



2nd GLOBAL CONFERENCE on BUSINESS, ECONOMICS, MANAGEMENT and  
TOURISM, 30-31 October 2014, Prague, Czech Republic

## A Comparison of Classification/Regression Trees and Logistic Regression in Failure Models

Irimia-Diequez, A.I.<sup>a\*</sup>, Blanco-Oliver, A.<sup>a</sup>, Vazquez-Cueto, M.J.<sup>a</sup>

<sup>a</sup>Faculty of Economics and Business Administration, University of Seville, Av. Ramon y Cajal, s/n, 41018 Seville (Spain)

---

### Abstract

The use of non-parametric statistical methods, the development of models geared towards the homogeneous characteristics of corporate sub-populations, and the introduction of non-financial variables, are three main issues analysed in this paper. This study compares the predictive performance of a non-parametric methodology, namely Classification/Regression Trees (CART), against traditional logistic regression (LR) by employing a vast set of matched-pair accounts of the smallest enterprises, known as micro-entities, from the United Kingdom for the period 1999 to 2008 that includes financial, non-financial, and macroeconomic variables. Our findings show that CART outperforms the standard approach in the literature, LR.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of Academic World Research and Education Center

*Keywords:* Micro-entities, Failure models, Financial ratios, Non-financial information, Macroeconomic variables, Logistic regression, Classification/Regression Trees.

---

### 1. Introduction

In recent years, three new research lines have appeared which strive to improve the performance of bankruptcy models: (a) introducing non-financial information as predictor variables (Grunert, Norden & Weber, (2005); (b) developing models specifically designed for each firm feature, such as size and sector (Altman, & Sabato, 2007); and (c) implementing non-parametric statistical techniques to fit the bankruptcy models (Jagric, Kracun & Jagric, 2011).

This study deals with these three advances developing a failure model specifically designed for the smallest micro-enterprises: micro-entities (hereafter, MEs), which have recently been defined by the Competitiveness Council of the European Union as those companies with an annual turnover of less than €700,000, total assets

---

\*Irimia-Diequez, A.I. Tel.: +34-954-559875; fax: +34-954-557570.  
E-mail address: [anairimia@us.es](mailto:anairimia@us.es)

of less than €350,000, and average number of employees during the financial year of no more than 10, (Official Journal of the European Union, 2012). The particular characteristics of MEs: (a) higher probability of failure (Carter, & Van Auken, 2006); (b) great limitation of publicly available financial information, due to the fact that they file abridged accounts; and (c) the inexistence of failure models adapted to this types of firms despite of their leading role in the economic activity worldwide, justify the need of the development of specific failure models. Furthermore, we also test the accuracy capacity to detect the failure of a non-parametric statistical technique Classification/Regression Trees (CART), in comparison with the classic logistic regression (LR) analysis. In this sense, previous literature shows that CART often outperforms LR in the failure environment [6].

After the implementation of the Basel Accord regulation (Basel II), considerable studies have been undertaken in an effort to predict the failure of SMEs. Whereas the importance of financial factors is widely accepted, the relevance of non-financial predictors appears to need more empirical evidence. Moreover, nothing is known about the applicability of default prediction models to MEs, and whether non-financial information improves the predictive capacity of models developed specifically for them due to the lack of research that deals with these kinds of firms.

One of the most relevant models specifically made for SMEs was developed by (Altman, & Sabato, 2007). Their study compares the traditional Z-score model with two new models which consider other financial variables and use traditional logistic regression. On a panel of data of over 2,000 US SMEs in the period 1994-2002, these authors find that the new models outperform the traditional Z-score model by almost 30 per cent, in terms of prediction power. Based on the above research, (Altman, & Sabato, 2007) they explore the effect of the introduction of non-financial information as predictor variables into the models developed by (Altman, & Sabato, 2007). Employing a large sample (5.8 million) of sets of accounts of unlisted firms from the U.K. in the period 2000-2007, they find that non-financial information makes a large contribution (by approximately 13% in terms of the area under the receiver operating characteristics curve, henceforth, AUC) towards increasing the default prediction power of risk models.

Therefore, the main objective of this study is to compare LR and CART in the building of a failure model designed for MEs which introduce financial, non-financial and macroeconomic variables. The large size of the sample (almost 40,000 set of accounts of MEs) is an important strength for the reliability of our findings. Moreover, the use of a parsimonious model constitutes a noteworthy improvement.

In Section 2, we provide details of our sample and methodology carried out. In Section 3, several failure models for MEs are developed, comparing LR and CART approaches. In Section 4, the results are shown and discussed them. Finally, Section 5 provides the main conclusions and future lines of research.

## 2. Data set

A dataset provided by a U.K. Credit Agency is used in this study. After eliminating missing and abnormal cases and selecting a random sample of MEs, 39,710 sets of accounts of MEs (50% non-failed) for the period 1999-2008 remained. In line with other studies, we define corporate failure as entry into liquidation, administration or receivership between 1999 and 2008. The accounts analyzed for failed companies are the last set of accounts filed in the year preceding insolvency. For each case, the dependent variable takes the value 1 when the ME failed, and 0 otherwise. Finally, to run the models, our final dataset was randomly split into three sub-sets; a training set of 60%, a validation set of 20%, and a test data set (or hold-out sample) of 20% (Hastie, Tibshirani & Friedman, 2009).

Table 1 describes the variables considered in this study and the theoretical relationship with the failure of the firm. All the financial ratios used in this study were employed in prior research, such as Altman, (1968); Altman, Sabato & Wilson, (2010); Ohlson, (1980); Taffler, (1984) and Zmijewski (1984). Moreover, based in the findings of Carter, Van Auken, (2006), it seems reasonable to assume that an adequate failure model made specifically for MEs should also introduce non-financial information. Finally, since several studies have shown a positive relationship between the adverse economic cycle and the number of corporate failures Moon, and Sohn (2010), we also include a macroeconomic variable (*Industry\_solvency*) which measure the financial health of the sector in which operate the firm and is inverse of the probability of bankruptcy of the sector.

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات