

Predicting SME's default: some old facts and a new idea

Lisa Crosato¹, Josep Domenech ², Caterina Liberati³

¹Department of Economics, Ca' Foscari University of Venice, Italy, ²Department of Economics and Social Sciences, Universitat Politècnica de Valencia, Spain, ³Department of Economics Management and Statistics, University of Milano-Bicocca, Italy.

Abstract

The Small Business Act of the European Commission in 2008 acknowledges the key role of Small and Medium Enterprises (SMEs) in the EU economy. This is particularly relevant for Italy, which has the largest share of SMEs in Europe, as well as for other countries such as Portugal, Spain and Greece. On the other hand, SMEs experience more difficulties in their early stages mainly due to high market competition and credit constraints, as highlighted by Fritsch and Weyh (2006). For these reasons, the study of SMEs default risk is always relevant. There are several papers studying firm default factors in a single country (see Ciampi, 2015, Fantazzini and Figini, 2009, Flix and dos Santos, 2018). The literature concentrates mainly on financial indicators built on businesses' balance sheets, which are available about two years late with respect to their reference period. This diminishes the significance of the results, both for credit risk and policy aims, and particularly in a forecasting perspective. The purpose of this paper is to provide a preliminary study on a sample of Spanish firms selected from the SABI, Sistema de Análisis de Balances Ibéricos, which is listed among Bureau van Dijk databases. The analysis will be carried out according to both parametric and non-parametric discrimination techniques, with the standard construction of a training set on which to build a model and a validation set to test the validity and robustness of the results, and, in the end, the reliability of the model in predicting default. Finally we present a new proposal: a scheme to understand to what extent firms' default can be predicted by substituting the traditional data sources (offline information) with data collected from their corporate websites (online information) in order to exploit more up-to-date information.

Keywords: Default risk; SMEs; Web scraping; Corporate websites.
