Book of Short Papers SIS 2020

Editors: Alessio Pollice, Nicola Salvati and Francesco Schirripa Spagnolo

PUBLISHED BY PEARSON

WWW.PEARSON.COM

ISBN 9788891910776

Contents

Specialized sessions

Solicited Sessions

Contributed papers and Posters

SIS 2020 Pisa

Preface

The COVID-19 pandemic is putting our society under incredible health, emotional, and economic stress. Facing its harmful effects and their uncertainty, the Executive Board of the Italian Statistical Society (SIS) and the Local Organizing Committee, to ensure the highest level of safety for members and delegates, deliberated to cancel the 50th Meeting of the Italian Statistical Society originally planned to be held in Pisa in June 2020 and to postpone the conference to June 2021. The Executive Board and the Local Organizing Committee continue to monitor closely the pandemic evolving situation, and keep the members of SIS and the researchers informed about the potential new dates for the next meeting. To give value to the work of those who prepared their presentation for the conference, the Program Committee decided to publish the volume *Book of short papers - SIS 2020* despite the conference cancellation.

The conference program included 4 plenary sessions, 16 specialized sessions, 24 solicited sessions, 32 contributed sessions and the poster exhibition. Plenary sessions concerned with robust statistics, human longevity, statistical models for climate changes and small area estimation for educational poverty. The meeting had to host also 2 round tables on data privacy and innovation in statistics. Activities focused on topics of interest for a wider audience included two round tables on Teaching Statistics and on the SIS journal Statistical Methods & Applications, and the Stats Under the Stars (SUS6) competition for young statisticians. The SUS6 event attracted many sponsors from statistical, financial and editorial firms as well as numerous students. The conference committee had registered 345 accepted submissions, including 143 to be presented in invited plenary, specialized and solicited sessions, and 202 spontaneously submitted for oral and poster sessions.

This book includes most of the scientific contributions that had to be presented at the 50th Meeting of the Italian Statistical Society. It is organized into 49 chapters corresponding to 15 specialized, 23 solicited sessions, and to 11 general topics for contributed papers and posters. All 268 contributions provide a wide overview of the state-of-the-art of the subjects, from methodological and theoretical contributions, to applied works and case studies. The result is a very lively picture of the Italian statistical community with its international connections.

We would like to thank all contributors for having submitted their work to the conference, the members of the Program Committee and the extra reviewers for their efforts in this difficult period. Although the Conference did not take place, the organization went on until cancellation was decided for safety reasons. It would have been impossible without the joint effort of Università di Pisa, Scuola Superiore Sant'Anna and National Research Council of Pisa. Members these three institutions took part actively in the Local Organizing Committee. Finally we wish to express our gratitude to the publisher Pearson Italia for all the support received.

This book is our contribution to encourage the scientific community and the network of the Italian Statistical Society to go on and transform this difficult period into an opportunity of scientific debate for better statistics in a better world.

Alessio Pollice
Università degli Studi di Bari Aldo Moro
Chair of the Program Committee

Nicola Salvati
Università di Pisa
Chair of the Local Organizing Committee

Francesco Schirripa Spagnolo
Università di Pisa

Program Committee: Alessio Pollice (Chair), Serena Arima, Marilena Barbieri, Alessandra Brazzale, Eugenio Brentari, Alessia Caponera, Antonio Lepore, Antonella Plaia, Tommaso Proietti, MarcoRiani, Nicola Salvati, Pasquale Sarnacchiaro, Mauro Scanu, Manuela Stranges, Valentina Tocchioni, Simone Vantini, Massimo Ventrucci, Paola Vicard, Donatella Vicari.

Local Organizing Committee: Nicola Salvati (Chair), Gaia Bertarelli, Bruno Cheli, Alessandra Coli, Paolo Frumento, Fosca Giannotti, Caterina Giusti, Piero Manfredi, Stefano Marchetti, Lucio Masserini, Vincenzo Mauro, Barbara Pacini, Dino Pedreschi, Francesco Schirripa Spagnolo, Chiara Seghieri.

Organizers of Specialized and Solicited Sessions: Giada Adelfio, Bruno Arpino, Emanuele Aliverti, Nicoletta Balbo, Mara Bernardi, Silvia Bozza, Pierpaolo Brutti. Annalisa Busetta, Michela Cameletti, Carlo Cavicchia, Fabrizio Durante, Leonardo Egidi, Pietro D. Falorsi, Francesco Finazzi, Livio Finos, Stefania Galimberti, Michele Gallo, Caterina Giusti, Francesca Greselin, Alessandra Guglielmi, Francesca Ieva, Tiziana Laureti, Achille Lemmi, Brunero Liseo, Fabio Massimo Lo Verde, Daria Mendola, Roberta Pappadà, Lea Petrella, Alessandra Petrucci, Alessia Pini, Sabrina Prati, Maria Giovanna Ranalli, Davide Risso, Fabrizio Ruggeri, Silvana Salvini, Monica Scannapieco, Francesco Stingo, Luca Tardella, Grazia Vicario, Susanna Zaccarin, Maroussa Zagoraiou.

Determinants of Firms' Default Risk after the 2008 and 2011 Economic Crises: a Latent Growth Models Approach

Rischio di bancarotta dell'ultimo decennio per le imprese manifatturiere italiane: un modello di curva a crescita latente

Lucio Masserini, Matilde Bini and Alessandro Zeli

Abstract Between 2008 and 2009, European countries became mired in the Great Recession, derived from the US subprime crisis, with financial contraction and bank failures spreading out across the Atlantic and causing damage to financial markets and the real economy. Then, starting in 2011, the crisis affected the euro zone and the sovereign debts of European countries until 2014. For Italy, the crisis period led to a deep negative conjuncture until 2009 and then to a slight recovery until the first half of 2011 and an intense recession from 2011 to 2015. The aims of this work are to identify some determinants of the defaults of Italian firms during the years 2008–2012 and to understand the effects of the Great Recession on the default probability. To perform this analysis, a Latent Growth Curve Model is proposed, using an important Italian private database of Italian firms.

Abstract *In Italia, il periodo di crisi economica è stato caratterizzato da una profonda congiuntura negativa fino al 2009 e da una leggera ripresa fino alla prima metà del 2011 e poi da un'intensa recessione dal 2011 al 2014. Lo scopo di questo lavoro è di studiare gli effetti della grande crisi sul rischio di fallimento delle imprese manifatturiere nel periodo post crisi (2017) e di misurare l'impatto di alcuni fattori su tale rischio. Viene proposto un modello di Curva di crescita latente, sui dati contabili delle società di capitale italiane.*

[1] Lucio Masserini, Department of Economics and Management, University of Pisa; email: lucio.masserini@unipi.it

Matilde Bini, Department of Human Sciences, European University of Rome; email: matilde.bini@unier.it

Alessandro Zeli, Division for Data Analysis and Economic, Social and Environmental Research, Italian National Statistical Institute; email: zeli@istat.it

Key words: Bankruptcy, Firms' default, Interest coverage ratio, Latent Growth Curve Model, Panel data

1 Introduction

In Italy, the crisis period from 2007 to 2014 can be divided into two sub-periods: the first, from 2008 to 2011, led to a deep negative conjuncture until 2009 with a subsequent slight recovery until the first half of 2011, while the second led to an intense recession, begun in 2011 and enduring until 2014. The aim of this paper is to analyse firms' defaults, using information on financial ratios from a large panel of Italian manufacturing firms. Since in the majority of financial literature major attention is devoted to financial conditions [1,2,3,6,7], this study intends to take into consideration more factors characterising a firm that, together with financial and economic conditions, can determine a firm's success or failure. Hence, the main motivation of this study is to assess whether classification variables such as a firm's size, industrial sector, and territorial location can better explain its default risk. An important feature of this study is the large number of enterprises included in the sample, covering a large range of sizes, economic activities, and geographic locations. The analysis is carried out using a statistical approach based on a Latent Growth Curve Model (LGCM) with an Italian private database containing the book-value data of the joint-stock company Italian firms. Unlike traditional longitudinal data analysis techniques, a LGCM allows inferences to be made about individual level effects as well as group effects [4]. Results we obtained allow us to answer the main research questions, such as whether the economic cycle influenced the default risk, how the default risk differs throughout different industries, and whether cumulative indicators have an effect on default probability.

2 Italian manufacturing database

We used a private database called *Analisi Informatizzata delle Aziende Italiane* (AIDA) that contains the book-value data of 16,383 joint-stock Italian firms over the period 2008–2017. First, we checked and controlled the data in order to avoid any inconsistencies in the items of the financial statement and strong outliers among all variables and indicators of interest. We included firms that underwent insolvency proceedings (e.g. failure, liquidation, and extraordinary administration). The target population is represented by Italian manufacturing firms (traced from 2008 to 2017). As a result, we ended up with a balanced panel of 7,689 companies that were used in the statistical analysis, repeated for five years.

3 Latent Growth Curve Model

The analysis was carried out using a Latent Growth Curve Model [5], which assumes the existence of latent trajectories for each firm, measured by the repeated values of the response variable (y_i) over time. In this analysis, the response variable is Interest Coverage, whose values are measured on a continuous scale and observed from 2008 to 2012. This approach was chosen for its flexibility in longitudinal data modelling: it allows for the estimation of different functional forms, the incorporation of both time-varying covariates (TVCs) and time-invariant covariates (TICs), as well as the assessment of the goodness of fit through indices (see, for more details, [9]) and the possibility of dealing with missing data [8]. In general, the LGCM can be expressed in terms of a trajectory equation and a structural model. In particular, given $\mathbf{y}_i$, a $T \times 1$ vector of repeated observed measures for firm *i* at time points $t = 1,2 \ldots T$, the trajectory equation, is defined as it follows:

$$\mathbf{y}_i = \boldsymbol{\Lambda}\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i$$

$\boldsymbol{\eta}_i$ is the $m \times 1$ vector of the underlying latent factors which identifying the latent growth factors (e.g. initial status and trend factors); $\boldsymbol{\Lambda}$ is a $T \times m$ matrix of factor loadings for T time points, whose elements are fixed to specify the functional form for the individual trajectories; $\boldsymbol{\varepsilon}_i$ is a random vector of time-specific error terms. On the other hand, the structural model can be formulated as follows:

$$\boldsymbol{\eta}_i = \boldsymbol{\mu}_\eta + \boldsymbol{\Gamma}\mathbf{x}_i + \boldsymbol{\zeta}_i \, .$$

$\boldsymbol{\mu}_\eta$ is an $m \times 1$ vector of factor means; $\boldsymbol{\Gamma}$ is an $m \times k$ matrix of regression coefficients between the latent factors and the observed covariates; $\boldsymbol{\zeta}_i$ is an $m \times 1$ vector of error terms, which represents the random components of the model. Finally, a further response variable u_i (Bankruptcy 2017) is added as a distal outcome, whose values can be predicted by the latent growth factors, as follows:

$$u_i = \tau + \boldsymbol{\beta}\boldsymbol{\eta}_i + \boldsymbol{\upsilon}_i$$

Here, the vector of regression coefficients $\boldsymbol{\beta}$ represents the effects of the growth factors on the distal outcome, τ is the intercept and $\boldsymbol{\upsilon}_i$ are the error terms. The final model is shown in Fig. 1, which summarizes graphically the whole system of equations. The rectangular boxes in this figure represent the observed variables (repeated outcome variables, as well as TICs and TVC), while the circles represent the latent variables (growth factors):

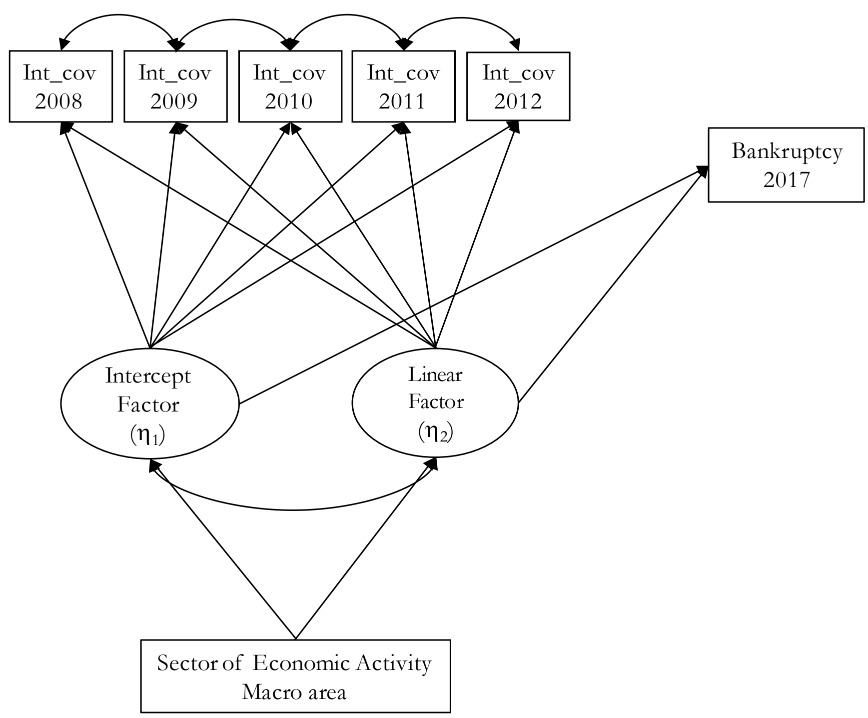

Figure 1: Path diagram of the LGCM

4 Results

Below, the results from the analysis performed by fitting a LGCM are shown. First, a set of alternative unconditional LGCMs with correlated measurement errors between adjacent time points were estimated (such as linear, quadratic, piecewise linear with two knots, and latent basis). These estimates allowed for the identification of a more suitable functional form for the individual latent trajectories. The model parameters were estimated by using the WLSMV (Weighted Least Square Mean and Variance adjusted) estimator with robust standard errors [10]. The model's goodness of fit was evaluated by RMSEA (Root Mean Square Error of Approximation), CFI (Comparative Fit Index), TLI (Tucker-Lewis Index) and SRMR (Standardized Root Mean Squared Residual) [9]. Based on the results, the linear form LGCM was preferred given the satisfactory values of the indices

(RMSEA = 0.041, CFI = 0.999, TLI = 0.997, SRMR = 0.023). The corresponding results show an adequate fit of the estimated model.

Table 1: LGCMs model parameter estimates for the response variable Interest coverage

Parameter	***Estimate***	***P-value***
Intercept Factor Mean (μ_1)	15.794	0.000
Intercept Factor Variance (ψ_{11})	1609.466	0.000
Linear Factor Mean (μ_2)	0.987	0.012
Linear Factor Variance (ψ_{22})	84.385	0.000
Correlation Intercept Factor vs. Linear Factor (ψ_{12})	0.084	0.429

Results in Table 1 show that the starting point of the response variable Interest Coverage is 15.794 with a significant and remarkable variability (1609.466). Moreover, Interest Coverage shows a linear growth trend (0.987) which also exhibits a significant variability. Differences in growth parameters can be analysed by sector of economic activity and a firm's geographical location in Table 2, where the asterisks show significant differences with respect to the average at the 0.05 significance level.

Table 2: Differences of growth curve parameters relative to baseline level for various sectors of economic activity and geographic location

Parameter	***Intercept***	***Slope***
Food and Tobacco	1.825	0.852
Textile and Leather	-6.775	1.886
Wood, Publishing, and Paper	6.171	-2.094
Refining	1.159	1.723
Chemistry and Rubber	-0.234	0.258
Metallurgy and Steel Industry	2.578*	-1.035
Electric Machines and Mechanical	7.628	0.486
Transport	4.762	-3.008*
Other Industries and Maintenance (Reference Category)	5.431	1.418
North	7.823*	0.197
Centre	6.515	0.174

Firms in the metallurgy and steel industry sector and those located in the North of Italy show a higher level of interest coverage at the beginning of the observed period. On the other hand, a significant and lower rate of growth is found in those firms that belong to the transport sector.

Table 3 shows the parameter estimates where Interest Coverage is used to predict a firm's situation in terms of risk of failure after five years, as measured by Bankruptcy 2017, introduced as a distal outcome. In particular, this approach allows for the evaluation of the Interest Coverage trend during the crisis period (2008–2012) and the impact of this trend on the risk of failure after five years.

Table 3: Effects of Interest coverage on Bankruptcy 2017

Parameter	***Estimate***	***P-value***
Intercept (τ)	1.360	0.000
Intercept Factor (β_1)	-0.004	0.002
Linear Factor (β_2)	-0.064	0.000

The Latent Growth Curve Model successfully detects a firm's outcome in terms of bankruptcy five years later. Indeed, the bankruptcy risk is well predicted by the linear function, and the model estimates show that in the considered period, industries in which Italian manufacturing is stronger in terms of interest coverage have a lower risk of bankruptcy.

References

1. Alaka, H.A., Oyedele, L.O., Owolabi, H.A., Kumar, V., Ajayi, S.O., Akinade, O.O., Bilal, M.: Systematic review of bankruptcy prediction models: Towards a framework for tool selection. *Expert Systems with Applications*, 94, 164–184 (2018)
2. Becchetti, L., Sierra, J.: Bankruptcy risk and productive efficiency in manufacturing firms. *Journal of Banking and Finance*, 27: 2099-2120 (2003)
3. Bellovary, J.L., Giacomino, D.E., Akers, M.D.: A review of bankruptcy prediction studies: 1930 to Present. *Journal of Financial Education*, 33, 1–42 (2007)
4. Bini, M., Masserini, L., Zeli, A.: A longitudinal analysis of riskiness indicators after the 2008 and 2011 economic crises: the case of Italian manufacturing. *Social Indicators Research*, accepted (2020).
5. Bollen, K. A., Curran, P. J.: *Latent Curve Models*. New York: Wiley (2006)
6. Bottazzi, G., Grazzi, M., Secchi, A., Tamagni, F.: Financial and economic determinants of firm default. *Journal of Evolutionary Economics*, vol. 21, issue 3, 373–406 (2011)
7. Dimitras, A. I., Slowinski, R., Susmaga R., Zopounidis, C.: Business failure prediction using rough sets. *European Journal of Operational Research*, vol. 114, issue 2, 263–280 (1999)
8. Enders, C.: A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modelling*. 8(1), 128–141 (2001)
9. Hu, L., Bentler, P. M.. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55 (1999)
10. Muthén, B., Satorra, A.: Technical aspects of Muthén's LISCOMP approach to estimation of latent variable relations with a comprehensive measurement model. *Psychometrika*, 60(4), 489–503 (1995)