:5). In this group, firms are mainly characterized by a high level of bank debt and modest economic performance. Hence, the syndrome has been labeled as an "underperformer with a high level of bank debt". In all clusters, the combination of the two phenomena determines serious difficulties in the coverage of financial charges. For this group of companies, financial fragility translates into high cost pressure – one of the most common causes of bankruptcy (Kucher *et al.*, 2020). Therefore, cost management is the key to mitigating the risk of bankruptcy in the short term; over the longer horizon, investing in more profitable products and services is perhaps the most suitable solution for these firms.

Group II: The second group includes six clusters related to five sectors (46:5, 25:6, 41:3, 47:4, 47:10 and 43:2) with prominent dependence on debt, especially of a banking nature, and low levels of capitalization ("undercapitalization" syndrome). All clusters are characterized by high indebtedness and unsatisfactory assets coverage values, shareholders' equity, and financial autonomy. Lack of equity appears to be one of the most prevalent internal causes leading to SME bankruptcy because it limits the possibilities of going concerns and makes it difficult to access alternative sources of financing (Carter and van Auken, 2006; Ooghe and De Prijcker, 2008; Kucher *et al.*, 2020; Mayr *et al.*, 2021).

Group III: The third group corresponds to four clusters (46: 2, 46: 3, 41: 4 and 43: 3) whose firms are negatively affected by inventories. All clusters are mainly characterized by high values for inventories and low values for invested capital turnover. Poor warehouse turnover indicates a difficulty in managing stocks, possibly due to lower commercial and/or planning capacity. Poor working-capital management is one of the four main blocks of SME bankruptcy (Gaskill *et al.*, 1993; Kucher *et al.*, 2020) because it could lead to a strong dependence on the banking channel (all lending indicators of Group III are under stress).

Group IV: The fourth group, including eight clusters (46:11, 25:8, 25:10, 41:7, 41:8, 47:13, 43:4 and 43:5), revealed high values for "degree of indebtedness" and "payables to suppliers on shareholders' equity", and low values for "net asset coverage", "financial autonomy" and "shareholders' equity on equity and inventories". This group presents conflicting economic signals and interestingly reports high values of ROI. However, this return is not effective, as the consistent incidence of debt costs absorbs a significant part of the revenue so that the

ROA goes down to values lower than the average of healthy firms. This group syndrome can be labeled as "supply-dependent with low capitalization". Excessive dependence on the supply chain, accompanied by a weak financial structure, can become dangerous in the event of an economic slowdown or sudden changes in the competitive environment (Kucher *et al.*, 2020) due to unpredictable dynamics as the recent COVID-19 pandemic has demonstrated.

Group V: Composed of four clusters distributed in four sectors (25:2, 41:2, 47:6 and 43:1), the fifth group encompasses firms mainly characterized by a high incidence of tangible fixed assets and a balanced capital structure. However, the low results in terms of turnover and returns show that the core business cannot produce satisfactory revenue. Firms in this group can be classified as slow movers, highlighting how, even in the presence of a solid financial position, a low turnover leads to a lower than average economic performance (Kucher *et al.*, 2020; Mayr *et al.*, 2021).

Regardless of the group to which they belong, all companies in default show more intense use of credit lines and more problematic relationships with the banking system, as indicated by a high incidence of anomalies. Therefore, bank-firm relationship information can be helpful in the credit-monitoring phase, flagging up early warning signs. Given this, identifying typical syndromes will enable us to verify which syndrome and what kinds of weaknesses a firm is suffering from and define a set of targeted interventions aimed at reducing specific risks. Just as a good doctor identifies a patient's pathology in time, in the bank-business relationship, a good bank is not necessarily one that grants endless credit but one that knows how to find the right solution to solve a firm's actual problems. Identifying specific, recurrent syndromes across sectors simplifies the problem of around one dimension order and can facilitate a more straightforward diagnosis of firms' pathologies.

6. Conclusions and further analyses

In this study, the characteristics of defaulted firms were analyzed to verify whether common patterns characterized the road to default and to detect the main syndromes affecting SMEs.

This approach constitutes a new direction in terms of analyzing default risk, as, on the one hand, it shows that some key syndromes can be defined, and on the other hand, it describes those syndromes and suggests that different problems can affect firms and induce different default risks.

This risk-splitting introduces an innovative point of view and highlights the need for differentiated approaches when dealing with firms affected by different risks. Similar to what a good doctor does, a course of treatment can be decided after identifying the cause of the patient's problem, or, at least, the syndrome affecting him, and not just generically treating him as someone who is sick.

To do something similar when dealing with firms, two steps are needed. The first one entails describing key syndromes, that is to say, the set of associated symptoms that characterize each specific condition. The second step, which is not fully developed in this paper, concerns the search for the best treatment of each syndrome to help the company improve its performance.

In this study, the first step has been built by performing a cluster analysis of defaulted firms in the most significant economic sector and presenting the results in the form of "symptoms" (significant variations in balance-sheet variables in the year before the default).

In technical terms, performing the cluster analysis separately for each considered sector means that the analysis was conducted for different samples in a parallel way, cross-validating the results. The results show that the clusters (syndromes) identified in each economic sector, and characterized through the significant variables, can be found in other sectors. These results suggest that the same causes induce similar effects when applied to different sectors, even if specific equilibriums and values characterize each sector.

Finally, this new point of view enlarges the perspective on the problems that afflict SMEs and can lead to their default.

From an academic point of view, the introduction of a new approach to understanding the weaknesses of SMEs can facilitate the search for more targeted solutions based on objective data, such as those relating to financial symptoms. From the point of view of banks and companies, identification and description of specific syndromes can significantly affect the bank-company relationship, allowing more effective interventions and possibly reducing the specific risk to which each company is subject. Identifying specific risks allows both banks and companies to intervene in time to activate preventive behaviors that can lead to better results in improving firms' prospective sustainability, compared to more generic treatments or treatments not adopted on time. In this perspective, the quantitative description of the syndromes facilitates activation of an automatic alarm system, which can also be applied in areas of the credit relationship where the availability of information on a company may be more limited due to a less profound credit relationship (e.g. fintech companies).

Although this study has obtained valuable insights into bankruptcy and default-prediction domains, our analysis was of an exploratory nature. The limits of the dataset and confidentiality restrictions meant that we could not allow for testing whether the syndromes were related to specific phases of a firm's life cycle or were dictated by the personal attitudes of a firm's senior management. The possibility of linking these two kinds of information would allow for a better understanding of the determinants leading firms down specific default paths, thereby (staying with the medicine metaphor) enabling decision-makers to go from identification of a syndrome to the identification of a disease. This approach would also benefit from a panel study about the evolution of key syndromes over time to generate more detailed information. Future analyses, possibly integrating a quantitative and qualitative approach, and coupling the grouping strategy used in this study to define the sample with an in-depth analysis of each defaulted firm in the sample, could help to verify whether these syndromes recur across countries and whether a common causality for each group can be confirmed.

Notes

1. https://www.lexico.com/definition/syndrome
2. The countries are Greece, Portugal, Slovenia, Malta, Cyprus, Spain, Italy, France, the UK, Belgium, Germany, Finland, Austria and the Netherlands; source: Eurostat.
3. We did this by the "complete linkage" clustering procedure with a Euclidean (dis)similarity measure, implemented in Stata13.0.

References

Altman, E.I. (1968), "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *The Journal of Finance*, Vol. 23 No. 4, pp. 589-609.

Altman, E.I. (2002), "Revisiting credit scoring models in a Basel 2 environment", NYU Finance Working Papers.

Altman, E.I. and Sabato, G. (2007), "Modelling credit risk for SMEs: evidence from the US market", *Abacus*, Vol. 33 No. 3, pp. 332-357.

Altman, E.I., Sabato, G. and Wilson, N. (2010), "The value of non-financial information in small and medium-sized enterprise risk management", *Journal of Credit Risk*, Vol. 6, pp. 1-33.

Andriosopoulos, D., Doumpos, M., Pardalos, P.M. and Zopounidis, C. (2019), "Computational approaches and data analytics in financial services: a literature review", *Journal of the Operational Research Society*, Vol. 70 No. 10, pp. 1581-1599.

Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J. and Vanthiene, J. (2003), "Benchmarking state-of-the-art classification algorithms for credit scoring", *Journal of the Operational Research Society*, Vol. 54, pp. 627-635.

Bauer, J. and Agarwal, V. (2014), "Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test", *Journal of Banking and Finance*, Vol. 40, pp. 432-442.

Beaver, W.H. (1966), "Financial ratios as predictors of failure", *Journal of Accounting Research*, Vol. 4, pp. 71-111.

Beck, T. and Demirgüc-Kunt, A. (2006), "Small and medium-size enterprises: access to finance as a growth constraint", *Journal of Banking and Finance*, Vol. 30, pp. 2931-2943.

Beck, T., Demirgüç-Kunt, A. and Martínez Pería, M.S. (2011), "Bank financing for SMEs: evidence across countries and bank ownership types", *Journal of Financial Services Research*, Vol. 39, pp. 35-54.

Beck, T., Degryse, H., Haas, R.D. and van Horen, N. (2018), "When arm's length is too far: relationship banking over the credit cycle", *Journal of Financial Economics*, Vol. 127 No. 1, pp. 174-196.

Behr, P. and Guttler, A. (2007), "Credit risk assessment and relationship lending: an empirical analysis of German small and medium-sized enterprises", *Journal of Small Business Management*, Vol. 45 No. 2, pp. 194-213.

Bellovary, J.L., Giacomino, D.E. and Akers, M.D. (2007), "A review of bankruptcy prediction studies: 1930 to present", *Journal of Financial Education*, Vol. 33, pp. 1-42.

Berger, A.N. and Udell, G.F. (2002), "Small business credit availability and relationship lending: the importance of bank organisational structure", *The Economic Journal*, Vol. 112 No. 477, F32–F53.

Berry, M. (2000), *Mastering Data Mining*, John Wiley & Sons, Hoboken, NY.

Blum, M.P. (1974), "Failing company discriminant analysis", *Journal of Accounting Research*, Vol. 12 No. 1, pp. 1-25.

Brealey, R.A., Myers, S.C. and Allen, F. (2020), *Principles of Corporate Finance*, 13th ed., McGraw-Hill, New York.

Brighi, P., Lucarelli, C. and Venturelli, V. (2019), "Predictive strength of lending technologies in funding SMEs", *Journal of Small Business Management*, Vol. 57 No. 4, pp. 1350-1377.

Burgstahler, D.C., Hail, L. and Leuz, C. (2006), "The importance of reporting incentives: earnings management in European private and public firms", *The Accounting Review*, Vol. 81 No. 5, pp. 983-1016.

Calabrese, R., Marra, G. and Osmetti, S. (2016), "Bankruptcy prediction of small and medium enterprises using a flexible binary generalized extreme value model", *Journal of the Operational Research Society*, Vol. 67 No. 4, pp. 604-615.

Canto-Cuevas, F.-J., Palacìn-Sanchez, M.-P. and Di Pietro, F. (2019), "Trade credit as a sustainable resource during an SME's life cycle", *Sustainability*, Vol. 11 No. 3, p. 670.

Carter, R. and van Auken, H. (2006), "Small firm bankruptcy", *Journal of Small Business Management*, Vol. 44 No. 4, pp. 493-512.

Cheng, C., Jones, S. and Moser, W. (2018), "Abnormal trading behavior of specific types of shareholders before US firm bankruptcy and its implications for firm bankruptcy prediction", *Journal of Business Finance and Accounting*, Vol. 45 No. 9, pp. 1100-1138.

Crook, J., Edelman, D. and Thomas, L. (2007), "Recent developments in consumer credit risk assessment", *European Journal of Operational Research*, Vol. 18, pp. 1447-1465.

Delli Gatti, D., Gallegati, M., Giulioni, G. and Palestrini, A. (2003), "Financial fragility, patterns of firms' entry and exit and aggregate dynamics", *Journal of Economic Behavior and Organization*, Vol. 51 No. 1, pp. 79-97.

Duarte, F.D., Gama, A.P.M. and Gulamhussen, M.A. (2018), "Defaults in bank loans to SMEs during the financial crisis", *Small Business Economics*, Vol. 51 No. 3, pp. 591-608.

Edminster, R.O. (1972), "An empirical test of financial ratio analysis for small business failure prediction", *Journal of Financial and Quantitative Analysis*, Vol. 7, pp. 1477-1493.

Eldridge, D., Nisar, T.M. and Torchia, M. (2021), "What impact does equity crowdfunding have on SME innovation and growth? An empirical study", *Small Business Economics*, Vol. 56, pp. 105-120.

European Commission, Executive Agency for Small and Medium-sized Enterprises (2020), *Annual Report on European SMEs 2018/2019: Research & Development and Innovation by SMEs*, Publications Office, Luxembourg.

Fiordelisi, F., Monferrà, S. and Sampagnaro, G. (2014), "Relationship lending and credit quality", *Journal of Financial Services Research*, Vol. 46 No. 3, pp. 295-315.

Franquesa, J. and Vera, D. (2021), "Small business debt financing: the effect of lender structural complexity", *Journal of Small Business and Enterprise Development*, Vol. 28 No. 3, pp. 456-474.

Gaskill, L.R., Van Auken, H.E. and Manning, R.A. (1993), "A factor Analytic study of the perceived causes of small business failure", *Journal of Small Business Management*, Vol. 34 No. 4, pp. 8-31.

Ghosal, V. and Ye, Y. (2019), "The impact of uncertainty on the number of businesses", *Journal of Economics and Business*, Vol. 105, p. 105840.

Grice, J.S. and Ingram, R.W. (2001), "Tests of the generalizability of Altman's bankruptcy prediction model", *Journal of Business Research*, Vol. 54 No. 1, pp. 53-61.

Habib, A. and Hasan, M.M. (2017), "Firm life cycle, corporate risk-taking and investor sentiment", *Accounting and Finance*, Vol. 57 No. 2, pp. 465-497.

Huang, S.C., Tang, Y.C., Lee, C.W. and Chang, M.-J. (2012), "Kernel local Fisher discriminant analysis-based manifold-regularized SVM model for financial distress predictions", *Expert Systems with Applications*, Vol. 39, pp. 3855-3861.

Jones, S. and Wang, G. (2019), "Predicting private company failure: a multi-class analysis", *Journal of International Financial Markets, Institutions and Money*, Vol. 61, pp. 161-188.

Jones, S., Johnstone, D. and Wilson, R. (2017), "Predicting corporate bankruptcy: an evaluation of alternative statistical frameworks", *Journal of Business Finance & Accounting*, Vol. 44 Nos 1-2, pp. 3-34.

Jorion, P. and Zhang, G. (2009), "Credit contagion from counterpart risk", *The Journal of Finance*, Vol. 64 No. 5, pp. 2053-2087.

Karoui, L., Khlif, W. and Ingley, C. (2017), "SME heterogeneity and board configurations: an empirical typology", *Journal of Small Business and Enterprise Development*, Vol. 24 No. 3, pp. 545-561.

Khelil, N. (2016), "The many faces of entrepreneurial failure: insights from an empirical taxonomy", *Journal of Business Venturing*, Vol. 31 No. 1, pp. 72-94.

Kücher, A., Mayr, S., Mitter, C., Duller, C. and Feldbauer-Durstmüller, B. (2020), "Firm age dynamics and causes of corporate bankruptcy: age-dependent explanations for business failure", *Review of Managerial Science*, Vol. 14, pp. 633-661.

Laitinen, E.K. and Gin Chong, H. (1999), "Early-warning system for crisis in SMEs: preliminary evidence from Finland and the UK", *Journal of Small Business and Enterprise Development*, Vol. 6 No. 1, pp. 89-102.

Laitinen, E.K., Lukason, O. and Suvas, A. (2014), "Are firm failure processes different? Evidence from seven countries", *Investment Management and Financial Innovations*, Vol. 11 No. 4, pp. 212-222.

Lim, M.K. and Sohn, S.Y. (2007), "Cluster-based dynamic scoring model", *Expert Systems with Applications*, Vol. 32, pp. 427-431.

Liu, Z., Shang, J., Wu and Chen, P.-Y. (2020), "Social collateral, soft information and online peer-to-peer lending a theoretical model", *European Journal of Operational Research*, Vol. 281 No. 2, pp. 428-438.

Louzada, F., Ara, A. and Fernandes, G.B. (2016), "Classification methods applied to credit scoring: systematic review and overall comparison", *Surveys in Operations Research and Management Science*, Vol. 21 No. 2, pp. 117-134.

Mayr, S., Mitter, C. and Aichmayr, A. (2017), "Corporate crisis and sustainable reorganization: evidence from bankrupt Austrian SMEs", *Journal of Small Business Management*, Vol. 55 No. 1, pp. 108-127.

Mayr, S., Mitter, C., Kücher, A. and Duller, C. (2021), "Entrepreneur characteristics and differences in reasons for business failure: evidence from bankrupt Austrian SMEs", *Journal of Small Business and Entrepreneurship*, Vol. 33 No. 5, pp. 539-558, doi: 10.1080/08276331.2020.1786647.

McGuinness, G., Hogan, T. and Powell, R. (2018), "European trade credit use and SME survival", *Journal of Corporate Finance*, Vol. 49, pp. 81-103.

Norden, L. and Weber, M. (2010), "Credit line usage, checking account activity, and default risk of bank borrowers", *Review of Financial Studies*, Vol. 23 No. 10, pp. 3665-3699.

Ohlson, J.A. (1980), "Financial ratios and the probabilistic prediction of bankruptcy", *Journal of Accounting Research*, Vol. 18, pp. 109-131.

Ooghe, H. and De Prijcker, S. (2008), "Failure processes and causes of company bankruptcy: a typology", *Management Decision*, Vol. 46 No. 2, pp. 223-242.

Peel, M.J. and Peel, D.A. (1989), "A multi-logit approach to predicting corporate failure: some evidence from the UK corporate sector", *Omega International Journal of Management Science*, Vol. 16 No. 4, pp. 309-318.

Pindado, J., Rodrigues, L. and de la Torre, C. (2008), "Estimating financial distress likelihood", *Journal of Business Research*, Vol. 61 No. 9, pp. 995-1003.

Ravi Kumar, P. and Ravi, V. (2007), "Bankruptcy prediction in banks and firms via statistical and intelligent techniques - a review", *European Journal of Operational Research*, Vol. 180, pp. 1-28.

Ryan, R.M., O'Toole, C. and McCann, F. (2014), "Does bank market power affect SME financing constraints?", *Journal of Banking and Finance*, Vol. 49, pp. 495-505.

Sette, E. and Gobbi, G. (2015), "Relationship lending during a financial crisis", *Journal of the European Economic Association*, Vol. 13 No. 3, pp. 453-481.

Shumway, T. (2001), "Forecasting bankruptcy more accurately: a simple hazard model", *The Journal of Business*, Vol. 74 No. 1, pp. 101-124.

Sigrist, F. and Hirnschall, C. (2019), "Grabit: gradient tree-boosted Tobit models for default prediction", *Journal of Banking and Finance*, Vol. 102, pp. 177-192.

Stiglitz, J. and Weiss, A. (1981), "Credit rationing in markets with imperfect information", *American Economic Review*, Vol. 71 No. 3, pp. 93-410.

Strotmann, H. (2007), "Entrepreneurial survival", *Small Business Economics*, Vol. 28 No. 1, pp. 87-104.

Tian, S., Yu, Y. and Guo, H. (2015), "Variable selection and corporate bankruptcy forecasts", *Journal of Banking and Finance*, Vol. 52 No. C, pp. 89-100.

Veganzones, D. and Severin, E. (2021), "Corporate failure prediction models in the twenty-first century: a review", *European Business Review*, Vol. 33 No. 2, pp. 204-226.

Xia, Y., Liu, C., Li, Y. and Liu, N. (2017), "A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring", *Expert Systems with Applications*, Vol. 78, pp. 225-241.

Corresponding author
Stefano Zedda can be contacted at: szedda@unica.it

Appendix

Variables	Variables
Other revenues on production value	Medium and long-term debt
Cash flow on production value	Debt to banks
Intangible asset on production value	Financial dependence index
Immediate liquidity on total assets	Short-term debt
Added value on production value	Quick ratio
(Current assets-inventories)/current liabilities	Debt burden index
Depreciation and devaluation on costs	Index of rigidity of assets
Financial autonomy	EBITDA on revenues
Payables to banks on current assets	Financial interest on revenues
Unit cash flow (on total revenues)	Financial interest on added value
Net active coverage	Shareholders' equity on (long-term equity and payables)
Fixed asset coverage	Shareholders' equity on equity and inventories
Financial coverage index	Leverage
Labor cost on revenues	Gross operating profitability (EBITDA on production value)
Labor cost on production value	Inventories on Total assets
Unit labor cost	Inventories on Short-term debt
Credits on Total assets	Inventories on Medium/long-term debt
Current assets/current liabilities	Inventories on Total debt
Short-term payables on amounts due to banks	Inventories on Bank debt
Short-term payables on Net worth	ROA (Return on Asset)
Payables on short-term debts	ROD (Return on Debt)
Debts on Net worth	ROE (Return on Equity)
Payables to suppliers on Net worth (shareholders' equity)	ROI (Return on Investment)
Payables to suppliers on Total debt	ROS (Return on Sales)
Inventory duration	Capital invested rotation
Degree of indebtedness	Working capital turnover
Intangible assets on shareholders' equity	Inventory turnover
Intangible assets on Total assets	Value of production on Total Assets
Tangible assets on shareholders' equity	Production value on Revenues
Tangible assets on Total assets	Production value on Inventory

Table A1. List of all available balance sheet variables (#60)

Table A2.
List of all available bank variables (#18)

Variables
Shareholder dummy
days of overdraft
days of continuous overdraft
Credit line usage
Violation share
% debit amounts on credit amounts
Average annual credit amount
Average annual debit amount
Average annual credit/debit amount
% annual number debt transactions debit on number credit transactions
% debit amounts on annual debit/credit amounts
% credit amounts on annual debit/credit amounts
Total unpaid checks at first presentation
Number of blank cheques
Total number of unpaid checks
Violation months
Credit limit violation flag
Continuous violation months

Table A3.
Clusters variables values for Sector 46 (Wholesale – excluding cars and motorcycles)

Cluster	No default		46_1	46_5	46_3	46_2	46_11
Cases: 3.401	Mean	Std. Err.	43	31	22	8	7
Continuous overdraft	1.01	0.19	6.12	18.52	11.14	14.75	16.00
Shareholder dummy	0.43	0.01	0.23	0.26	0.18	0.25	0.14
Credit line usage	46.24	1.75	81.93	91.53	107.04	108.91	95.87
Violation share	8.19	1.58	5.23	14.49	20.73	24.08	13.95
Violation months	1.10	0.04	3.14	3.81	5.64	4.63	5.71
Credit limit violation flag	0.34	0.01	0.72	0.77	1.00	0.63	0.86
Blank cheques	0.05	0.01	0.23	1.23	0.82	0.00	1.71
Degree of indebtedness	74.05	0.30	76.69	91.29	89.12	72.66	93.44
Added value on production value	23.01	0.26	28.56	20.10	26.09	37.88	23.10
Net assets coverage	20.00	0.28	18.74	6.25	6.97	21.51	3.54
Bank debt on total liabilities	32.60	0.30	47.19	38.69	49.56	48.89	32.68
ROA	1.63	0.08	−1.06	0.66	−0.14	−1.19	−2.48
ROI	16.32	0.32	14.08	14.72	12.72	13.86	18.74
Tangible assets on Total assets	19.33	0.32	23.74	10.39	14.22	14.43	6.00
Inventories on Total assets	22.12	0.28	17.33	22.35	31.78	47.98	20.63
Invested Capital turnover	147.47	1.61	105.91	144.92	85.66	47.34	145.27
Financial autonomy	20.79	0.29	19.90	6.38	7.27	22.35	3.58
Shareholders' equity on equity and inventories	47.57	0.45	53.76	31.65	21.30	31.32	20.29
Payables to suppliers on shareholders' equity	598.81	47.23	116.75	645.31	409.61	77.48	1214.18
Inventory duration	81.27	1.79	78.50	76.05	154.65	374.16	50.58
Debt burden index	20.53	0.93	34.20	25.27	57.73	62.56	38.88

For sector 46, the clusters with the highest number of firms in default are those identified by numbers 1, 2, 3, 5 and 11.

The first cluster (46_1) is mainly characterized by high values for bank debt and low values for profitability measures. Coverage of financial charges is weak, while use of bank credit lines is excessive, as evidenced by the number of months of annual overrun. The syndrome is therefore that of an "underperformer with a high level of bank debt".

The second cluster (46_5) is mainly characterized by high values for the degree of indebtedness, low values for shareholders' equity, and low levels of financial autonomy. The asset coverage is not balanced, mainly due to low levels of capitalization. The firms in this cluster can be identified as "undercapitalized".

The third cluster (46_3) is characterized by high values for inventory and bank dependency, and low values for capitalization, financial autonomy and capital turnover. The low turnover of the warehouse indicates difficulty in managing stocks and leads to greater dependence on the banking channel. The firms in this cluster are "negatively affected by inventories".

The fourth cluster (46_2) is characterized by high values for added value, bank debt and inventory duration, and low values for payables to suppliers and capital turnover. It has similar characteristics to the previous cluster, so firms in this cluster are "negatively affected by inventories".

The fifth cluster (46_11) is characterized by a high degree of indebtedness and payables to suppliers, and low values for asset coverage, financial autonomy and inventory duration. It shows mixed economic signals (ROI shows positive values, while ROA is negative) and a strong incidence of operating and financial debts. The firms in this cluster are "supply-dependent with low capitalization".

Cluster	No default		25_1	25_6	25_8	25_2	25_10
Cases: 1.822	Mean	Std. Err.	47	10	7	5	5
Continuous overdraft	1.0	0.2	12.7	23.5	26.9	3.2	0.0
Shareholder dummy	0.5	0.0	0.2	0.3	0.4	0.2	0.4
Credit line usage	41.8	2.3	96.1	117.6	104.0	93.9	76.4
Violation share	8.9	2.0	19.2	44.1	16.4	17.2	6.7
Violation months	1.3	0.1	4.6	4.5	6.3	5.0	4.6
Credit limit violation flag	0.4	0.0	0.8	0.9	1.0	0.8	0.8
Blank cheques	0.0	0.0	0.1	0.0	0.1	0.4	0.0
Degree of indebtedness	68.6	0.4	76.0	91.7	89.4	85.7	89.9
Added value on production value	45.7	0.5	34.9	40.2	46.7	31.4	29.5
Net assets coverage	21.2	0.4	18.6	4.0	2.5	6.8	1.9
Bank debt on total liabilities	32.1	0.4	46.4	49.1	42.6	42.7	48.8
ROA	2.3	0.1	0.0	−0.6	−3.1	1.3	0.1
ROI	20.5	0.5	8.5	13.8	20.5	5.1	5.3
Tangible assets on Total assets	31.4	0.5	30.0	23.8	22.9	34.0	18.0
Inventories on Total assets	14.2	0.3	19.4	27.2	18.9	8.6	19.0
Invested Capital turnover	104.3	1.0	77.8	83.0	92.3	69.7	111.9
Financial autonomy	22.2	0.4	18.8	4.2	2.6	7.7	1.9
Shareholders' equity on equity and inventories	58.0	0.6	51.3	18.1	14.7	46.9	19.3
Payables to suppliers on shareholders' equity	379.8	37.1	148.7	773.6	1199.3	459.7	1715.5
Inventory duration	57.4	1.4	94.4	134.9	78.7	37.5	58.6
Debt burden index	13.8	0.6	28.0	34.6	23.2	31.1	45.2

Table A4. Clusters variables values for Sector 25 (Manufacture of metal products – excl. machinery and equipment)

For sector 25, the clusters with the highest number of firms in default are those identified by numbers 1, 2, 6, 8 and 10.

The first cluster (25_1) is mainly characterized by high values for bank debt on total liabilities, inventory duration and debt burden index, and low values for added value on production value, ROA, ROI and payables to suppliers on shareholders' equity. It shows similar characteristics to cluster 46_1, and these firms can therefore be defined as "underperformers with a high level of bank debt".

The second cluster (25_6) is mainly characterized by a high degree of debt, both banking and commercial, and by a long duration of stocks, levels of equity coverage, financial autonomy and equity are low. Despite the relative incidence of tangible fixed assets, low capitalization does not help firms properly cover assets. Therefore, this syndrome is termed "undercapitalized".

The third cluster (25_8) shows mixed economic signals (ROI value is, on average, the same as for healthy firms, while ROA is particularly negative), low capitalization and a strong incidence of operating debts, which affects its financial autonomy. The firms in this cluster are characterized as being "supply-dependent with low capitalization".

The fourth cluster (25_2) is mainly characterized by a high incidence of tangible fixed assets, which negatively affects the invested capital turnover. The economic results appear weak. The firms in this cluster can be described as having "slow mover" syndrome.

The fifth cluster (25_10) is mainly characterized by high values for payables to suppliers and debt burden index, and very low values for shareholder equity, financial autonomy and net asset coverage. All the indicators show weak economic performance on the part of firms in this cluster. The syndrome for this cluster is "supply-dependent with low capitalization".

Table A5.
Clusters variables values for Sector 41 (Construction of buildings)

Cluster Cases: 804	No default Mean	 Std.Err	41_3 56	41_4 15	41_5 7	41_1 6	41_2 4	41_7 4	41_8 4
Continuous overdraft	2.5	0.6	18.0	7.6	98.7	28.5	42.0	0.0	18.3
Shareholder dummy	0.4	0.0	0.3	0.4	0.3	0.0	0.5	0.3	0.5
Credit line usage	65.3	1.2	96.3	97.0	113.6	96.1	94.3	87.6	107.8
Violation share	2.5	0.4	8.3	4.0	20.8	3.5	1.6	0.6	13.6
Violation months	1.6	0.1	4.6	4.9	5.1	4.0	5.3	3.8	7.5
Credit limit violation flag	0.5	0.0	0.9	0.8	0.9	0.8	0.5	1.0	1.0
Blank cheques	0.0	0.0	0.3	0.1	0.0	0.0	0.5	0.0	0.3
Degree of indebtedness	77.8	0.6	88.1	84.4	85.5	89.7	82.1	95.2	96.8
Added value on production value	59.7	0.9	61.2	69.9	65.4	87.6	78.0	66.5	45.9
Net assets coverage	18.2	0.6	9.2	13.7	12.1	9.7	15.0	1.6	0.6
Bank debt on total liabilities	36.0	0.8	50.3	56.2	51.6	67.0	56.0	54.8	30.3
ROA	1.0	0.1	0.8	−1.7	−1.6	−0.3	−1.6	−3.4	−0.6
ROI	28.0	1.6	24.7	17.4	7.6	7.8	26.7	37.6	30.6
Tangible assets on Total assets	17.2	0.7	14.3	5.8	14.3	11.4	39.6	13.3	8.5
Inventories on Total assets	49.8	1.1	48.3	80.3	71.6	61.0	50.6	33.8	52.4
Invested Capital turnover	69.8	2.4	56.6	26.0	13.7	4.2	5.2	86.1	81.0
Financial autonomy	18.3	0.6	9.2	13.7	12.1	9.7	15.0	1.6	0.6
Shareholders' equity on equity and inventories	31.5	1.0	22.9	12.5	14.8	21.0	30.9	9.2	1.6
Payables to suppliers on shareholders' equity	374.9	40.6	355.2	237.9	83.5	27.6	17.1	1906.5	5029.6
Inventory duration	6115.8	3819.9	378.7	1172.5	1926.0	5236.0	3474.8	163.7	322.0
Debt burden index	20.5	1.7	21.5	24.1	41.4	40.4	70.5	9.5	25.2

For sector 41, the clusters with the highest number of firms in default are those identified by numbers 1, 3, 4 and 5. Since some clusters only include four cases, the selection also covers clusters 2, 7 and 8.

The first cluster (41_3) presents higher dependence on debt, especially of a banking nature, and lower levels of capitalization. The firms in this cluster have the "undercapitalized" syndrome.

The second cluster (41_4) is characterized by high values for bank debt and inventories, and low values for invested capital turnover and tangible assets. It is similar to clusters 46_2 and 46_3 and can therefore be classified as "negatively affected by inventories".

The third cluster (41_5) is characterized by a strong incidence of financial charges on EBITDA due to both a higher incidence of bank debt and a modest ability to generate income (the relative ROI and ROA values are low even in comparison with other clusters in the sector). The firms in this cluster are "underperformers with a high level of bank debt".

The fourth cluster (41_1) has several similarities to the previous cluster, and both have modest income performance and consequent difficulty in covering financial charges given the high level of bank debt. As with the previous one, firms in this cluster can be classified as “underperformers with a high level of bank debt”.

The fifth cluster (41_2) is characterized by high values for tangible assets and debt burden index, and low values for invested capital turnover and payables to suppliers. This cluster comprises firms classified as “slow movers”.

The sixth (41_7) and seventh clusters (41_8) are characterized by high values for ROI and payables to suppliers, and low values for financial autonomy, asset coverage and equity. The firms in these two clusters are “supply-dependent with low capitalization”.

Cluster	No default		47_10	47_6	47_4	47_3	47_7	47_13
Cases: 1.078	Mean	Std. Err.	12	7	6	5	5	5
Continuous overdraft	0.8	0.3	6.6	0.4	10.5	2.8	43.6	10.6
Shareholder dummy	0.4	0.0	0.2	0.3	0.7	0.0	0.4	0.4
Credit line usage	55.1	1.1	92.4	97.2	92.6	85.2	102.8	98.9
Violation share	2.0	0.2	6.2	6.1	3.2	13.3	10.4	3.9
Violation months	1.4	0.1	3.1	4.9	5.2	5.6	4.8	6.8
Credit limit violation flag	0.4	0.0	0.8	0.9	0.8	1.0	0.8	1.0
Blank cheques	0.1	0.0	0.8	0.0	2.5	2.0	0.2	0.2
Degree of indebtedness	73.3	0.6	91.6	61.7	87.5	87.3	74.2	94.7
Added value on production value	27.3	0.5	22.2	28.7	27.7	16.0	30.8	34.1
Net assets coverage	19.1	0.5	4.8	29.2	8.5	10.1	20.8	2.2
Bank debt on total liabilities	30.4	0.6	45.7	29.4	41.0	14.8	26.5	39.6
ROA	0.7	0.1	−0.9	−0.4	0.0	−0.1	−2.1	−4.3
ROI	18.1	0.8	13.4	5.5	7.1	4.9	26.8	28.7
Tangible assets on Total assets	27.5	0.7	32.9	33.8	17.0	14.8	23.7	10.4
Inventories on Total assets	36.3	0.8	15.0	32.1	39.7	71.4	23.5	58.5
Invested Capital turnover	153.4	4.1	144.5	53.5	71.5	76.3	154.7	93.9
Financial autonomy	20.2	0.5	5.0	31.5	9.4	10.3	20.8	2.2
Shareholders’ equity on equity and inventories	36.9	0.8	32.7	50.2	18.7	12.3	51.3	3.5
Payables to suppliers on shareholders’ equity	498.8	31.4	706.6	84.8	314.7	520.3	156.9	1533.3
Inventory duration	275.5	124.4	50.9	206.8	218.9	342.4	57.1	267.4
Debt burden index	18.7	0.8	26.1	59.2	46.6	29.5	10.9	36.6

Table A6. Clusters variables values for Sector 47 (Retail – excluding cars and motorcycles)

For sector 47, the clusters with the highest number of firms in default are those identified by numbers 4, 6 and 10. As with sector 41, as three clusters only included five cases (namely clusters 3, 7 and 13), the selection includes all these clusters.

The first cluster (47_10) is mainly characterized by high values for degree of indebtedness, and low values for net asset coverage and financial autonomy. The amount of indebtedness is high, both to banks and suppliers. The firms in this cluster are “undercapitalized”.

The second cluster (47_6) is mainly characterized by a balanced capital structure. The level of capitalization contains a degree of financial leverage and favors more than adequate coverage of assets, despite the consistency of fixed assets. However, the main economic indicators are weak, and these firms struggle to properly cover financial charges. The firms in this cluster can be described as having “slow mover” syndrome.

The third cluster (47_4) shows similar characteristics to those of cluster 47_10, having in common a high ratio of debt, low capital turnover and low level of financial autonomy. Compared to the homologous cluster, the asset side is more elastic, thanks to the greater weight of current assets. The firms in this cluster can also be classified as “undercapitalized”.

For the fourth cluster (47_3) and the fifth cluster (47_7), no specific classification is suggested.

The sixth cluster (47_13) is mainly characterized by high return values (ROI and value added on production value) and debt (payables to suppliers on shareholders' equity and degree of indebtedness) but also by low shareholders' equity on equity and inventories and low levels financial autonomy. It scores the worst results for ROA. The syndrome is "supply-dependent with low capitalization".

Cluster	No default		43_2	43_1	43_4	43_5	43_3
Cases: 1.205	Mean	Std. Err.	35	19	15	7	4
Continuous overdraft	1.5	0.4	24.6	31.8	3.9	5.1	0.5
Shareholder dummy	0.3	0.0	0.2	0.3	0.1	0.1	0.3
Credit Line Usage	48.7	1.4	98.0	95.6	97.8	99.2	96.6
Violation share	4.6	0.8	17.8	11.9	18.5	15.1	6.5
Violation months	1.5	0.1	4.8	4.8	4.6	6.0	7.0
Credit limit violation flag	0.4	0.0	0.8	0.9	0.9	1.0	1.0
Blank cheques	0.0	0.0	0.6	0.8	0.0	0.1	2.8
Degree of indebtedness	73.1	0.5	86.5	74.3	91.6	89.4	82.0
Added value on production value	47.1	0.6	47.5	51.5	57.2	65.7	57.3
Net assets coverage	17.5	0.4	7.6	15.5	3.2	1.7	13.0
Bank debt on total liabilities	30.6	0.5	42.2	38.4	41.4	48.8	33.4
ROA	2.4	0.2	−0.3	0.2	−0.2	−2.9	0.8
ROI	24.3	0.8	21.9	17.5	31.4	35.2	11.2
Tangible assets on Total assets	23.6	0.6	15.1	33.7	17.1	19.7	12.3
Inventories on Total assets	16.9	0.5	19.1	9.9	16.6	12.1	46.7
Invested Capital turnover	107.8	1.6	114.9	81.2	114.1	96.1	42.6
Financial autonomy	18.6	0.4	7.8	16.0	3.2	1.7	13.1
Shareholders' equity on equity and inventories	53.4	0.9	37.2	63.3	23.3	26.0	27.3
Payables to suppliers on shareholders' equity	442.9	27.5	446.6	136.8	1065.2	1730.2	198.8
Inventory duration	128.3	34.7	86.6	50.6	57.2	50.8	406.6
Debt burden index	14.0	0.7	19.7	20.4	22.8	9.7	22.6

Table A7. Clusters variables values for Sector 43 (Specialized construction works)

For sector 43, cluster numbers 1, 2, 3, 4 and 5 are associated with a high number of cases.

The first cluster (43_2) is generally characterized by high total debt values, mainly of banking origin, which negatively affects financial autonomy. Inadequate capitalization levels negatively affect asset coverage and the ability to service bank debt. Therefore, firms in this cluster can be classified as "undercapitalized".

The second cluster (43_1) shows adequate levels of capitalization that favor a contained degree of financial leverage and correct coverage of assets despite a higher incidence of tangible assets. Compared to other clusters in the economic sector in question, the economic indicators show modest values. The firms in this cluster can be classified as "slow movers".

The third (43_4) and fourth clusters (43_5) are mainly characterized by an unbalanced capital structure which highlights the weight of financial liabilities and even more so of operational ones. Asset management appears to be sufficiently elastic, but the low levels of capitalization do not allow it to be properly hedged. Positive economic performances measured by the ROI and added value allow firms to mitigate the incidence of financial charges on the EBITDA, with particular reference to cluster 43_5. The firms in these clusters are characterized as being "supply-dependent with low capitalization".

The fifth cluster (43_3) is mainly characterized by high values for inventories on total assets, inventory duration, added value on production value and debt burden index, and by low values for ROI, tangible assets on total assets and payables to suppliers on shareholders' equity. The firms in this cluster can be characterized as being "negatively impacted by inventory".