A new tool for failure analysis in small firms: frontiers of financial ratios based on percentile differences (PDFR)

ABSTRACT

This paper proposes an innovative methodology to compute scores and distances to failure, specially oriented to SMEs. It is based on statistical differences between the group of failed firms and the population to which the failed firms belong (industry, period, and geographical zone selected). Our results for the selection of the most discriminant variables are consistent with previous literature; and the hit rates of failed and non-failed firms outperform those of the commonly used traditional methodologies. In addition, the proposed methodology allows us to compute distances to failure of both individual firms and groups of firms. Finally, this methodology identifies which ones of the financial drivers used are strengths or weaknesses for the specific firms or group of firms under study, for purposes of a potential reorganization.

KEYWORDS

Business failure, financial ratios, SMEs, percentile differences, mathematical model.

JEL: G17, G33, L25.

1. INTRODUCTION

The role played by SMEs in the economy is indisputable. The report ‘EU SMEs in 2012: at the crossroads. Annual report on small and medium-sized enterprises in the EU, 2011/12’ estimates that in 2012, across the 27 countries in the European Union, SMEs counts for 99.8% of total firms, 67.4% of employment and 58.1% of gross added value. We take into consideration that SMEs show particular characteristics and behavior in respect to big firms, requiring specific tools for risk management. Furthermore, during recessions of the economic cycle a high rate of SMEs enter in critical situations or bankrupt. All these reasons justify our selection of SMEs as the main subject for our model to be applied

Two general drawbacks can be identified in business failure analysis methodologies commonly applied to date: 1) Distances to failure are not computed, as the classification tends to be dichotomous: failed or non-failed. 2) Financial information extracted by ratios does not capture changes in the environment quickly enough. Models do not take macroeconomic or industry changes into consideration, except for the effect caused over the financial information, even though the firm’s evolution may be partially caused by the evolution of the industry or that of the geographic zone instead of by the specific firm’s performance, hampering the comparison of results obtained in different periods or geographical zones.

To overcome such drawbacks, our first objective is the development of a business failure model oriented to SMEs that obtain scores to quantify the distance to failure of individual firms or groups of firms. By quantifying the role of every financial driver to take part of a distance to failure, financial weaknesses and strengths are detected and may be used as a management tool for a potential reorganization of firms identified as next to failure.

If a firm’s group of reference is well chosen, the methodology takes into consideration environmental factors that influence the whole group (industry and macroeconomic factors). The desirable aim would be studying how the benchmarks vary as both types of environmental factors do, but in the short run our approach can provide with realistic benchmarks for healthy firms and failed firms in that period, as well as for the distance a certain firm is from those two benchmarks.

We contribute an innovative methodology, Percentile Difference Frontier of Ratios (PDFR), that allows:

- To select the most discriminant financial drivers, concerning bankrupt or business failure, as a function of the population under study: geographical zone, industry, and period.

- To determine frontiers of failure with the selected financial drivers during the previous step.

- To compute the distance to failure for specific firms or groups of firms, for each financial driver. Thus, by-driver distances help to detect deficiencies in those specific areas of the firm (or group) proxied by each financial driver, becoming a potentially useful management tool.

- To use drivers with non-homogeneous value ranges (as it is the case of financial ratios). This is the most interesting characteristic of the methodology since it produces rankings of firms according to their distance to failure for a given population and period.

We test the model on a specific sample of small Spanish firms belonging to the construction sector during the period 2006 to 2008. Our final sample is made up of 12,197 firms, of which 268 are failed.

The paper is structured as follows. Section 2 gathers some aspects of previous evidence on business failure of SMEs that are determinant for our research. Section 3 develops our methodological proposal. Section 4 describes the sample and displays our results on variable selection, graphical frontier of bankruptcy, scores and distances to failure, both for specific firms or groups of firms. Section 5 concludes.

2. LITERATURE REVIEW ON BUSINESS FAILURE IN SMES

The most frequently used methodologies to assess business failure using financial information are discriminant analysis (DA), such as linear DA and quadratic DA; and binomial models such as logit and probit (Aziz y Dar, 2006; Premachandra *et al*., 2009).

Table 1 gathers the pioneer works on the main methodologies proposed in this research line (Tascon and Castaño, 2012).

Table 1. Business Failure Methodologies

Dates	Methodologies
20th century, 30s	Univarite basic models
1966	Beaver. Univariate analysis. Analysis of variances and test of dichotomous classification
1968	Altman. Multiple discriminant analysis. Z-score model
1977	Martin. Logistic regression. Logit and Probit
1984	Marais et al. Iterative and recursive partitioning algorithms
1990	Bell et al. Artificial.Intelligence. Neuronal networks
1991	Mar Molinero y Ezzamel. Multidimensional escalation techniques
1996	Serrano-Cinca. Artificial intelligence. Self-organizing maps
2002	Park y Han. Multi-criteria analysis
2002	Shin y Lee. Artificial Intelligence. Genetic algorithms
2004	Paradi et al. Data envelopment analysis (DEA)

Out of the methods in Table 1, only DEA allows us to compute distances to failure, even though it is difficult to be applied to big samples like the one studied in this work. As for the environment, it plays a very poor or none role in all the methods mentioned. In general, financial drivers coming from sources other than financial information are difficult to obtain, given the scarce external information SMEs report.

Pioneer works such as Edmister (1972) and Lincoln (1984) show how useful are financial ratios to identify and predict failure in SMEs. However, the number of empirical studies on big firms is remarkably higher than that on SMEs. The reasons adduced for that lack of attention are difficulties in accessing, less quality, and lower reliability of information (Edmister, 1972; Labatut et al., 2009), although recent literature generally recognize that SMEs require specific tools for risk management considering their particular characteristics (Dietsch and Petey, 2002; Altman and Sabato, 2007; Behr and Guttler, 2007; Altman et al., 2008; Davydenko and Franks, 2008).

Due to their simpler structures, SMEs can answer faster to changes in economic conditions and satisfy the needs of local consumers, what makes the economy more dynamic (Altman and Sabato, 2005). Consequently, more volatile performance in SMEs than in big firms may induce extraordinary growth rates during boom periods, but fast failure in adverse conditions, giving rise to periodic distress for a significant number of SMEs (Dannreuther and Kessler, 2010).

Among the empirical evidence on business failure, Dietsch and Petey (2004) identify this higher risk in a sample of German and French SMEs. Pompe and Bilderbeek (2005) confirm that bankrupt is more difficult to predict in younger firms, after analyzing a sample of more than a thousand failed Belgian firms. These results confirm the need to develop specific tools to analyze business failure in SMEs.

Two elements treated in literature are determinant for our design of the experiment and the later discussion: (1) whether heterogeneous populations including several industries or homogeneous groups made up of firms in the same industry; and (2) whether several causes of failure or an only category of failed firms should be analyzed.

The evidence is unanimous (Lincoln, 1984; López et al., 1998; Psillaki et al., 2010) as for by-sector studies to improve the identification or prediction ability of failure models. Or study focuses on a certain industry and a geographical area, what provides us with homogeneous macroeconomic conditions. Additionally, as firms have been selected by size, similar conditions are set concerning economic risk.

The lack of an ultimate theory that unambiguously defines the concept of failure makes of officially declarations of bankruptcy the most commonly used criterion. Thus, distressed or failed firms are classified as non-failed, according to less strict criteria, despite they are likely to fail (Baixauli and Módica-Milo, 2010). The broader the definition of failure, the higher the rate of firms that is included in this category (Watson y Everett, 1996; Altman et al., 2008). However, officially dissolved firms may have not failed. This is the case of firms sold at fair prices, those absorbed due to strategic reasons (Headd, 2003) or those who change owners due to illness or retirement (Cochran, 1986).

Some authors distinguish two categories of failure: economically non-viable firms, and those other suffering transient financial distress (Franks y Torous, 1992). The distinction is relevant because next-to-failure or bankrupt firms should be reorganized when they are able to generate at least a normal performance (Cook et al., 2012). Laitinen (2008) includes both financial data and non-financial information in models able to identify the SMEs that could avoid the failure after a reorganization process. Cook et al. (2012) find strength in resources useful to identify the economically viable firms among the failed SMEs.

Our work takes into consideration these previous results in several ways. On the one hand, we are aware that our concept of failure (bankrupt, dissolved and extinct) includes as non-failed firms some ones that may suffer difficult o even critical financial situations; while some of the dissolved or extinct firms may have been in good financial or even profitable situations. On the other hand, that is the reason for our proposed methodology to establish three benchmarks: the median value of the group of failed firms, the median value of the population under analysis, and mirror values (healthy firms).

As for the application of scores to SME business failure analysis, we can mention Baixauli and Módica-Milo (2010) as a reference. In order to reduce heterogeneity of the group of firms qualified as healthy these authors build an indicator of financial strength[1]. They identify a sample of strong firms opposite to failed firms, to obtain more accurate models of failure prediction. Therefore, the score proxies for the probability of business failure in healthy firms. It is made up of four profitability measures (return on assets, turnover, return on equity, and change of annual net value) and one of financial information quality, based on the auditor's opinion.

Unlike the score proposed[2] by Baixauli y Módica-Milo (2010), our work poses differences of percentiles in respect to the population values, and takes the values on the percentiles as continuous variables. Thus, our scores better discriminate between groups of individual firms. Another difference is that Baixauli and Módica-Milo (2010) take only a fixed benchmark (percentile 5) while our work uses three benchmarks: the percentile on the central values of failed firms, the central percentile of the population, and the mirror percentile opposite to the percentile of failed firms. This way, our model is flexible and includes the environment conditions in the period under study: the levels in which firms are failing determine the levels in which we can qualify non-failed firms as more healthy. For example, if failures take place in percentiles far from the median, the economic setting of the population (country, region, industry) is worse and any firm would obtain bad results except if it is in a very good situation or makes a big

[1] Baixauli and Módica-Milo's (2010) indicator is computed: $S = \sum_{j=1}^{4} I\left(FR_j \geq P_5\right) + d_5$, where *FR* is a financial ratio, P_5 is the fith percentile, d_5 is a dichotomous variable that takes value *1* if the auditors' opinion is positive and *0* otherwise, and *I(.)* is the indicator function, that takes value *1* if *FR* is higher or equal to *P5* and *0* otherwise.

[2] As seen, they take percentiles inside the group of not failed firms.

competitive effort. In this setting, it makes sense to choose healthy firms far from the median in the opposite direction, that is, those with the best values for each ratio.

3. METHODOLOGICAL PROPOSAL

The frontiers of ratios based on differences of percentiles (PDFR) show some methodological strengths. It is a non-parametric method, as the relative importance of individual variables is measured as differences of percentiles, unlike the commonly used regression models (DA, logit, probit). PDFR is distribution-free, in the sense that the frontiers are made up of distances from the population to the failed group. No distribution is assumed for the location of observations in respect to the failure frontier. Furthermore, PDFR does not need a priori probabilities on the relative assignment of observations to the failed or the non-failed group. The main drawback, as in the case of other non-parametric methods (Premachandra et al., 2009), is that PDFR does not have significance tests.

To describe the methodology, we start by taking two groups in a population G of firms. G_s includes a number n_s of non-failed firms, and G_q includes a number n_q of failed firms. The population G is made up of n firms ($i = 1, \ldots, n$), being $n = n_s + n_q$. Now we define a group f of financial drivers ($f = 1, \ldots h$), selected by their ability to discriminate failed firms against non-failed firms. Financial drivers can be ratios, as in our empirical analysis, but could also be other types of financial measures. The only condition is that percentiles are good descriptive statistics of the measures used. The first step in the empirical process is describing how the selected financial drivers are distributed in the population under study (geographical area, industry, and period) by computing and displaying the frequency distributions of each financial driver by year, both for the whole population G, and for the failed firms, G_q. For each financial driver, we identify the values of the group of failed firms inside the distribution of values for the whole population where the failed firms are included. That is, we identify the values of failed firms and the percentiles of the distribution in which they are included.

The value of any firm's financial driver is noted as x_{if}, and the value of any failed firm's financial driver as z_{if}. We define the column vectors which include the values for the fth financial driver: X_f, for the whole group of firms, Y_f, for non-failed firms, and Z_f, for failed firms, being $X_f = (Y_f \cap Z_f)$.

We compute the median value $\tilde{z}_f$ of vector Z_f and the percentile of the value $\tilde{z}_f$ over the group of values in the whole population for each financial driver, $\pi(q)_f = p(\tilde{z}_f, X_f) \in [0,1]$. Similarly, we specify the place of an individual firm i for the financial driver f inside the population, $\pi_{if} = p(x_{if}, X_f) \in [0,1]$.

The percentile distance, d_{if}, for a firm i and a financial driver f is defined as the difference between the percentile placed by the firm's driver inside the values taken by the whole population for this financial driver, and 0.5, that is the percentile of the population's median value for this driver. We have adopted median instead of mean values to represent the population, and the group of failed firms G_q, considering the high dispersion of financial drivers in small firms (Castaño and Tascón, 2012).

$$\begin{cases} Si\ \pi(q)_f < 0{,}5\ ; & d_{if} = 0{,}5 - \pi_{if} \\ Si\ \pi(q)_f > 0{,}5\ ; & d_{if} = \pi_{if} - 0{,}5 \end{cases}$$

d_{if} equals 0 when a firm's value for the financial driver f equals the median value of the driver for the whole population. The distance is positive (negative) when the percentile for the failed group is lower than 0.5 and it means a higher (lower) failure risk; or when the percentile is higher than 0.5 and it places the firm nearer (farther) the failure. For any financial driver, the distance falls in the interval [-0.5 , 0.5].

Similarly, we can compute the percentile distance for a certain group in the population. Our interest group is G_q: the failed firms. Thus, we can homogeneously measure the discriminant power between failed and non-failed firms, for any financial driver. Homogeneous measures are especially important in financial ratios, as they show very different ranges and distributions.

$$\begin{cases} Si\ \pi(q)_f < 0{,}5\ ; & d(q)_f = 0{,}5 - \pi(q)_f \\ Si\ \pi(q)_f > 0{,}5\ ; & d(q)_f = \pi(q)_f - 0{,}5 \end{cases}$$

The bigger the distance, $d(q)_f$, between the percentile of failed firms, G_q, and that of the population, G, the better the discriminant ability of the *f*th financial driver.

According to Labatut *et al*. (2009), a specific financial driver taking a value far from the benchmark of a comfortable business situation does not mean that the firm is close to failure. To confirm that proximity to failure, several financial drivers have to simultaneously deviate from the financial health benchmark. Our model builds a score by adding the percentile differences of all those financial drivers found discriminant for the group of failed firms in respect to the population. That is, the score is made up of the higher distances, $d(q)_f$, computed for those financial drivers with discriminant power to analyze the proximity to failure. The score takes values in the interval $[-0.5 \times h\, , 0.5 \times h]$, being h the number of financial drivers. Thus, for 10 financial drivers the score for the group of failed firms (q) would take a value between -5 and 5.

$$S(h, q) = \sum_{f=1}^{h} d(q)_f$$

We can select a lower number of drivers ($h_1 < h$) when a part of the initially chosen are not discriminant for the population and period under study. The score computed for failed firms would be:

$$S(h_1, q) = \sum_{f=1}^{h_1} d(q)_f$$

It is possible to compute scores for groups of firms in the population (by size, industry, etc.) and also for individual firms. Thus, the score computed with all the selected financial drivers (h) for an individual firm (i) would be:

$$S(h)_i = \sum_{f=1}^{h} d_{if}$$

Using a lower number of financial drivers, h_1, the score for an individual firm (i) would be:

$$S(h_1)_i = \sum_{f=1}^{h_1} d_{if}$$

In order to avoid the size effect induced by the number of financial drivers used, an average percentile distance can be computed in any of the previous cases, just dividing by the number of drivers. For the group of failed firms G_q, the mean score falls in the interval [0 , 0.5]. For any other group, or an individual firm, the mean score falls in the interval [-0.5 , 0.5] whatever the number of drivers. For example, the mean score computed with all the selected drivers (h) for the group of failed firms (q) would be $\bar{S}(h,q)$

$$\bar{S}(h,q) = \frac{\sum_{f=1}^{h} d(q)_f}{h};$$

It is easy to deduce that financial drivers with low discriminant power, that is, with distances close to 0, will contribute little to the failed firms score, $S(h,q)$, or to the scores of individual firms, $S(h)_i$. Also, when mean scores ($\bar{S}$) are used, less discriminant drivers will reduce the total value. In that case, a lower number of drivers would be more useful to compare groups and firms.

A weighted average score can be computed by using the distance between the group of failed firms and the population $d(q)_f$ as the weighting in order to give more importance to the most discriminant drivers and less importance to the less discriminant drivers.

Finally, the third benchmark is formed by the mirror values of the failed firms' percentiles, that are computed as the values on the percentiles opposed to the median values of the group of failed firms, $\tilde{z}_f$. These values allow us to take values better than the median of the population as reference for each financial driver, considering the heterogeneous nature of non-failed firms.

$$\pi(esp) = 1 - \pi(q)_f \, \epsilon[0\,,1]$$

Therefore, we obtain three benchmarks:

- The score computed as the median of the failed firms' value for each financial driver.
- The score computed as the median of the population's value for each financial driver. The population includes both failed and non-failed firms.
- The score computed as the value on the mirror percentile in respect to the median of the failed firms' value for each financial driver.

When the position of a subgroup of the population, or the position of a firm, is analyzed, the score obtained is compared to: the score of the failed firms, $S(h,q)$, the score of the total population, $S(h,g)$, and the score of the mirror values, $S(h,esp)$. The last one is a better benchmark of healthy firms than G_s. We can place the firm in several levels of proximity to failure or to financial health. The scores consist of distances, hence the bigger the score is in positive values, the higher the firm's failure risk.

1. Very high failure risk, $S(h)_i \geq S(h,q)$;
2. Moderate to high failure risk, $(h,q) > S(h)_i \geq S(h,g) = 0$;
3. Medium financial situation, $S(h,g) = 0 > S(h)_i \geq S(h,esp)$;
4. Good financial situation, $S(h)_i < S(h,esp)$.

For each type of score proposed, distances to failure can be computed. Thus, the distance to failure for a group of firms, $G_a \in G$, (including failed firms and/or non-failed firms), is defined as the difference between the score of the group G_a minus the score of the group of failed firms G_q. To obtain distance results coherent with economic reasoning, we change signs, so that a distance to failure is negative when the group under study is worse than the reference group (failed firms); and a distance to failure is positive when the group under study is better than the group of failed firms.

- The distance to failure, D, using all the selected financial drivers (h) for group G_a, is:

$D(h, a) = -\big(S(h, a) - S(h, q)\big) =$

$\sum_{f=1}^{h} d(q)_f - \sum_{f=1}^{h} d(a)_f = \sum_{f=1}^{h} (d(q)_f - d(a)_f)$

- The mean distance to failure, $\overline{\mathrm{D}}$, using all the selected financial drivers (h) for group G_a, is:

  $\bar{D}(h, a) = -\big(\bar{S}(h, a) - \bar{S}(h, q)\big) = \frac{\sum_{f=1}^{h} d(q)_f}{h} - \frac{\sum_{f=1}^{h} d(a)_f}{h} = \frac{\sum_{f=1}^{h}(d(q)_f - d(a)_f)}{h}$

- The weighted average distance to failure, $\overline{\mathrm{Dw}}$, using all the selected financial drivers (h) for group G_a, is:

  $\overline{Dw}(h, a) = -(\overline{Sw}(h, a) - \overline{Sw}(h, q)) =$

  $\frac{\sum_{f=1}^{h} d(q)_f \cdot d(q)_f}{S(h,q)} - \frac{\sum_{f=1}^{h} d(a)_f \cdot d(q)_f}{S(h,q)} = \frac{\sum_{f=1}^{h}(d(q)_f - d(a)_f) \cdot d(q)_f}{S(h,q)}$

Using a lower number of financial drivers, $h_1 < h$, the formulation would be the same but replacing h by h_1.

Concerning the distances to failure for individual firms included in the population, G,

- Distance to failure, D, for firm i, using all the financial drivers selected (h):

  $D(h)_i = -\big(S(h)_i - S(h, q)\big) = \sum_{f=1}^{h} d(q)_f - \sum_{f=1}^{h} d_{if} = \sum_{f=1}^{h}(d(q)_f - d_{if})$

- Mean distance to failure, $\bar{D}$, for firm i, using all the financial drivers selected (h):

  $\bar{D}(h)_i = -\big(\bar{S}(h)_i - \bar{S}(h, q)\big) = \frac{\sum_{f=1}^{h} d(q)_f}{h} - \frac{\sum_{f=1}^{h} d_{if}}{h} = \frac{\sum_{f=1}^{h}(d(q)_f - d_{if})}{h}$

- Weighted average distance, $\overline{Dw}$, for firm i, using all the financial drivers selected (h):

  $\overline{Dw}(h)_i = -\big(\overline{Sw}(h)_i - \overline{Sw}(h, q)\big) =$

  $\frac{\sum_{f=1}^{h} d(q)_f \cdot d(q)_f}{S(h,q)} - \frac{\sum_{f=1}^{h} d_{if} \cdot d(q)_f}{S(h,q)} = \frac{\sum_{f=1}^{h}(d(q)_f - d_{if}) \cdot d(q)_f}{S(h,q)}$

Again, the formulation for a lower number of financial drivers, $h_1 < h$, is the same, but replacing h by h_1.

Once distances to failure are formulated, we can describe the position of an individual firm (and the position of any subgroup of the population).

1. The firm is in very high failure risk when the distance is lower than that of the group of failure firms (0). The risk is higher when the distance is more negative, $D(h)_i \; < \; D(h,q) = 0$;

2. The firm is in moderate to high failure risk when the distance ranges from zero to the population's distance. The risk is higher when the distance to failure is lower, $D(h,q) = 0 \;\; \leq \; D(h)_i \; < \; D(h,g)$;

3. The firm is in medium failure risk when the distance to failure is higher than that of the population but lower than the mirror distance (double distance). The risk is lower when the distance to failure takes higher positive values, $D(h,g) \; \leq \; D(h)_i \; < \; D(h,esp)$;

4. The firm is in a good financial situation when the distance to failure is higher than the mirror distance. The failure risk is lower when the distance takes higher positive values, $D(h)_i > D(h,esp)$.

4. EMPIRICAL STUDY: AN APPLICATION OF PDFR TO THE CONSTRUCTION INDUSTRY

As for the construction industry, we can perceive its relevance in the Spanish Economy, as well as the effect that the crisis started in 2007 had on business failure in this industry (National Institute of Statistics webpage). The bubble burst affected primarily the construction industry, becoming an interesting group to be analyzed in the first place.

Table 2. Employment and GDP rates in construction firms. Spain. 2007 and 2011

	2007 (a)	2011 (b)	Difference 2011-2007	% Dif.
Employed (thousands):				
· Total CNAE-93	20.356,2	18.654,1	-1.702,1	-8.36%
· Construction	2.680,6	1.497,9	-1.182,7	-44,12%
· % Construction/Total	13,17%	8,03%	69,48%	
GDP market prices (millions):				
· Total CNAE-93	1.053.161	1.063.355	+10.194	+0,97%
· Construction	131.074	98.546	-32.528	-24.82%
· % Construction/Total	12,45%	9,27%		

(a) Data on employed people referred to 2007, first trimester
(b) Data on employed people referred to 2011, second trimester
Source: www.ine.es.

To test the PDFR, we perform an empirical study on a homogeneous population: Spanish small firms in the construction industry for the period 2006 to 2008. We select a certain geographical zone, the Autonomous Community of Castille and Leon, to work with all the failed and non-failed firms for an industry and period, thus avoiding the sample bias. Data are provided by *Iberinform*, covering both financial information and the failure status of firms (bankrupt, dissolved, and extinct).

The population consists of 4.263 firms in 2006, 5.037 firms in 2007, and 3.427 firms in 2008. Given that several ratios are computed, we have eliminated those firms with indeterminate values. Thus, the final population consists of 4.142 firms in 2006, 4.825 in 2007, and 3.230 in 2008.

As for the variables to analyze the firms' situation, we take the eight most frequently used financial ratios in literature plus the most frequent ratios to proxy for profit margin and asset turnover, as found by Tascon and Castaño (2012). Table 3 shows statistics of the variables used in 2008 both for the group of failed firms and the group of non-failed firms. Dispersion of the variables and lack of normality can be appreciated[3].

Table 3. Descriptive statistics

Failure	variable	n	mean	median	stand. dev.	bias
no	R1= TD/TA	3193	1.464248	0.7783218	35.13914	56.26852
no	R2= CA/CL	3193	4.992888	1.274884	65.81954	30.69193
no	R3= EBIT/TA	3193	-0.0345993	0.0444596	2.19314	-41.647
no	R4= NI/TA	3193	-0.0876411	0.0159537	3.849023	-51.79087
no	R5= CA/TA	3193	0.7268043	0.7864155	0.3425963	9.195928
no	R6= FE/TD	3193	0.0319486	0.0214282	0.0583653	13.70667
no	R7= RP/TA	3193	0.810277	0.1062388	32.54369	56.31467
no	R8= CF/TD	3193	0.1890471	0.0608538	2.319081	23.02114

[3] Statistics for 2006 and 2007 (not tabulated) offer similar values.

no	R9= NI/SL	3193	-0.5259805	0.0124977	26.23841	-48.29314
no	R10= SL/TA	3193	1.815489	1.381083	22.46602	18.59462
yes	R1= TD/TA	37	1.09844	0.9751387	1.010565	3.7524
yes	R2= CA/CL	37	3.23058	1.057278	11.17866	5.727102
yes	R3= EBIT/TA	37	-0.2554667	-0.0164002	0.7124744	-2.656331
yes	R4= NI/TA	37	-0.2745696	-0.0354618	0.6439843	-2.230388
yes	R5= CA/TA	37	0.774168	0.8938961	0.2734852	-1.460193
yes	R6= FE/TD	37	0.4296611	0.032181	2.422027	5.832577
yes	R7= RP/TA	37	0.1619858	0.0299024	0.6085422	3.317699
yes	R8= CF/TD	37	1.418611	-0.0286602	14.66338	4.637255
yes	R9= NI/SL	37	-0.2360054	-0.0441338	0.5619837	-3.127511
yes	R10= SL/TA	37	2.612674	1.475444	5.620027	5.245937

NOTAS: TD/TA is Total Debt/Total Assets; CA/CL is Current Assets/Current Liabilities; EBIT/TA is Earnings Before Interests and Taxes/Total Assets; NI/TA is Net Income/Total Assets; CA/TA is Current Assets/Total Assets; RP/TA is Non Distributed Benefits/ Total Assets; FE/TD is Financial Expenses/Total Debt; CF/TD is CashFlow/Total debt; NI/SL is Net Income/Sales; and SL/TA is Sales/Total Assets.

4.1. Selection of the most discriminant ratios and graphical frontier of failure

Charts 1, 2 and 3 show, from high to low discriminant power, the 10 ratios used to identify failed firms inside the population under study. In the lower part of the Charts the values of the percentiles corresponding to the median of the ratio in the group of failed firms and the mirror percentiles are shown. The values of the ratios corresponding to the failure frontier are included in red.

Chart 1. Failure frontier. Construction firms. 2006

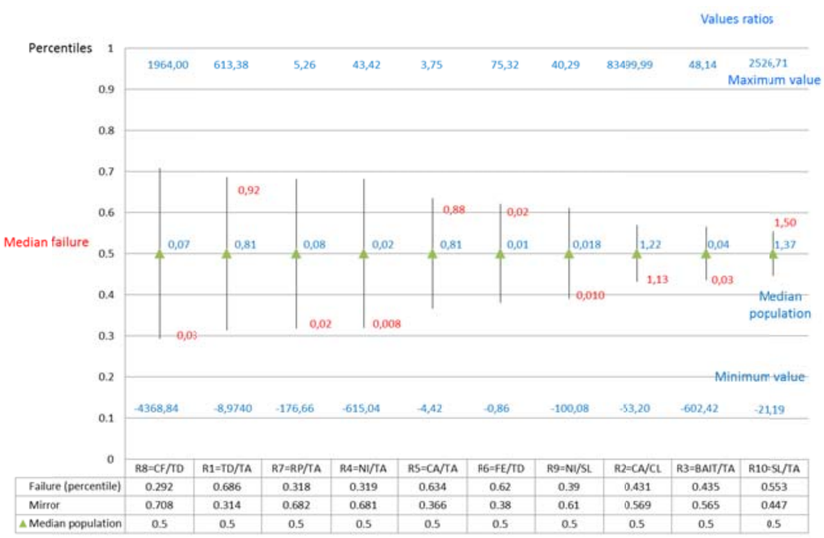
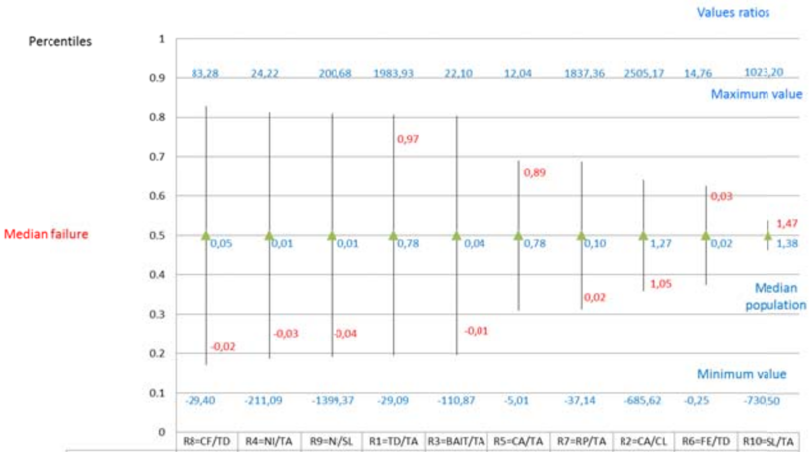


Chart 2. Failure frontier. Construction firms. 2007

	R1=TD/TA	R8=CF/TD	R9=NI/SL	R4=NI/TA	R5=CA/TA	R7=RP/TA	R6=FE/TD	R10=SL/TA	R2=CA/CL	R3=BAIT/TA
Failure (percentile)	0.662	0.338	0.339	0.343	0.627	0.378	0.61	0.591	0.431	0.435
Mirror	0.338	0.662	0.661	0.657	0.373	0.622	0.39	0.409	0.569	0.565
▲ Median population	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

Chart 3. Failure frontier. Construction firms. 2008



	R8=CF/TD	R4=NI/TA	R9=N/SL	R1=TD/TA	R3=BAIT/TA	R5=CA/TA	R7=RP/TA	R2=CA/CL	R6=FE/TD	R10=SL/TA
Failure (percentile)	0.172	0.187	0.191	0.806	0.196	0.69	0.313	0.359	0.625	0.537
Mirror	0.828	0.813	0.809	0.194	0.804	0.31	0.687	0.641	0.375	0.463
▲ Median population	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

Our results show that using percentile differences the most discriminant ratios are: in 2006, R8, R1, R7, R4, followed by R5, R6, R9; in 2007, R1, R8, R9, R4, followed by R5, R7, R6; and in 2008, R8, R4, R9, R1, R3, followed by R5, R7. The resulting selected variables are the same as using the Rank Sum Test[4] methodology. The only difference is that the Rank Sum Test compares failed firms with non-failed firms, while the proposed methodology compares failed firms with the population where they are included.

[4] This methodology, also called Mann-Whitney-Wilcoxon Test has been recently introduced in business failure analysis (Premachandra et al., 2009; Sueyoshi y Goto, 2009). It selects the most discriminant variables by computing median differences between the variables of the compared groups.

Second, the evolution of the frontiers of ratios over time is appreciated. Vertical lines indicate the distance from the median of the population to the percentile of failure, and in the opposite dirction, from the median of the population to the mirror percentile. The median value of each ratio for the group of failed firms is shown in red. The minimum, median, and maximum values for each ratio are shown in blue. In the table below the charts, the values of the percentiles for failed firms, the values of mirror percentiles, and the value of the percentile for the population (0.5) are shown.

We obtain similar values in 2006 and 2007, but the percentile distances remarkably increase in 2008 (except for ratio 10) and the median values for the whole population's ratios deteriorate. Some ratios deteriorate as they reduce when less value implies closer to failure; and the other ratios deteriorate as they increase when more value implies again closer to failure. Our results are consistent with the evolution of the industry in the country and period analyzed (Gill de Albornoz and Giner, 2010).

4.2. Robustness analysis

In order to check the robustness of our results for the construction industry in Spain, we have repeated the study in 2010 for a sample of 11.889 Spanish firms. The firms are required to be active during the previous two years, what reduced the number of failed firms to just 28. Thus, the new sample is taken after the adaptation of the country to international standards, with firms located in any Autonomous Community, and excluding small firms with only one or two years life. The most discriminant ratios obtained are: r9, r3, r4, r1 and r8, exactly the same as those obtained for 2008, thus confirming the importance of factors related to cash flow, return on assets and indebtedness in this industry.

4.3. Distance to a failure frontier computed with scores of financial ratios

The *score*, S, or distance to a failure frontier, is computed by adding the percentile differences of all the considered ratios (h), or those ratios whose percentile for the median value in failed firms is significantly different from 0,5. An equal number of ratios will make the scores comparable.

A second score, $\bar{S}$, easier to interpret, is the mean distance to failure. This value is the distance between the percentiles of the mean values of the selected ratios in both groups under comparison, the failed firms and the whole population.

Given that the discriminant power may be different for each ratio, a firm will be closer to or farther from failure if the most discriminant ratios indicate it. A big distance to failure in a low discriminant ratio should not be so important to qualify the firm as in a high discriminant ratio. Therefore, we propose a weighted failure distance, where the weighting measures how important the ratio is in the population and period analyzed[5].

We can observe that computing the frontier starting from 10 ratios, the distance is higher in absolute terms, but the mean distance is lower[6] considering the inclusion criterion for ratios. Due to it, a weighted average distance becomes useful.

Comparing the frontier values obtained with the scores for the group of failed firms in each year, we confirm that distances to failure are similar in 2006 and 2007 but substantially increase in 2008. Results are consistent whatever of the three scores proposed is used. Distances to failure allow us to be more precise in comparing firms or groups. Thus, we quantify the slight reduction of the distance to failure from 2006 to 2007, and the substantial increase of about 100% from 2007 to 2008.

Table 4. Scores for failed firms computed with 5 and 10 ratios in 2006, 2007, and 2008

Scores	**2006**	**2007**	**2008**
$S(5,q)$	0,8910	0,7690	1,5600
$S(10,q)$	1,3080	1,2260	2,2400
$\bar{S}(5,q)$	0,1782	0,1538	0,3120
$\bar{S}(10,q)$	0,1308	0,1226	0,2240
$\overline{Sw}(5,q)$	0,1814	0,1549	0,3122
$\overline{Sw}(10,q)$	0,1529	0,1333	0,2656

4.4. Relative position of specific firms

Given a certain population, once we have obtained the frontier to failure and the mirror frontier (above and below the median values of selected ratios for that population), it is

[5] Weightings are positive by definition, because distances between the group of failed firms and the whole population are defined in this way for all ratios. However, scores and distances computed for individual firms, or other groups of firms within the population, may be negative. In these cases, as the weighting is always positive, it does not eliminate the effect of negative signs.

[6] We allude here to the case in which the group of failed firms is compared to the whole population. However, when two groups of firms within the population are compared, all types of cases are possible.

possible to locate any firm (or group of firms) under analysis. For each ratio, distances to failure can be computed and graphically displayed. In addition, the firm's distance to each of both frontiers can be computed numerically through scores, in order to characterize the firm (or group of firms) inside the population. We have selected a failed firm and a non-failed firm in 2006 to show the characterization potential of the methodology.

In Chart 4, the orange line joints the percentile values of each ratio for the failed firm taken as an example, while the green line joints the percentile values of each ratio for the non-failed firm. Graphically, it can be appreciated the distance in percentiles to the median value for failed firms in that period and population. This is especially interesting for the failed firm taken as an example. In the opposite sense, it can be appreciated the distance to the median value of the whole population, and further to the mirror percentile. Both values become benchmarks of interest for healthy firms.

Chart 4. Business failure. 2006. Characterization of specific firms

	R8=CF/TD	R1=TD/TA	R7=RP/TA	R4=NI/TA	R5=CA/TA	R6=FE/TD	R9=NI/SL	R2=CA/CL	R3=BAIT/TA	R10=SL/TA
Failure (percentile)	0.292	0.686	0.318	0.319	0.634	0.62	0.39	0.431	0.435	0.553
Mirror	0.708	0.314	0.682	0.681	0.366	0.38	0.61	0.569	0.565	0.447
▲ Median population	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

Table 5 shows the three types of scores computed with 5 and 10 ratios, as well as the distances to failure, for both examples of firms displayed in Chart 4, a failed firm and a non-failed firm. At first sight, the chart shows that:

- The failed firm is in a clear failure position, as all the ratios locate in a worse position o close to (R7 y R6) the benchmark for the group of failed firms, G_q.

- For the healthy firm, the scores are negative, because almost all the ratios are in a better position than the median value for the whole population (except R5 and R10).

Table 5. Scores G_q and examples of failed and non-failed firms. 2006

***Scores* and Distances**	**2006**	**Ex. Failed**	**Distance to failure**	**Ex. Non-failed**	**Distance to failure**
$S(5,q), S(5)_i, D(5)_i$	0,8910	1,4250	-0,5340	-0,5650	1,4560
$S(10,q), S(10)_i, D(10)_i$	1,3080	3,0190	-1,7110	-0,9690	2.2770
$\bar{S}(5,q), \bar{S}(5)_i, \bar{D}(5)_i$	0,1782	0,2850	-0,1068	-0,1130	0,2912
$\bar{S}(10,q), \bar{S}(10)_i, \bar{D}(10)_i$	0,1308	0,3019	-0,1711	-0,0969	0,2277
$\overline{Sw}(5,q), \overline{Sw}(5)_i, \overline{Dw}(5)_i$	0,1814	0,2965	-0,1151	-0,1387	0,3201
$\overline{Sw}(10,q), \overline{Sw}(10)_i, \overline{Dw}(10)_i$	0,1529	0,2972	-0,1443	-0,1360	0,2889

We characterize the selected firms according to the levels proposed in the methodological section, both using scores (S, $\bar{S}$, y $\overline{Sw}$) and distances to failure (D, $\bar{D}$, y $\overline{Dw}$). All the scores computed for the failed firm are higher than the average values of the group of failed firms, meaning a very high failure risk.

The failed firm takes negative values for the distance to failure in every case, indicating worse percentiles than the medians of the group of failed firms. It is the equivalent position to very high failure risk.

As for the non-failed firm taken as an example, the scores are negative, and the distances to failure are positive. All the firm’s scores are lower than the score for the medians of the population (which is zero by definition), but higher than the scores for the mirror percentiles, $S(h,g) > S(h)_i \geq S(h,esp)$. The firm is in a good medium position, closer to the mirror values than to the population’s median values.

All distances to failure are higher than the population’s distance but lower than the mirror distance (which would be just twice the reference score in 2006): $D(h,g) \leq D(h)_i < D(h,esp)$. Consistently, all measures place the firm in a medium financial situation, next to the financial position of the best firms in the population under study.

4.5. Comparison to some traditional methodologies

In order to compare the effectiveness of PDFR in this paper to identify failed and non-failed firms, we have applied the most commonly used traditional methodologies over the same sample taken for 2008: linear discriminant analysis, quadratic discriminant analysis, logit, and probit. Table 6 shows the results of the horse race among the methodologies for the five variables found more significant by the PDFR and the Rank Sum Test. Unlike the rest of methodologies, PDFR obtains good results in both groups, failed and non-failed firms, with hit rates of 65% and 68%, respectively. Table 6 also shows percentages of Type I error (non-failed firms classified as failed) and Type II error (failed firms classified as non-failed).

Table 6. Comparison of PDFR with traditional methodologies

METHODS	NON-FAILED FIRMS				FAILED FIRMS								
	Classified as NON-FAILED		Classified as FAILED		Classified as NON-FAILED		Classified as FAILED		TOTAL WELL CLASSIFIED		TOTAL MISSCLASIFIED		Average % WELL
	Number	Percentage	Number	Percentage	Number	Percentage	Number	Percentage	Number	Percentage	Number	Percentage	CLASSIFIED
LDA	3102	97.15%	91	2.85%	31	83.78%	6	16.22%	3108	96.22%	122	3.78%	56.68%
QDA	98	3.07%	3095	96.93%	2	5.41%	35	94.59%	133	4.12%	3097	95.88%	48.83%
Logit	3193	100.00%	0	0.00%	37	100.00%	0	0.00%	3193	98.85%	37	1.15%	50.00%
Probit	3193	100.00%	0	0.00%	37	100.00%	0	0.00%	3193	98.85%	37	1.15%	50.00%
PDFR	2156	67.52%	1037	32.48%	13	35.14%	24	64.86%	2180	67.49%	1050	32.51%	66.19%

Type II Error

Type I Error

Due to the big overlap zone in which the values of the financial variables may belong to failed an non-failed firms, parametric models are strongly biased. As shown in Table 6, those methods performing well in identifying failed (non-failed) firms are very poor or incapable in identifying non-failed (failed) firms.

5. CONCLUSIONS

This work proposes a model to measure the distance to failure of small firms, the group of firms less studied in literature. The variables used are financial ratios, commonly used for this type of analysis, what allows us to compare the results of identification of failed firms obtained with the proposed methodology with those obtained using alternative methodologies. The ratios are computed with available information in mandatory financial statements. Using financial information mandatory for any size of firms is especially important in a model to be applied to small firms, given the limited information available in this type of firms.

The model computes a score by firm. Instead of distances between ratios, which are heterogeneous variables (exacerbated in SMEs), we use distances between percentiles, what makes the variables homogeneous. Hence, the variables can be aggregated to compute scores, independently of their ranges and statistical distributions. Besides, percentiles are computed starting from medians, which are more representative central values, considering the high dispersion of financial variables in SMEs.

This way of variable comparison is our main contribution as it allows to make classifications (or rankings) of the most relevant variables to discriminate between two groups; the score provides a number of points by firm or group of firms, that we interpret in respect to three benchmarks: the population' medians, the failed firms' medians, and the healthy firms' medians. Also, it is possible to form weightings for a certain score in respect to a subgroup (for example, an industry), where specific variables may be more relevant than others for discriminating failed from non-failed firms; starting from the scores, distances to failure, to a common situation, or to a healthy situation, are computed.

The methodology can be applied to any group where the number of failed firms is high enough to obtain benchmarks on the position of the ratios. We test the methodology on a sample of Spanish small firms located in Castille and Leon during the period 2006-08. Our results evidence that scores and distances to failure computed with differences of percentiles are efficient to identify failed, close-to-failed, and non-failed small firms, finding comparable or better results than those methodologies commonly used in previous literature.

Ratios on cash flow, return on assets, and indebtedness are found the most discriminant both before and after the change of standards. Results obtained on a specific geographical zone for 2006, 2007, and 2008 are confirmed on a sample taken from the whole country in 2010. Besides, our findings are in line with previous results found in literature for construction firms, and for SMEs in general.

Finally, unlike most of the other methodologies under study, ours offers a potential application in reorganization of close-to-failure firms. By computing individual distances to median values of failed firms (or to median values of the whole population, or to the mirror values representing healthy firms), the method allows us to identify each firms' strengths and weaknesses among the financial drivers used. Thus, in a specific period and a certain geographical zone, we identify those ratios contributing to increase the distance to failure and/or to reduce that distance.

A remarkable advantage of scores and distances to failure through percentiles is that it takes into account the whole population situation. By doing so, the effect of generalized movements of the population under study (industry, country, region) on scores and distances is avoided. The evolution of the whole population partially explains the evolution of financial drivers in failed and next-to-failure firms. That is why analyzing firms in a contemporary industry context is important.

REFERENCES

Altman, E.I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 568-609.

Altman, E.I., and Sabato, G. (2005). Effects of the new Basel capital accord on bank capital requirements for SMEs. *Journal of Financial Services Research*, 28(1-3), 15-42.

Altman, E.I., and Sabato, G. (2007). Modeling credit risk for SMEs: evidence from the U.S. market. *Abacus*, 43(3), 332-357.

Altman, E.I., Sabato, G., and Wilson, N. (2008). The value of qualitative information in SME risk management. *Working Paper*. Leonard N. Stern School of Business, New York University.

Aziz, M.A., and Dar, H.A., (2006). Predicting corporate bankruptcy: Where we stand? *Corporate Governance*, 6(1), 18-33.

Baixauli, J.S., and Módica-Milo, A. (2010). The bias of unhealthy SMEs in bankruptcy prediction models. *Journal of Small Business and Enterprise Development*, 17(1), 60-77.

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

Behr, P., and Güttler, A. (2007). Credit risk assessment and relationship lending: an empirical analysis of German small and medium-sized enterprises. *Journal of Small Business Management*, 45(2), April, 194-213.

Bell, T.B., Ribar, G.S., and Verchio, J. (1990). Neural nets versus logistic regression: a comparison of each model's ability to predict commercial bank failures, en Srivastava, R.P. (ed.) *Auditing Symposium on Auditing Problems*, págs. 29-53.

Castaño, F.J., and Tascón, M.T. (2012). Selection of variables in business failure analysis: Mean selection vs. median selection". *European Accounting Association 35th Annual Congress*, 9-11 May, Ljubljana (Slovenia).

Cochran, B. (1986). Small business mortality rates: a review of the literature. *Journal of Small Business Management*, 19(4), 50-9.

Cook, G.A.S., Pandit, N.R., and Milman, D. (2012). A resource-based analysis of bankruptcy law, SMEs and corporate recovery. *International Small Business Journal*, 30(3), 275-293.

Dannreuther, C., and Kessler, O. (2010): *Small firm finance and the political economy of risk*. London. Routledge.

Davydenko, S.A., and Franks, J.R. (2008). Do bankruptcy codes matter? A study of defaults in France, Germany, and the U.K.. *The Journal of Finance*, 63(2), April, 565-608.

Dietsch, M., and Petey, J. (2002). The credit risk in SME loans portfolios: Modeling issues, pricing, and capital requirements. *Journal of Banking and Finance*, 26(2-3), 303-322.

Dietsch, M., and Petey, J. (2004). Should SME exposures be treated as retail or corporate exposures? A comparative analysis of default probabilities and asset correlations in French and German SMEs. *Journal of Banking and Finance*, 28(4), 773-788.

Edmister, R.O. (1972). An empirical test of financial ratio analysis for small business failure prediction. *Journal of Financial and Quantitative Analysis*, 7(2), March, 1477-1493.

Franks, J.R., and Torous, W.N. (1992). Lessons from a comparison of US and UK insolvency codes. *Oxford Review of Economic Policy*, 8(3), 70-82.

Gill de Albornoz, B., and Giner, B. (2010). El fracaso empresarial en los sectores inmobiliario y de la construcción. 2008-2009", en Gill de Albornoz, B (Dir.), Fernández de Guevara, J, Giner, B. y Martínez, L. *Las empresas del sector de la construcción e inmobiliario en España: del boom a la recesión económica*, Madrid, Funcas, 173-231.

Headd, B. (2003). Redefining business success: Distinguishing between closure and failure. *Small Business Economics*, 21(1), 51-61.

Labatut, G., Pozuelo, J., and Veres, E.J. (2009). Modelización temporal de los ratios contables en la detección del fracaso empresarial de la pyme española. *Spanish Journal of Finance and Accounting*, 38(143), julio-septiembre, 423-448.

Laitinen, E.K. (2008). Data system for assessing probability of failure in SME reorganization. *Industrial Management and Data Systems*, 108(7), 849-866.

Lincoln, M. (1984). An empirical study of the usefulness of accounting ratios to describe levels of insolvency risk. *Journal of Banking and Finance*, 8(2), 321-340.

López, J., Gandía, J.L., and Molina, R. (1998). La suspensión de pagos en las pymes: una aproximación empírica. *Spanish Journal of Finance and Accounting*, 27(94), enero-marzo, 71-97.

Mar Molinero, C., and Ezzamel, M. (1991). Multidimensional scaling applied to corporate failure. *Omega*, 19(4), 259-274.

Marais, M., Patell, J., and Wolfson, M. (1984). The experimental design of classification models: An application of recursive partitioning and bootstrapping to commercial bank loan classifications. *Journal of Accounting Research*, 22(1), 87-118.

Martin, D. (1977). Early warning of bank failure. *Journal of Banking and Finance*, 1(3), 249-276.

Paradi, J.C., Asmild, M., and Simak, P.C. (2004). Using DEA and worst practice DEA in credit risk evaluation. *Journal of Productivity Analysis*, 21(2), March, 153-165.

Park, C.S., and Han, I. (2002). A case-based reasoning with the feature weights derived by analytic hierarchy process for bankruptcy prediction, *Expert Systems with Applications*, 23(3), 255-264.

Pompe, P.P.M., and Bilderbeek, J. (2005). The prediction of bankruptcy of small and medium sized industrial firms. *Journal of Business Venturing*, 20(6), November, 847–868.

Premachandra, I.M., Bhabra, G.S., and Sueyoshi, T. (2009). DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique. *European Journal of Operational Research*, 193(2), 412-424.

Psillaki, M., Tsolas, I.E., and Margaritis, D. (2010). Evaluation of credit risk based on firm performance. *European Journal of Operational Research*, 201(3), 873-881.

Serrano Cinca, C. (1996). Self organizing neural networks for financial diagnosis". *Decision Support Systems*, 17(3), 227-238.

Shin, K.S., and Lee, Y.J. (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications*, 23(3), 321-328.

Sueyoshi, T., and Goto, M. (2009). DEA–DA for bankruptcy-based performance assessment: Misclassification analysis of Japanese construction industry. *European Journal of Operational Research*, 199(2), 576-594.

Tascón, M.T., and Castaño, F.J. (2012). Variables y modelos para la identificación y predicción del fracaso empresarial: revisión de la investigación empírica reciente. Spanish Accounting Review, 15(1), enero-junio, 7-58.

Watson, J., and Everett, J. (1996). Do small business have high failure rates? *Journal of Small Business Management*, 34(4), 45-52.