Risk Scoring Models for Trade Credit in Small and Medium Enterprises

Manuel Terradez, Renatas Kizys, Angel A. Juan, Ana M. Debon, Bartosz Sawik

Abstract Trade credit refers to providing goods and services on a deferred payment basis. Commercial credit management is a matter of great importance for most small and medium enterprises (SMEs), since it represents a significant portion of their assets. Commercial lending involves assuming some credit risk due to exposure to default. Thus, the management of trade credit and payment delays is strongly related to the liquidation and bankruptcy of enterprises. In this paper we study the relationship between trade credit management and the level of risk in SMEs. Despite its relevance for most SMEs, this problem has not been sufficiently analyzed in the existing literature. After reviewing some previous works, we use a large database of enterprises to analyze data and propose a multivariate decision-tree model which aims at explaining the level of risk as a function of several variables, both of financial and non-financial nature. Decision trees replace the equation in parametric regression models with a set of rules. This feature is an important aid for the decision process of risk experts, as it allows them to reduce time and then the economic cost of their decisions.

Manuel Terradez
Universitat Politcnica de Valncia, Valencia, Spain, e-mail: mterrade@eio.upv.es

Renatas Kizys
Portsmouth University, Portsmouth, UK, e-mail: renatas.kizys@port.ac.uk

Angel A. Juan
IN3 Open University of Catalonia, Barcelona, Spain, e-mail: ajuanp@uoc.edu

Ana M. Debon
Universitat Politcnica de Valncia, Valencia, Spain, e-mail: andeau@eio.upv.es

Bartosz Sawik
AGH University of Science and Technology, Krakow, Poland e-mail: bsawik@zarz.agh.edu.pl

1 Introduction

The current financial crisis has renewed the interest in research and development of failure prediction models for all of the corporate and retail sectors ([2]). The literature on the modeling of credit risk for large, listed companies is extensive, and it either uses historical accounting data to predict insolvency or models that rely on market information. However, market information is not available for small and medium enterprises (SMEs), which require risk management tools and methodologies specifically developed for them. Research on credit risk management for SMEs is relatively scarce. This research aims to partially fill this void by analyzing the risk of trade credit operations in SMEs. Trade credit (TC) involves supplying goods and services on a deferred payment basis; that is, giving the customer time to pay. Because TC represents an important share of total assets or liabilities, managing it is critical for the businesses, especially for SMEs. Thus, TC is an implicit short-term loan from non-financial suppliers to their clients. It occupies a prominent place in the world of business and is one of the most important forms of credit available to businesses.

While the actual cost of institutional credit remains close to the nominal cost, the cost of TC can vary widely. In effect, if significant discounts for early payment are considered, TC can become an expensive way of borrowing. TC is reflected in both the level of credit (the amount purchased on credit) and the length of the credit period (the number of days taken before payment is made). We discuss here some basic concepts and models to predict default risk in SMEs based on TC indicators. We explore the use of several models, including both classical statistical and econometric models (e.g., logistic regression and multiple discriminant analysis) as well as data-mining techniques (e.g., decision trees, neural networks, nearest neighbor). As discussed in the paper, our results suggest that decision trees have the best fit, where CHAID (Chi-squared Automatic Interaction Detector) provides better prediction of defaults than CART (Classification and Regression Tree). We also find that the most important predictor is the ratio of accounts payable over total liabilities, with larger values of this ratio implying a greater risk of default. Other important predictors are the ratio of accounts payable over accounts receivable and sales growth.

The remainder of this paper is organized as follows. Section 2 contains a literature review on the topic. Section 3 addresses measurement and estimation issues of default risk. Section 4 describes the data employed in this study. Section 5 provides an overview of our methodology. Finally, Section 6 offers some concluding remarks.

2 Literature Review

Early research into corporate failure prediction involved determining which accounting ratios best predict failure, primarily employing Multiple Discriminant Anal-

ysis (MDA) or Logit/Probit models. Usually, ratios are calculated a year before bankruptcy or default and thus these are static models. Altman ([1]) and Ohlson ([20]) pioneered models to predict failure using financial ratios. Altman used MDA, which was echoed by Deakin ([11]) and Micha ([19]), inter alia. Ohlson introduced a logistic regression, which has several advantages over MDA (see next section for a discussion on this matter) and a wealth of studies followed this line ([4], [5]).

Credit risk models for private companies are limited by data availability. Market data is not available for unlisted firms. Moreover, some of the data required to calculate accounting ratios in studies of the failure of listed companies is not available for SMEs. Other studies, using a variety of statistical techniques, have contributed to the knowledge of the insolvency indicators that anticipate in SMEs, both financial ([9], [3]) and non-financial ([14], [2]). In particular, ([12]) propose a non-parametric survival approach with a random-forest model, but they also conclude that a simpler logit model outperforms the random-forest model in the out-of-sample validation.

As we discussed in the previous section, managing trade credit is critical for the business, especially for SMEs, as it represents an important share of total assets. Therefore, it is not surprising that recent research has focused on the links between the management of TC (and delays in payment) and the liquidation or bankruptcy of enterprises, or even the refinancing (or restructuring) of debt ([29], [10]). Commercial lending involves credit risk due to exposure to default that can have negative effects on probability and liquidity ([8]). According to ([21]), the proper management of the TC offered (as a supplier) is critical to the survival and success of business. These authors also conclude that most SMEs are not proactive in their management of credit, and that they do not employ risk models (according to them, about 83% of SMEs do not classify their customers using risk categories).

As noticed by ([7]), credit-constrained firms facing liquidity problems from their customers are more likely to not paying their suppliers. However, because of the difficulty of obtaining data, the line of research that studies the relation between the management of TC and risk is not sufficiently developed.

3 Measuring Default Risk in SMEs

According to European standards, SMEs have less than 250 employees and sales figures under 50 million EUR (or total assets under 43 million EUR). As pointed out by ([2]), two of the main factors behind failed SMEs are insufficient capitalization and lack of planning. In the related literature it is frequent to find terms related to high levels of risk, such as: insolvency, bankruptcy, failure, default, etc. All of these terms are quite similar, albeit with small differences. In fact, they can be used interchangeably in a modeling framework, since they are usually transformed into a binary variable that takes on value 1 if the event occurs and takes on value 0 otherwise. In this paper we use the term default. Accordingly, we use the probability of default (PD) as a measure of risk. Notice, however, that failure and closure are different concepts: while failure generally implies closure, the inverse is not true

a firms closure may be due to other reasons. Several factors can affect a PD, such as the firms leverage, profitability or cash flows. A scoring model specifies how to combine the different pieces of information in order to get an accurate assessment of the PD.

Assume we have collected firm-level data at annual frequency on default factors and default behavior. The binary variable of default will take on value 1 if the firm eventually defaults in the year following the one observed for the factors, and zero otherwise. Ideally, the scoring model should predict a high PD for those firms that eventually will default and a low PD for those that will not. Logistic regression models can be used to predict default because the response variable is binary and they yield a score between 0 and 1, which can be interpreted as the clients PD. The model coefficients signal the importance of each predictor in the explanation of the estimated probability of default. A score summarizes the information contained in factors that affect the PD, e.g.: $score = b_0 + b_1x_1 + b_2x_2 + + b_kx_k$. The logistic function, $P = 1/(1 + exp(score))$, is usually applied to link scores to PDs. The ratio $P/(1P)$ is called odds-ratio and facilitates the model interpretation. $Log[P/(1P)]$ is called $lo - git(P)$ and thus the associated models are called logit models. A natural way of estimating the coefficients of the model is throughout the maximum likelihood method, i.e., the coefficients are chosen such that the probability of observing the given default behavior is maximized.

Alternative nonlinear techniques that can be used to approach this problem include decision trees and neural networks, among others ([6], [16]). A decision tree is a set of conditions organized in a hierarchical structure, so that the final decision can be determined by the fulfillment of the rules from the root of the tree to one of its end nodes. One of the great advantages of this technique is that the possible options from a given condition are exclusive, which allows analyzing a situation and, following the decision tree properly, get an action or take a decision. Two of the main techniques for developing trees are CART (Classification and Regression Tree) and CHAID (Chi-squared Automatic Interaction Detector): CART performs binary partitions and assigns a mean and variance to each node, trying to select partitions that reduce the variance of the child nodes; CHAID performs non-binary partitions and uses a Chi Square test to determine the optimal partition.

Instead of the well-known R^2 statistic, suitable for linear models, in the case of nonlinear models we can report the $Pseudo - R^2$, which is also bounded by 0 and 1 with higher values indicating a better fit. However, in this work we use an alternative measure of fit which is frequently used in binary models: the ROC curve. A ROC curve is a technique for visualizing, organizing and selecting classifiers based on their performance. It has its origin in Signal Detection Theory ([28]) and has been widely accepted and commonly used in fields such as Psychology ([18]) and Medicine ([30]). It has also been introduced in other fields that are more related with our work, such as Economics ([26]) and Data-mining ([15]). ROC curves are particularly useful for comparing the classification power of different estimated models. Details of ROC curves are provided in ([13]). The degree of predictability of the model is defined by the area under the ROC curve (AUC), which is constructed for all possible cutoff points to classify positive or negative events. Since the AUC is

a portion of the area of the unit square, its value will always be between 0 and 1, where the random guessing procedure has an area of 0.5. As with the $Pseudo - R^2$, the greater is the AUC the better is the classifier.

Alternative measures of risk that can be used to approach the problem of trade credit risk are Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR). VaR and CVaR have been widely used in the field of financial engineering (e.g. [25], [27]). CVaR is used in conjunction with VaR and is applied for estimating the risk with non-symmetric cost/return distributions. Rockafellar and Uryasev ([23], [24]) introduced a new approach to select a set of investments with the reduced risk of high losses. The problem was solved by calculating VaR and minimizing CVaR simultaneously. For trade credit VaR can be defined as accepted threshold of risk by decision-maker, in that case CVaR would be worst case accepted level of TC risk. Optimization models with VAR and CVaR could be used to shape the distribution of TC risk in a favorable way for a decision maker.

4 Data Sources

The Iberian Balance sheet Analysis System (SABI) is a database including balance sheets information of more than 1.2 million Spanish and more than 0.4 million Portuguese companies. A random sample (extracted from SABI) with more than 5,000 active Spanish SMEs has been used. We selected firms with these conditions: active Spanish SMEs (see definition in the previous section) with information of the last accounting year (that is, 2011) and variables Accounts receivable or Accounts payable informed. The SABI database provides MORE, a proprietary scoring model. One of our goals is to develop a scoring model, alternative to MORE, based on predictors related to TC. Our scoring model should have a transparent and replicable methodology under the assumption that SMEs typically apply homogeneous risk rules to all their TC customers. In contrast to MORE, in this research we also propose a model that allows for customized TC rules that suit different customers and, thus, leads to reduced levels of default risk. In our model we use a binary dependent variable (response) which, based on the MORE score, classifies SMEs into two categories: risky and non-risky companies. The former comprise companies with standard CC or lower rating featuring a relatively high probability of default the MORE rating consists of 10 categories indicated by traditional symbols used by rating agencies: AAA to D, being CC the 8th one ordered from most creditworthy to least creditworthy. The non-risky companies comprise all the remaining SMEs in the sample with a relatively low probability of default.

As for the independent variables (predictors) most of them are variables related to TC. Specifically, we used: (a) DAR: Days accounts receivable (debtors); (b) DAP: Days accounts payable (creditors); (c) AR_Assets: Ratio (Accounts receivable/Total assets); (d) AP_Liab: Ratio (Accounts payable/Total liabilities); (e) AP_AR: Ratio (Accounts payable/Accounts receivable); (f) APGrowth: Ratio (Accounts payable [last year]/Accounts payable [previous year]); and (g) ARGrowth: Ratio (Accounts

receivable [last year]/Accounts receivable [previous year]). We also used some other factors that, although not directly related to default on TC, help us to account for the existing heterogeneity of SMEs. These factors are: number of employees, age (years in operation), activity sector, and sales growth in the last year. All the variables were obtained for the last accounting year, that is, 2011. Growth variables were derived comparing 2011 and 2010. Figure 1 shows some descriptive statistics of the variables.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Employees	5094	1	244	9.20	18.673
Age	5093	3.0056	87.7889	15.038873	8.9831294
Log_AR_Assets	5094	-15.26	.00	-2.2883	1.90025
Log_AP_Liab	5094	-12.56	4.05	-1.7992	1.52436
Log_AP_AR	5048	-10.04	13.56	.5515	1.88502
Log_SalesGrowth	5094	-9.34	6.24	-.0826	.62420
Log_ARGrowth	5094	-9.71	10.74	-.0681	1.14850
Log_APGrowth	5094	-8.34	7.05	-.0333	.81845
Log_DAP	5094	-2.66	13.20	4.8826	1.45649
Log_DAR	5094	-2.30	14.05	3.5273	2.01879
Valid N (listwise)	5047				

Fig. 1 Descriptive statistics of the considered variables

Fig. 1 .
Also, the associated correlations are given in Table 1.

5 Methodology and Results

All enterprises with less than three years of operation were excluded from the study, since their accounting ratios and business behavior are not consolidated enough and their inclusion could mislead the results of the analysis. In the first step, we performed a log transformation of our continuous independent variables featuring high concentration on low values but long positive tails. We also used dummy predictors to include one categorical variable, activity sector (agriculture, manufacturing, building, services). In a second step we estimated various models, including both classical statistical and econometric models (logistic regression and multiple discriminant analysis) as well as data-mining techniques (CART and CHAID decision trees, neural networks, and nearest neighbor). Our results suggest that decision trees

Table 1 Correlations between paris of variables

Pearson Correlations

	Employees	Age	Log_AR_Assets	Log_AP_Liab	Log_AP_AR	Log_SalesGrowth	Log_ARGrowth	Log_APGrowth	Log_DAP	Log_DAR
Employees	1	.242^a	.113^a	.059^a	-.077^a	.048^a	.020	.005	.001	.096^a
Age	.242^a	1	-.001	-.094^a	-.086^a	-.017	-.014	-.047^a	-.014	.086^a
Log_AR_Assets	.113^a	-.001	1	.338^a	-.676^a	.041^a	.299^a	.037^a	-.003	.512^a
Log_AP_Liab	.059^a	-.094^a	.338^a	1	.411^a	.054^a	.028^b	.258^a	.309^a	.027
Log_AP_AR	-.077^a	-.086^a	-.676^a	.411^a	1	.003	-.272^a	.167^a	.280^a	-.580^a
Log_SalesGrowth	.048^a	-.017	.041^a	.054^a	.003	1	.155^a	.223^a	-.065^a	-.157^a
Log_ARGrowth	.020	-.014	.299^a	.028^b	-.272^a	.155^a	1	.236^a	.015	.217^a
Log_APGrowth	.005	-.047^a	.037^a	.258^a	.167^a	.223^a	.236^a	1	.243^a	.020
Log_DAP	.001	-.014	-.003	.309^a	.280^a	-.065^a	.015	.243^a	1	.156^a
Log_DAR	.096^a	.086^a	.512^a	.027	-.580^a	-.157^a	.217^a	.020	.156^a	1

^{a} Correlation is significant at the 0.01 level (2-tailed).
^{b} Correlation is significant at the 0.05 level (2-tailed).

show the best fit. An appealing feature of decision trees is that they are easy to implement and interpret. Neural networks provided equivalent results in our case, but they are much more difficult to interpret so we preferred to go with the decision trees. Results obtained with CHAID and CART decision trees are very similar, but CHAID seems to provide slightly better predictions for the target category (defaults). Therefore, our preferred model is a decision tree based on the CHAID technique. Our model has a maximum depth of 5 levels, a minimum nodes size of 20 individuals the Pearson Chi-Square statistic is used to decide the joining and division of nodes, and 6 intervals for the continuous predictors. The tree has 51 terminal nodes, which can be considered too many and could lead to an over-fitting problem. In order to minimize this problem, we carry out a 10-fold Cross Validation to validate the results.

Figures 2 and 3 suggest that our CHAID model shows a quite acceptable performance: AUC = 0.808 with respect to MORE, and 84% of success in predicting the right category. In fact, it shows a reasonable success in predicting both categories (96.2% of non-defaults and 29.2% of defaults were successfully predicted). Furthermore, it shows a better goodness of fit than the logit model.

Fig. 2 .
Figure 2. ROC curves for the CHAID and logit models.
Fig. 3 .
Figure 3. Numerical performance of the CHAID and logit models.
Fig. 4 .
Figure 4 shows the first two branches (13 nodes) of the final tree due to space limitations we do not depict the whole tree here.
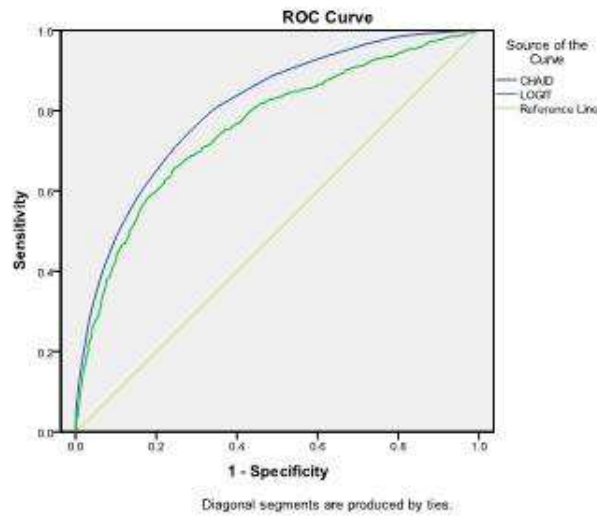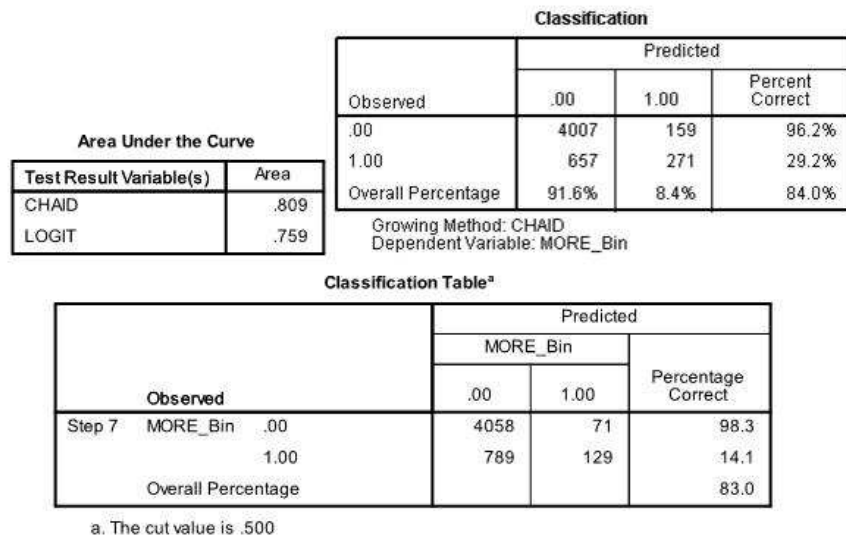
Fig. 2 ROC curves for the CHAID and logit models

Area Under the Curve

Test Result Variable(s)	Area
CHAID	.809
LOGIT	.759

Classification

	Predicted		
Observed	.00	1.00	Percent Correct
.00	4007	159	96.2%
1.00	657	271	29.2%
Overall Percentage	91.6%	8.4%	84.0%

Growing Method: CHAID
Dependent Variable: MORE_Bin

Classification Table[a]

			Predicted		
			MORE_Bin		
Observed			.00	1.00	Percentage Correct
Step 7	MORE_Bin	.00	4058	71	98.3
		1.00	789	129	14.1
	Overall Percentage				83.0

a. The cut value is .500

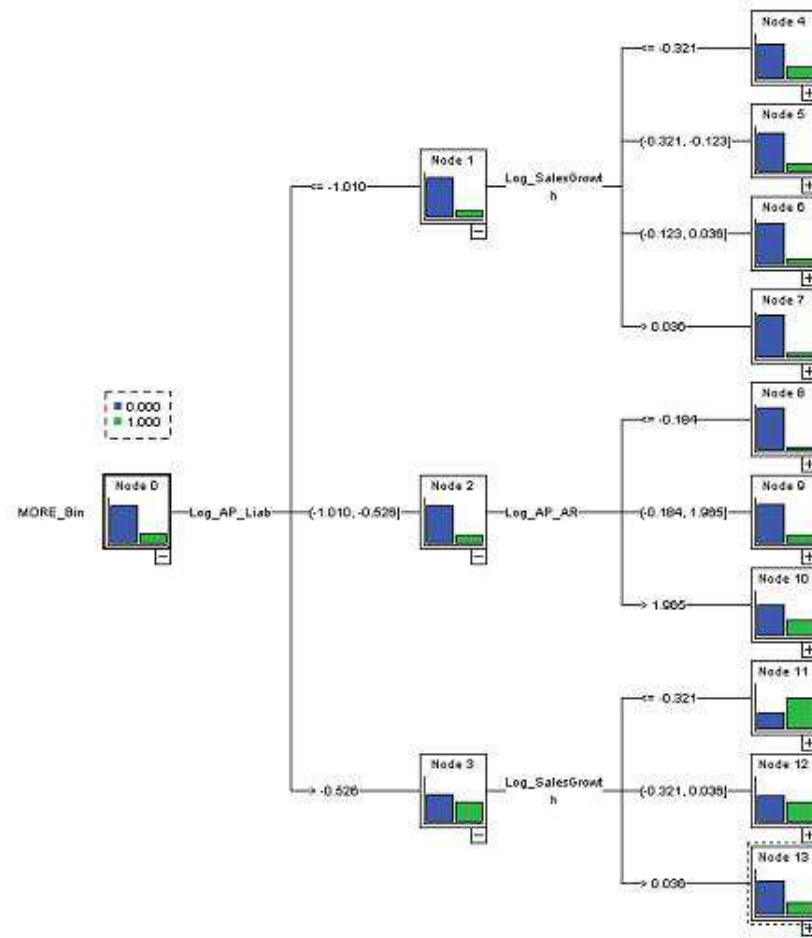**Fig. 3** Numerical performance of the CHAID and logit models

Fig. 4 First two branches of our CHAID tree model

The most important predictor is the accounts payable/total liabilities ratio. Larger values of this ratio imply a greater risk of default, with thresholds located at 0.36 and 0.59 (after transforming the model back to the levels). Other important variables are the ratio of accounts payable over accounts receivable and sales growth (SG). Our analysis of nodes 1 and 3 and their child nodes indicates that when AP_Liab is high or low, the second important variable to look at is SG. As expected, the lower is the value of SG the riskier is the company. However, when AP_Liab is medium (node 2), the second important variable is AP_AR, with larger values relating to riskier

firms, too. All of the initial variables are included in the final model, but those with less influence are accounts payable growth and activity sector.

In order to test the performance of our model, we used a reference model: the classical logit model proposed by ([22]). This logit model offers results which are fairly close to those obtained with MORE (AUC = 0.92 with respect to MORE). The logit model takes into account the main financial dimensions, such as: solvency, liquidity, profitability, leverage, etc. This logit model provides a good benchmark because it uses the same database (SABI) and the same population (Spanish SMEs) as in our study. An advantage of this model over the Altman model is that financial rations used by Altman are not very common in Spanish balance sheets. Accordingly, some studies ([17]) emphasize the limited applicability of the Altman model on Spanish firms.

6 Conclusions and Future work

Our research, based on a credit scoring model, provides evidence that trade credit is a good proxy of risk for Spanish SMEs. This means that firms can reduce their risk by managing TC properly, which implies adjusting the ratio of accounts payable over total liabilities as a first step. Our research also makes a significant contribution to the relatively scarce literature on the application of decision trees to credit risk analysis. The decision tree is distinguished by several aspects that provide better practical results than the parametric models. The results obtained using tree methods for classification or regression can be summarized in a series of (usually few) logical if-then conditions (tree nodes). This makes it easy to understand and interpret the model.

This paper aims at being a first step into a project which main goal is to help companies to reduce their global risk by customizing TC rules to different customers. In future work, we plan to combine the scoring model developed here with metaheuristic algorithms in order to support TC risk decision-making in SMEs. We also plan to develop new optimization models with VaR and CVaR as a risk measures for TC. These kind of models will provide a decision maker with a tool for evaluating the relationship between expected and worst-case TC risk level.

References

1. Altman, E.I.: Financial ratios, discriminant analysis and prediction of corporate bank-ruptcy. The Journal of Finance, **23**, 589–609 (1968).
2. Altman, E.I., Sabato, G., Wilson, N.: The value of non-financial information in SME risk management. The Journal of Credit Risk, **6**, 1–33 (2010).
3. Altman, E.I., Sabato, G.: Modeling credit risk for SMEs: evidence from the US market. ABACUS, **43**, 332–357 (2007).
4. Aziz, A., Emanuel, D.C., Lawson, G.H.: Bankruptcy prediction: an investigation of cash flow based models. Journal of Management Studies, **25**, 419–437 (1988).

5. Becchetti, L., Sierra, J.: Bankruptcy risk and productive efficiency in manufacturing firms, Journal of Banking and Finance, **27**, 2099–2120 (2002).
6. Berry, M.J.A., Linoff, G. Data Mining Techniques. John Wiley & Sons (1997).
7. Boissay, F., Gropp, R.: Payment defaults and interfirm liquidity provision. Review of Finance, 1–42, doi:10.1093/rof/rfs045 (2013).
8. Cheng, N., Pike, R.: The trade credit decision: evidence of UK firms. Managerial and Decision Economics, **24**, 419–438 (2003).
9. Correa, A., Acosta, M., Gonzlez, A.L.: La insolvencia empresarial: un anlisis emprico para la pyme. Revista de Contabilidad, **6**, 47–79 (2003).
10. Cuat, V.: Trade credit: suppliers as debt collectors and insurance providers. Review of Financial Studies, **20**, 491–527 (2007).
11. Deakin, E.B.: A discriminant analysis of predictors of business failure. Journal of Accounting Research, **10**, 167–179 (1972).
12. Fantazzini, D., Figini, S.: Random Survival Forest Models for SME Credit Risk Measurement, Methodology and Computing in Applied Probability, **11**, 29–45 (2009).
13. Fawcett, T.: An introduction to ROC analysis. Pattern Recognition Letters, **27**, 861–874 (2006).
14. Grunert, J., Norden, L., Weber, M.: The Role of Non-Financial Factors in Internal Credit Ratings, Journal of Banking and Finance, **29**, 509–531 (2004).
15. Hastie, T, Tibshirani, R., Friedman, J.: The elements of statistical learning: data mining, inference and prediction. Springer (2001).
16. Hernndez, J., Ramrez, M.J., Ferri, C.: Introduccin a la minera de datos, Pearson Prentice Hall (2004).
17. Lizarraga, F.: Modelos de previsin del fracaso empresarial: funciona entre nuestras empresas el modelo de Altman de 1968?, Revista de Contabilidad, **1**, 137–164 (1998).
18. Metz, C.E., Kronman, H.B.: Statistical significance tests for binormal ROC curves. Journal of Mathematical Psychology, **22**, 218–243 (1980).
19. Micha, B.: Analysis of business failures in France. Journal of Banking and Finance, **8**, 281–291 (1984).
20. Ohlson, J.: Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research, **18**, 109–131 (1980).
21. Poutziouris, P., Michaelas, N., Soufani, K.: Financial management of Trade Credits in SMEs. Working paper. Concordia University.
http://www.efmaefm.org/efma2005/papers/241-soufani_paper.pdf.
22. Pozuelo, J., Labatut, G., Veres, E.: Anlisis descriptivo de los procesos de fracaso empresarial en microempresas mediante tcnicas multivariantes. Revista Europea de Direccin y Economa de la Empresa, **19**, 47–66 (2010).
23. Rockafellar R.T., Uryasev S.: Optimization of conditional value-at-risk. The Journal of Risk, **2**(3), 21–41 (2000)
24. Rockafellar R.T., Uryasev S.: Conditional value-at-risk for general loss distributions. The Journal of Banking and Finance, **26**, 1443–1471 (2002)
25. Sarykalin S., Serraino G., Uryasev S.: Value-at-Risk vs. Conditional Value-at-Risk in Risk Management and Optimization. In: Z-L. Chen, S. Raghavan, P. Gray (Eds.) Tutorials in Operations Research, INFORMS Annual Meeting, Washington D.C., USA, October 12–15 (2008)
26. Sobehart, J.R., Keenan, S.C.: A practical review and test of default prediction models. RMA Journal, **84**, 54–59 (2001).
27. Sawik, B.: Downside Risk Approach for Multi-Objective Portfolio Optimization. In: Klatte, D., Lthi, H.-J., & Schmedders, K. (Eds.) *Operations Research Proceedings 2011*, Operations Research Proceedings, Springer-Verlag, Berlin, Heidelberg, pp. 191–196 (2012).
28. Swets, J.A.: Signal detection theory and ROC analysis in psychology and diagnostics. Collected papers. Lawrence Erlbaum Associates (1996).
29. Wilner, B.: The exploitation of relationships in financial distress: the case of trade credit. The Journal of Finance, **55**, 153–178 (2000).
30. Zweig, M.H., Campbell, G.: Receiver-Operating Characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical Chemistry, **39**, 561–577 (1993).