Springer Proceedings in Mathematics & Statistics

Eugenio Brentari
Marcello Chiodi
Ernst-Jan Camiel Wit *Editors*

Models for Data Analysis

SIS 2018, Palermo, Italy, June 20–22

SIS Società Italiana di Statistica

Springer

Springer Proceedings in Mathematics & Statistics

Volume 402

This book series features volumes composed of selected contributions from workshops and conferences in all areas of current research in mathematics and statistics, including data science, operations research and optimization. In addition to an overall evaluation of the interest, scientific quality, and timeliness of each proposal at the hands of the publisher, individual contributions are all refereed to the high quality standards of leading journals in the field. Thus, this series provides the research community with well-edited, authoritative reports on developments in the most exciting areas of mathematical and statistical research today.

Eugenio Brentari · Marcello Chiodi ·
Ernst-Jan Camiel Wit
Editors

Models for Data Analysis

SIS 2018, Palermo, Italy, June 20–22

Springer

Editors
Eugenio Brentari
Department of Economics and Management
University of Brescia
Brescia, Italy

Marcello Chiodi
Department of Economics, Business and Statistics
University of Palermo
Palermo, Italy

Ernst-Jan Camiel Wit
Statistical Computing Laboratory, Faculty of Informatics
Università della Svizzera Italiana
Lugano, Switzerland

ISSN 2194-1009 ISSN 2194-1017 (electronic)
Springer Proceedings in Mathematics & Statistics
ISBN 978-3-031-15884-1 ISBN 978-3-031-15885-8 (eBook)
https://doi.org/10.1007/978-3-031-15885-8

Mathematics Subject Classification: 62-06, 62-07, 62H, 62J, 62P12, 62P20, 62P25

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Organization

Program Chairs

Ernst Wit, Universita della Svizzera italiana, Switzerland
Marcello Chiodi, University of Palermo, Italy
Francesca Ferrari, Springer
Eugenio Brentari, University of Brescia, Italy

Program Committee

Eugenio Brentari, University of Brescia, Italy
Filomena Racioppi, Sapienza, University of Rome, Italy
Monica Pratesi, (chair SIS) University of Pisa, Italy
Marcello Chiodi, University of Palermo, Italy
Annamaria Bianchi, University of Bergamo, Italy
Francesco Maria Chelli, Marche Polytechnic University, Italy
Andrea Cutillo, Italian National Institute of Statistics, Italy
Alessandra De Rose, Sapienza University of Rome, Italy
Daniele Durante, Bocconi University, Italy
Fedele Greco, University of Bologna, Italy
Leonardo Grilli, University of Florence, Italy
Piero Manfredi, University of Pisa, Italy
Marica Manisera, University of Brescia, Italy
Antonietta Mira, University of Svizzera italiana, Switzerland and University of Insubria, Italy
Marcella Niglio, University of Salerno, Italy
Francesco Palumbo, University of Naples, Italy
Elvira Romano, Second University of Naples, Italy
Nicola Tedesco, University of Cagliari, Italy

Preface

The 49th Scientific meeting of the Italian Statistical Society was held in June 2018 in Palermo, with more than 450 attendants.

The meeting covered a wide area of applied as well as theoretical issues, according to the modern trends in statistical sciences. There were three plenary sessions on Rating data, Global climate dynamics and Computational methods in biomedicine; twenty specialized and twenty-five solicited sessions, as well as oral and poster-contributed sessions. The whole program is still available at the conference site: http://meetings3.sis-statistica.org/index.php/sis2018/49th.

This book presents a selection of contributions of Authors who presented an extended version of their conference presentation, which are original unpublished works. As usual, every contribution has been submitted to standard double-blind refereeing, and the final result is a picture of the state of the art of statistical research in recent years.

Unfortunately, there was some delay in publishing the final version also due to the COVID-19 pandemic situation, but finally, we are here to present a collection of twenty papers that deal with modern methodological as well as applied topics of models for data analysis.

To have a graphical idea of what SIS 2018 has been like, perhaps the best thing is to reproduce the conference word cloud.

Palermo, Italy
June 2022

Eugenio Brentari
Marcello Chiodi
Ernst-Jan Camiel Wit

Contents

Environmental Vibration Data Analysis for Damage Detection on a Civil Engineering Structure

Gianna Agrò

Abstract Management of the useful life and safety performance of the infrastructure of a motorway network is currently a topic of great interest. The dynamic behavior of civil engineering structures has usually been studied by means of ambient vibration observations, analyzing them with methods of Operational Modal Analysis and the so-called Peak-Picking technique. The present paper reports the results of an alternative multivariate statistical approach, specifically Principal Component Analysis (PCA), for ambient vibration data, collected to detect suspected structural damage on some specific highway bridge spans in Sicily. The method consists in comparing the system structures of undamaged spans with that coming from a suspected damaged one, as designated by optimal subspaces determined by PCA. The distance between the subspaces, measured by means of the maximum principal angle between them, provides evidence regarding damage in the span under investigation.

Keywords Principal component analysis · Damage detection · Subspace angles · Operational modal analysis · Fast fourier transform · Peak picking technique

1 Introduction

The most common procedure for damage detection in engineering structures relies on an ambient vibration test, requiring the gathering of natural frequencies and obtaining the so-called mode shapes and structural dampings that they exhibit. The test is conducted using a predetermined number **p** of uniaxial piezometric accelerometers connected to a control unit for the acquisition of **n** environmental acceleration times. The time series data are transferred to the frequency domain by means of a Fast Fourier Transform (FFT), and the "Peak-Picking technique" is used to extract the dynamic parameters from the spectral density matrices. However, sometimes the method is not sufficiently sensitive in identifying damage when it is present in the structure.

G. Agrò (✉)
DSEAS, University of Palermo, Palermo, Italy
e-mail: gianna.agro@unipa.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_1

A more recent multivariate statistical approach to damage identification in engineering structures is Principal Components Analysis (PCA). This can be usefully applied to investigate the existence of any change between a suspected damaged structure and an undamaged similar one adopted as a reference model. The analysis is conducted using the same data gathered by the control unit for the acquisition of environmental acceleration, but leaving them in time domain.

Section 2 of this paper outlines the principle components method as might be used in such an application, highlighting its use in practical statistical inference. In Sect. 3, the method is applied to the case of a Sicilian motorway bridge, and the encouraging results are illustrated. Section 4 presents inferential aspects in terms of significance tests obtained by bootstrap methods. Finally, Sect. 5 reviews the conclusions regarding improvements in analysis achieved using the principle components technique.

2 The Statistical Method to Detect Structural Damage

As a starting point for data analysis, an environmental acceleration time series over **n** periods is obtained from each of **p** sensors, constituting a data matrix $X_{n,p}$. The aim of PCA is to reduce the space of **p** correlated variables into **k < p** uncorrelated ones, in such a way that the size of the analytic problem is reduced while the bulk of information contained in the original data is preserved.

In this endeavour, a correlation matrix $\mathbf{R_{p,p}}$ is calculated from data in the matrix $\mathbf{X_{n,p}}$ in such a way that the eigenvectors and eigenvalues of **R** can be used to identify an optimal subspace of dimension **k < p** which retains a fixed amount of the system variance [1]. This procedure is adopted for the matrices of two observed series, one generated from the system structure known to be healthy, $\mathbf{X_{n,p}}$, and the other from the structure that is suspected to be damaged, $\mathbf{Y_{n,p}}$. The optimal subspace matrices of rotation coefficients for the healthy and damaged systems, denoted by H and D respectively, are compared so to assess evidence whether the suspect structure is really damaged or not. The assessment is evaluated on the basis of the distance between two subspaces.

The distance between the two subspaces is defined as the largest principal angle θ between the subspaces defined in R^k in which they reside [2, 3]. The principle angles $\theta_1 \ldots \theta_k \in [0, \pi/2]$ between the subspaces H and D in R^k are defined recursively by:

$$cos(\theta_i) = \max_{d \in D} \max_{h \in H} d^T h = d_i^T h_i \tag{1}$$

subject to:

$$|\mathrm{d}| = |\mathrm{h}| = 1$$

$$d^T d_j = 0 \, and \, h^T h_j \; j = 1 : i - 1$$

Note that the principal angles satisfy $0 \leq \theta_1 \ldots \leq \theta_k \leq \pi/2$; the vectors $\{d_1,\ldots,d_k\}$ and $\{h1,\ldots,h_k\}$ are called the principal vectors between the subspaces H and D. The largest principal angle is related to the notion of distance between equidimensional subspaces; so the distance between the subspaces is defined as:

$$\text{dist(H, D)} = \sqrt{1 - \cos(\theta_k)^2}$$

If the largest principal angle θ_k is not zero, then the compared subspaces are different. In our context, this may mean that if one of them is healthy then the other may be damaged.

To formulate a hypothesis test for the significance of $\cos(\theta_k)$, the sampling distribution of $\cos(\theta_k)$, may be determined by means of a bootstrap procedure. In the computations I present below, specifically the "moving blocks" bootstrap was adopted [4]. Briefly, to generate bootstrap realizations of a time series, a block length *l* is chosen and a number *m* of blocks of this length are sampled from within the time series. Overlap is allowed. Continuing this procedure, a bootstrap series is recomposed having the same length as the original one.

In the application reported below, the moving blocks bootstrap was used to obtain the bootstrapped matrices $s\mathbf{X_i}$ and $s\mathbf{Y_i}$ i = 1,..,5000. In particular, the columns of matrices $\mathbf{X}_{n,p}$ and $\mathbf{Y}_{n,p,}$ with $\mathbf{n} =$ 240,000 and $\mathbf{p} =$ 6, as they will detail in the next paragraph, were shuffled, choosing block length l = 30,000 rows (5 min), to obtain simulated matrices $s\mathbf{X}_i$ and $s\mathbf{Y}_i$ respectively.

3 Damage Detection: The Case of a Sicilian Highway Bridge

The case study concerns a Sicilian highway bridge [5], built in the 70 s. Along its 1533 m length it includes 35 reinforced concrete spans. The bridge consists of isostatic spans, each of which has a deck consisting of four pre-stressing R.C. beams with a double T section. Each span is 45 m long and 9.8 m wide. Ambient vibration tests were conducted on two adjacent spans: n. 8, considered to be intact and n. 9, suspected of being damaged. Figure 1 displays the layout of the sensors. Although $\mathbf{p} = 10$ piezoelectric accelerometers which were used as sensors, the present study was conducted using only six vertical sensors, shown in red colour. The other four shown in the Figure are horizontal, paired in orthogonal directions. The environmental acceleration time histories were recorded over a time duration of 2400 s, at intervals of 0.01 s. So data were collected in matrices of dimension (240,000 × 6).

To begin a standard analysis, by performing a Fast Fourier Transform (FFT) the time series were transferred to the frequency domain, and by means of Operational Modal Analysis the natural frequencies [6, 7] and mode shapes were identified. In particular, the so-called Peak Picking technique consists of observing the graphic representation of the spectrum and extracting the values of the dynamic parameters

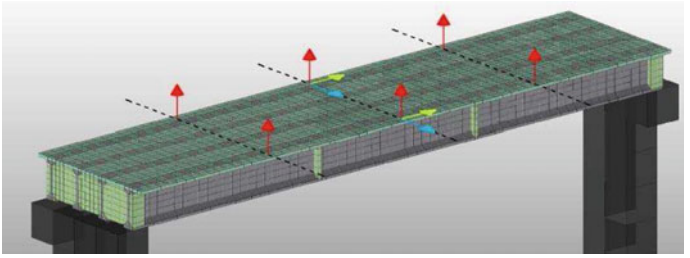


Fig. 1 Location of the sensors on the bridge deck

of the structures at the peaks. In order to investigate the existence of any change between the suspected damaged structure span n.9 and the undamaged similar one span n.8, the frequency of the spectra of the two spans (Fig. 2) are compared. As is evident, the difference in abscissa values of the peaks is very low. Consequently, it is difficult to detect any substantial difference between the structural parameters of the spans using this method of analysis. In conclusion, nothing can be inferred about damage in the questionable span n.9.

An alternative analysis can be provided by the multivariate statistical method of principle components. The application of the PCA to study the behaviour of the healthy span n.8 (matrix X) and the suspected damaged span n.9 (matrix Y) led to the identification of optimal four-dimensional subspaces for H and D, since the percentages of the total variance explained were 75% for healthy span and 80% for damaged span using this dimensional reduction.

In order to measure the difference among the spans, the principal angles between the four-dimensional subspaces were calculated by formula (1) where $D_{(6\times4)}$ and $H_{(6x4)}$ are the matrices of eigenvectors for the damaged and healthy span respectively.

As seen in Table 1 which shows the values of $\cos(\theta_k)$ obtained, it is apparent that a sizeable difference between the two spans exists, allowing us to conclude that the span n.9 is damaged: $\cos(\theta_k) = 0.7429$ and the angle $\theta = \arccos(0.7429) = 42°$.

Since the main angle θ between the four-dimensional subspaces is between PC1_H versus PC1_D, then the angle could be highlighted by representing the projection of

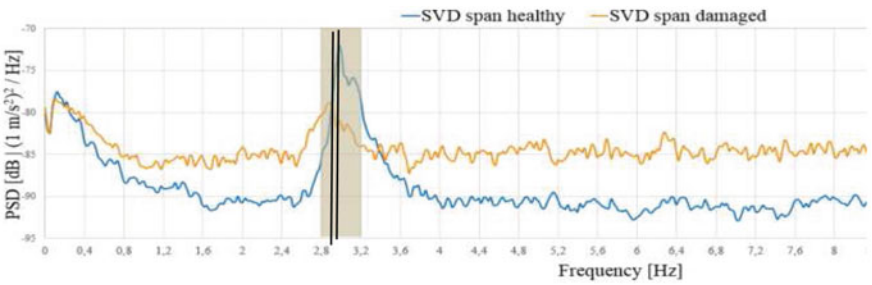


Fig. 2 Singular Values of Spectral Densities of Test Setup

Table 1 Cos(θ) for principal angles θ

H vs D	*PC1*	*PC2*	*PC3*	*PC4*
PC1	**0.7429298**	0.2386534	0.3945228	−0.3908226
PC2	0.5139214	0.2406920	−0.2453405	0.3750808
PC3	0.3328730	−0.1497690	−0.6219700	0.3301877
PC4	0.1773170	−0.6197478	−0.3634516	−0.6058055

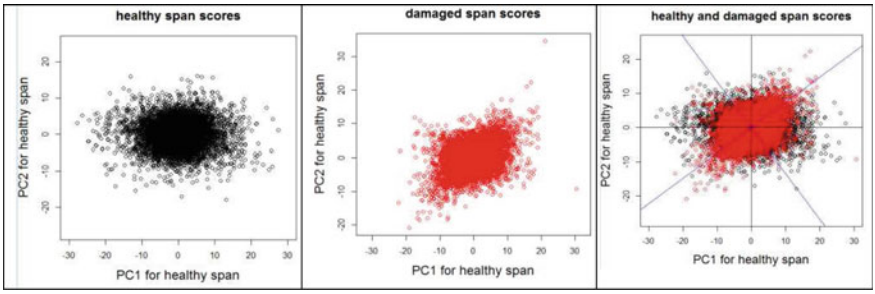


Fig. 3 Scores from healthy and damaged spans projected on healthy PC1-PC2 subspace

the data scores obtained for healthy and damaged spans within a smaller subspace spanned by PC1-PC2.

Figure 3 shows the resulting graphs for healthy and damaged data, respectively, on the left and in the center. These scores are both represented on the same diagram on the right, where the angle $\theta \cong 42°$ is well highlighted.

4 Sampling Distribution and Hypothesis Testing of Cos(θ_k)

In order to perform inference about the parameter $\cos(\theta_k)$ the moving blocks bootstrap method [4] was adopted for bootstrapping time series, in successive steps:

Step1. Shuffle the columns of $\mathbf{X_{n,p}}$ for group of n = 30,000 rows (5 min) (overlapping is admitted) to obtain $\mathbf{sX_{n,p}}$ and do the same procedure on $\mathbf{Y_{n,p}}$ to obtain $\mathbf{sY_{n,p}}$.

Step2. Detect the structural subspaces for each simulated matrix $\mathbf{sX_{n,p}}$ and $\mathbf{sY_{n,p}}$ by means of PCA, and quote the among of variance remaining over a threshold.

Step3. Calculate the $\cos(\theta_k)$ between the subspaces **sH** and **H** and between the subspaces **sD** and **D** obtaining $\text{hcos}(\theta_k)$ and $\text{dcos}(\theta_k)$ respectively.

Repeat this sequence for N = 5000 times beginning with Step.1.

The sampling distribution of $\text{hcos}(\theta_k)$ and $\text{dcos}(\theta_k)$ are in Fig. 7, on the left and right respectively.

The inferential problem can be expressed as:

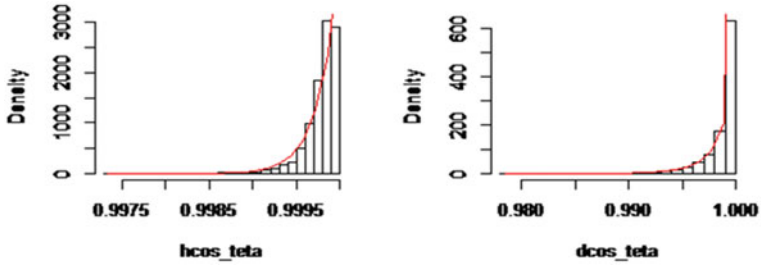


Fig. 7 Distributions of maxcos(θ) of (sH vs H) and (sD vs D)

Is the observed value $\cos(\theta) = 0.7429$ consistent with the hypothesis that $\cos(\theta_k) = 1$?

Or more formally:

Hypothesis testing H_0: $\cos(\theta_k) = 1$ **vs** H_1: $\cos(\theta_k) < 1$.

According to the simulation results, the observed value $\cos(\theta_k) = 0{,}7429$ isn't a plausible value under the hypothesis $H0_0$as the p-value $< 10^{-5}$ for both simulated sampling distributions.

The same simulation plane, repeated considering in step 1 n = 6000 rows (1 min) instead of n = 30,000, gave similar results related to the hypothesis testing.

5 Conclusion

The paper has presented an investigation of the damage detection capability of the Principal Component Analysis applied to the response structural vibration tests [8] in time domain in the study of statistical data for the spans of a particular bridge. Whereas the analysis of the vibration tests using a Peaking technique did not produce easily interpretable evidence, the analysis of the subspaces resulting from the reduction by means of PCA gave evidence of the damage present in the structure.

Acknowledgements The author thanks Frank Lad for the useful suggestions that have improved the writing and the linguistic quality of the article

References

1. Anderson, T.W.: An Introduction to Multivariate Statistical Analysis. Wiley (2003)
2. Golub, G., Van loan C.: Matrix Computations. The Johns Hopkins University Press, Baltimore (1996)
3. Nguyen, V.H. et al.: Damage detection on the champangshiehl bridge using blind source separation. In: Proceedings of the Third International Symposium on Life-Cycle Civil Engineering, IALCCE'12, pp. 172–176 (2012)

4. Efron, B., Tibshirani, R.J.: An introduction to the Bootstrap Chapmann and Hall (1994)
5. Technical Report n 51216–2, DISMAT Laboratorio per la sperimentazione sulle strutture e sui materiali da costruzione (2017). https://dismat.it/
6. Brincker, R.,Ventura, C.: Introduction to Operational Modal Analysis. Wiley (2015)
7. Poncelet, F., Kerschen, G., Golinval, J.C.: Experimental modal analysis using blind source separation techniques. In: International Conference on Noise and Vibration Engineering, Leuven (2006)
8. Zhang, G., et al.: Principal component analysis method with space and time windows for damage detection. Sensors (Basel) **19**, 2521 (2019)

Using Differential Geometry for Sparse High-Dimensional Risk Regression Models

Luigi Augugliaro, Ernst C. Wit, Hassan Pazira, Javier González, Fentaw Abegaz, and Angelo M. Mineo

Abstract With the introduction of high-throughput technologies in clinical and epidemiological studies, the need for inferential tools that are able to deal with fat data-structures, i.e., relatively small number of observations compared to the number of features, is becoming more prominent. In this paper we propose an extension of the dgLARS method to high-dimensional risk regression models. The main idea of the proposed method is to use the differential geometric structure of the partial likelihood function in order to select the optimal subset of covariates.

Keywords dgLARS · High-dimensional data · Risk regression model · Sparsity · Survival analysis

1 Introduction

Advances in genomic technologies have meant that many new clinical studies in cancer survival include a variety of genomic measurements, ranging from gene expression to SNP data. Studying the relationship between survival and genomic markers can be useful for a variety of reasons. If a genomic signature can be found, then patients can receave more accurate survival information. Furthermore, treatment and care may be adjusted to the prospects of an individual patient. Eventually, the genomic signature combined with information from other studies may be used to identify drug targets.

L. Augugliaro (✉) · A. M. Mineo
Department of Economics, Business and Statistics, University of Palermo, Palermo, Italy
e-mail: luigi.augugliaro@unipa.it

E. C. Wit
Institute of Computational Science, USI, Via Buffi 13, 6900 Lugano, Switzerland

H. Pazira · F. Abegaz
Bernoulli Institute, University of Groningen, Groningen, The Netherlands

J. González
Amazon Research Cambridge, Cambridge, UK

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_2

In the study of the dependence of survival time on covariates, the Cox proportional hazards model [6] has proved to be a major tool in many clinical and epidemiological applications. However, when the number of features is large, the simple Cox proportional model breaks down. This has led, in the last two decades, to the development of a variety of penalized methods, that is methods based on the idea to penalize convex likelihoods by functions of the parameters inducing solutions with many zeros. The Lasso [36], elastic net [38], l_0 [30] and the SCAD [11] penalties are examples of such penalties that, depending on some tuning parameter, conveniently shrink estimates to exact zeros. Also in survival analysis these methods have been introduced. Tibshirani [37] applied the Lasso penalty to the Cox proportional hazards model. References [14, 15, 33] suggested important computational improvements to make the calculation of the Lasso estimator in the Cox proportional hazards model more efficient. Although the Lasso penalty induces sparsity, it is well known to suffer for possible inconsistent selection of variables.

Whereas penalized inference is convenient, justification of the penalty is somewhat problematic. Furthermore, the methods suffer from being not invariant under scale transformations of the explanatory variables, then most penalized regression methods start their exposition by assuming that the variables are appropriately renormalized.

In this paper, we will approach sparsity directly from a likelihood point of view. The angle between the covariates and the tangent residual vector within the likelihood manifold provides a direct and scale-invariant way to assess the importance of the individual covariates. The idea is similar to the least angle regression approach proposed by [10]. However, rather than using it as a computational device for obtaining Lasso solutions, we view the method in its own right as in [1]. Moreover, the method extends directly the Cox proportional hazard model.

The remaining part of this paper is structured as follows. In Sect. 2, we briefly review the Cox regression model and its extensions and in Sect. 3 we derive its differential geometric structure. In Sect. 4 we propose the extension of the differential geometric least angle regression (dgLARS) method [1]. In Sect. 5, by simulation studies, we compare the performance of the proposed method with other sparse survival regression approaches, especially in the presence of correlated predictors. In Sect. 6 we employ the proposed method to find out a genetic signature for cancer survival in skin, colon, prostate and ovarian cancer. Finally, in Sect. 7 we draw some conclusions.

2 Cox Regression Model and Its Extensions

In analyzing survival data, one of the most important tool is the hazard function, which is used to express the risk or hazard of failure at some time t. Formally, let T be the (absolutely) continuous random variable associated with the survival time and let $f(t)$ be the corresponding probability density function, the hazard function is defined as $\lambda(t) = f(t)/\{1 - \int_0^t f(s)\mathrm{d}s\}$; $\lambda(t)$ specifies the instantaneous rate at which failures

occur for subjects that are surviving at time t. Suppose that the hazard function can depend on a p-dimensional, possibly time-dependent, vector of covariates, denoted by $\mathbf{x}(t) = (x_1(t), \ldots, x_p(t))^\top$. Cox regression model [6] is based on the assumption that $\mathbf{x}(t)$ influences the hazard function through the following relation

$$\lambda(t; \mathbf{x}) = \lambda_0(t) \exp\{\boldsymbol{\beta}^\top \mathbf{x}(t)\}, \tag{1}$$

where $\boldsymbol{\beta} \in B \subseteq \mathbb{R}^p$ is a p-dimensional vector of unknown fixed parameters and $\lambda_0(t)$ is the base line hazard function at time t, which is left unspecified. Model (1) was extended in [35] introducing the general class of models called *relative risk regression models*, which is defined assuming that

$$\lambda(t; \mathbf{x}) = \lambda_0(t) \psi\{\mathbf{x}(t); \boldsymbol{\beta}\}, \tag{2}$$

where $\psi : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$ is a differentiable function, called the *relative risk function*, such that $\psi\{\mathbf{x}(t); \boldsymbol{\beta}\} > 0$ for each $\boldsymbol{\beta} \in B$. It is also assumed that $\psi(\mathbf{0}; \boldsymbol{\beta}) = 1$, that is the relative risk function is normalized. Model (2) allows us to work with applications in which the exponential form of the relative risk function is not the best choice. This issue was observed in [23] and further underlined in [8]. As a motiving example for the generalization (2), several authors have also noted that the linear relative risk function $\psi\{\mathbf{x}(t); \boldsymbol{\beta}\} = 1 + \boldsymbol{\beta}^\top \mathbf{x}(t)$ provides a natural framework to assess departures from an additive relative risk model when two or more risk factors are studied in relation to the incidence of a disease (see for example [27, 28, 35], among the others.

Suppose that n observations are available and let t_i be the ith observed failure time. Assume that we have k uncensored and untied failure times and let D be the set of indices for which the corresponding failure time is observed; the remaining failure times are right censored. As explained in [9], if we denote by $R(t)$ the risk set, i.e., the set of indices corresponding to the subjects who have not failed and are still under observation just prior to time t, under the assumption of independent censoring, inference about the $\boldsymbol{\beta}$ can be carried out by the partial likelihood function

$$L_p(\boldsymbol{\beta}) = \prod_{i \in D} \frac{\psi\{\mathbf{x}_i(t_i); \boldsymbol{\beta}\}}{\sum_{j \in R(t_i)} \psi\{\mathbf{x}_j(t_i); \boldsymbol{\beta}\}}. \tag{3}$$

When the exponential relative risk function is used in model (2) and we work with fixed covariates, (3) is clearly equal to the original partial likelihood introduced in [6] and discussed in great detail in [7].

3 Geometrical Structure of a Relative Risk Regression Model

In this section we extend the dgLARS method [1] to the relative risk regression models. The basic idea underlying this method is to use the differential geometrical structure of a Generalized Linear Model (GLM) [20] to generalize the LARS method [10]. This means that a critical step in our method is to relate the partial likelihood (3) with the likelihood function of a specific GLM. This can be done using the identity existing between the partial likelihood and the likelihood function of a logistic regression model for matched case-control studies [34].

The partial likelihood (3) can be seen as arising from a multinomial sample scheme. Consider an index $i \in D$ and let $\mathbf{Y}_i = (Y_{ih})_{h \in R(t_i)}$ be a multinomial random variable with sample size equal to 1 and cell probabilities $\boldsymbol{\pi}_i = (\pi_{ih})_{h \in R(t_i)} \in \Pi_i$, i.e., $p(\mathbf{y}; \boldsymbol{\pi}_i) = \prod_{h \in R(t_i)} \pi_{ih}^{y_{ih}}$. Assuming that $\mathbf{Y}_i$ are independent, the joint probability density function is an element of the set $S = \left\{ \prod_{i \in D} \prod_{h \in R(t_i)} \pi_{ih}^{y_{ih}} : (\boldsymbol{\pi}_i)_{i \in D} \in \bigotimes_{i \in D} \Pi_i \right\}$, called the ambient space. Letting

$$E_{\boldsymbol{\beta}}(Y_{ih}) = \pi_{ih}(\boldsymbol{\beta}) = \frac{\psi\{\mathbf{x}_h(t_i); \boldsymbol{\beta}\}}{\sum_{j \in R(t_i)} \psi\{\mathbf{x}_j(t_i); \boldsymbol{\beta}\}}, \tag{4}$$

and assuming that for each $i \in D$, the observed y_{ih} is equal to one if h is equal to i and zero otherwise; it can be shown that (3) is formally equivalent to the likelihood function associated to the model space $M = \left\{ \prod_{i \in D} \prod_{h \in R(t_i)} \{\pi_{ih}(\boldsymbol{\beta})\}^{y_{ih}} : \boldsymbol{\beta} \in B \right\}$. Let $\ell(\boldsymbol{\beta}) = \sum_{i \in D} \sum_{h \in R(t_i)} Y_{ih} \log \pi_{ih}(\boldsymbol{\beta})$ be the log-likelihood function associated to the model space M and let $\partial_m \ell(\boldsymbol{\beta}) = \partial \ell(\boldsymbol{\beta}) / \partial \beta_m$. The tangent space $T_{\boldsymbol{\beta}} M$ of M at the model point $\prod_{i \in D} \prod_{h \in R(t_i)} \{\pi_{ih}(\boldsymbol{\beta})\}^{y_{ih}}$ is defined as the linear vector space spanned by the p elements of the score vector, formally, $T_{\boldsymbol{\beta}} M = \text{span}\{\partial_1 \ell(\boldsymbol{\beta}), \ldots, \partial_p \ell(\boldsymbol{\beta})\}$. Under standard regularity conditions, it is easy to see that $T_{\boldsymbol{\beta}} M$ is the linear vector space of the random variables $v_{\boldsymbol{\beta}} = \sum_{m=1}^{p} v_m \partial_m \ell(\boldsymbol{\beta})$ with zero expectation and finite variance. Applying the chain rule, for any tangent vector belonging to $T_{\boldsymbol{\beta}} M$ we have

$$\begin{aligned} v_{\boldsymbol{\beta}} = \sum_{m=1}^{p} v_m \partial_m \ell(\boldsymbol{\beta}) &= \sum_{i \in D} \sum_{h \in R(t_i)} \left\{ \sum_{m=1}^{p} v_m \frac{\partial \pi_{ih}(\boldsymbol{\beta})}{\partial \beta_m} \right\} \partial_{ih} \ell(\boldsymbol{\beta}) = \\ &= \sum_{i \in D} \sum_{h \in R(t_i)} w_{ih} \partial_{ih} \ell(\boldsymbol{\beta}), \end{aligned}$$

where $\partial_{ih} \ell(\boldsymbol{\beta}) = \partial \ell(\boldsymbol{\beta}) / \partial \pi_{ih}$; the previous expression shows that $T_{\boldsymbol{\beta}} M$ is a linear sub vector space of the tangent space $T_{\boldsymbol{\beta}} S$ spanned by the random variables $\partial_{ih} \ell(\boldsymbol{\beta})$. To define the notion of angle between two given tangent vectors belonging to $T_{\boldsymbol{\beta}} M$, say $v_{\boldsymbol{\beta}} = \sum_{m=1}^{p} v_m \partial_m \ell(\boldsymbol{\beta})$ and $w_{\boldsymbol{\beta}} = \sum_{n=1}^{p} w_n \partial_n \ell(\boldsymbol{\beta})$, we shall use the information metric [29], i.e.,

$$\langle v_{\boldsymbol{\beta}}; w_{\boldsymbol{\beta}} \rangle_{\boldsymbol{\beta}} = E_{\boldsymbol{\beta}}(v_{\boldsymbol{\beta}} w_{\boldsymbol{\beta}}) = \sum_{m,n=1}^{p} E_{\boldsymbol{\beta}}\{\partial_m \ell(\boldsymbol{\beta}) \partial_n \ell(\boldsymbol{\beta})\} v_m w_n = \mathbf{v}^\top I(\boldsymbol{\beta}) \mathbf{w}, \quad (5)$$

where $\mathbf{v} = (v_1, \ldots, v_p)^\top$, $\mathbf{w} = (w_1, \ldots, w_p)^\top$ and $I(\boldsymbol{\beta})$ is the Fisher information matrix evaluated at $\boldsymbol{\beta}$. As observed in [21], the matrix $I(\boldsymbol{\beta})$ used in (5) is not exactly equal to the Fisher information matrix of the relative risk regression model, however it has the same asymptotic inferential properties. Finally, to complete our differential geometric framework we need to introduce the tangent residual vector $r_{\boldsymbol{\beta}} = \sum_{i \in D} \sum_{h \in R(t_i)} r_{ih}(\boldsymbol{\beta}) \partial_{ih} \ell(\boldsymbol{\beta})$, where $r_{ih}(\boldsymbol{\beta}) = y_{ih} - \pi_{ih}(\boldsymbol{\beta})$, which is an element of $T_{\boldsymbol{\beta}} S$ and which measures the difference between a model in M and the observed data. For more details see [1].

4 Extending dgLARS to Relative Risk Regression Models

dgLARS is a sequential method developed for constructing a sparse path of solutions indexed by a positive parameter γ and theoretically founded on the following characterization of the mth signed Rao score test statistic, i.e.:

$$r_m^u(\boldsymbol{\beta}) = I_{mm}^{-1/2}(\boldsymbol{\beta}) \partial_m \ell(\boldsymbol{\beta}) = \cos\{\rho_m(\boldsymbol{\beta})\} \|r_{\boldsymbol{\beta}}\|_{\boldsymbol{\beta}}, \quad (6)$$

where $\|r_{\boldsymbol{\beta}}\|_{\boldsymbol{\beta}}^2 = \sum_{i \in D} \sum_{h,k \in R(t_i)} E_{\boldsymbol{\beta}}\{\partial_{ih} \ell(\boldsymbol{\beta}) \partial_{ik} \ell(\boldsymbol{\beta})\} r_{ih}(\boldsymbol{\beta}) r_{ik}(\boldsymbol{\beta})$ and $I_{mm}(\boldsymbol{\beta})$ is the Fisher information for β_m. The quantity $\rho_m(\boldsymbol{\beta})$ is a generalization of the Euclidean notion of angle between the mth column of the design matrix and the residual vector $\mathbf{r}(\boldsymbol{\beta}) = (r_{ih}(\boldsymbol{\beta}))_{i \in D, h \in R(t_i)}$. Characterization (6) gives us a natural way to generalize the equiangularity condition (?: two predictors, say the mth and nth, satisfy the generalized equiangularity condition at the point $\boldsymbol{\beta}$ when $|r_m^u(\boldsymbol{\beta})| = |r_n^u(\boldsymbol{\beta})|$. Inside dgLARS theory, the generalized equiangularity condition is used to identify the predictors that are included in the model.

The nonzero estimates are formally defined as follows. For any data set there is a finite sequence of transition points, say $\gamma^{(1)} \geq \cdots \geq \gamma^{(K)} \geq 0$, such that for any fixed γ between $\gamma^{(k+1)}$ and $\gamma^{(k)}$ the sub vector of the nonzero dgLARS estimates, denoted as $\hat{\boldsymbol{\beta}}_{\hat{A}}(\gamma) = (\hat{\beta}_m(\gamma))_{m \in \hat{A}}$, satisfies the following conditions:

$$r_m^u\{\hat{\boldsymbol{\beta}}_{\hat{A}}(\gamma)\} = s_m \gamma, \quad m \in \hat{A} \quad (7)$$

$$|r_n^u\{\hat{\boldsymbol{\beta}}_{\hat{A}}(\gamma)\}| < \gamma, \quad n \notin \hat{A} \quad (8)$$

where $s_m = \text{sign}\{\hat{\beta}_m(\gamma)\}$ and $\hat{A} = \{m : \hat{\beta}_m(\gamma) \neq 0\}$, called active set, is the set of the indices of the predictors that are included in the current model, called active predictors. In any transition point, say for example $\gamma^{(k)}$, one of the following two conditions occurs:

1. there is a non active predictor, say the nth, satisfying the generalized equiangularity condition with any active predictor, i.e.,

$$|r_n^u\{\hat{\boldsymbol{\beta}}_{\hat{A}}(\gamma^{(k)})\}| = |r_m^u\{\hat{\boldsymbol{\beta}}_{\hat{A}}(\gamma^{(k)})\}| = \gamma^{(k)}, \tag{9}$$

   for any m in $\hat{A}$, then it is included in the active set;
2. there is an active predictor, say the mth, such that

$$\text{sign}[r_m^u\{\hat{\boldsymbol{\beta}}_{\hat{A}}(\gamma^{(k)})\}] \neq \text{sign}\{\hat{\beta}_m(\gamma^{(k)})\}, \tag{10}$$

   then it is removed from the active set.

Given the previous definition, the path of solutions can be constructed in the following way. Since we are working with a class of regression models without intercept term, the starting point of the dgLARS curve is the zero vector: this means that, at the starting point, the p predictors are ranked using $|r_m^u(\mathbf{0})|$. Suppose that $a_1 = \arg\max_m |r_m^u(\mathbf{0})|$, then $\hat{A} = \{a_1\}$, $\gamma^{(1)}$ is set equal to $|r_{a_1}^u(\mathbf{0})|$ and the first segment of the dgLARS curve is implicitly defined by the nonlinear equation $r_{a_1}^u\{\hat{\beta}_{a_1}(\gamma)\} - s_{a_1}\gamma = 0$. The proposed method traces the first segment of the dgLARS curve reducing γ until we find the transition point $\gamma^{(2)}$ corresponding to the inclusion of a new index in the active set, in other words, there exists a predictor, say the a_2th, satisfying condition (9), then a_2 is included in $\hat{A}$ and the new segment of the dgLARS curve is implicitly defined by the system with nonlinear equations:

$$r_{a_i}^u\{\hat{\boldsymbol{\beta}}_{\hat{A}}(\gamma)\} - s_{a_i}\gamma = 0, \quad a_i \in \hat{A},$$

where $\hat{\boldsymbol{\beta}}_{\hat{A}}(\gamma) = (\hat{\beta}_{a_1}(\gamma), \hat{\beta}_{a_2}(\gamma))^\top$. The second segment is computed reducing γ and solving the previous system until we find the transition point $\gamma^{(3)}$. At this point, if condition (9) occurs a new index is included in $\hat{A}$, otherwise condition (10) occurs and an index is removed from $\hat{A}$. In the first case the previous system is updated adding a new nonlinear equation while, in the second case, a nonlinear equation is removed. The curve is traced as previously described until parameter γ is equal to some fixed value that can be zero, if the sample size is large enough, or some positive value if we are working in a high-dimensional setting, i.e., the number of predictors is larger than the sample size. Table 1 reports the pseudo-code of the developed algorithm to compute the dgLARS curve for a relative regression model. From a computational point of view, the entire dgLARS curve can be computed using the predictor-corrector algorithm proposed in [1]; for more details about this algorithm the interested reader is referred to [2, 3] or [25]. The latter extends the dgLARS method to GLM based on the exponential dispersion family proposing an improved predictor-corrector algorithm.

Table 1 Pseudocode of the dgLARS algorithm for a relative risk regression model

Step	Description
0.	Let $r_m^u(\boldsymbol{\beta})$ be the Rao score statistic associated with the partial likelihood
1.	Let $\gamma^{(1)} = \max_m \lvert r_m^u(\mathbf{0})\rvert$ and initialize the active set $\hat{A} = \arg\max_m \lvert r_m^u(\mathbf{0})\rvert$
2.	Repeat the following steps
3.	Trace the segment of the dgLARS curve reducing γ and solving the system
	$r_m^u\{\hat{\boldsymbol{\beta}}_{\hat{A}}(\gamma)\} - s_m\gamma = 0, \quad m \in \hat{A}$
4.	Until γ is equal to the next transition point
5.	If condition (9) is met then include the new index in $\hat{A}$
6.	Else (condition (10) is met) remove the index from $\hat{A}$
7.	Until γ reaches some small positive value

5 Simulation Study

In this section we compare the proposed method with three popular competitors: the coordinate descent method developed by [32], named CoxNet, the predictor-corrector algorithm developed by [24], named CoxPath, and the gradient ascent algorithm proposed by [14], named CoxPen. These algorithms are implemented in the `R` packages `glmnet`, `glmpath` and `penalized`, respectively. Considering that these methods have been implemented only for Cox regression model, our comparison will focus on this kind of relative risk regression model. In the following of this section, dgLARS method applied to the Cox regression model is referred to as the dgCox model.

We simulated one hundred datasets from a Cox regression model, where the survival times t_i $(i = 1, \ldots, n)$ follow an exponential distributions with parameter $\lambda_i = \exp(\boldsymbol{\beta}^\top \mathbf{x}_i)$, and $\mathbf{x}_i$ is sampled from a p-variate normal distribution $\mathrm{N}(\mathbf{0}, \Sigma)$; the entries of Σ are fixed to $\mathrm{corr}(X_m, X_n) = \rho^{|m-n|}$ with $\rho \in \{0.3, 0.9\}$. The censorship is randomly assigned to the survival times with probability $\pi \in \{0.2, 0.4\}$. The number of predictors is equal to 100 and the sample size is equal to 50 and 150. The first value is used to evaluate the behaviour of the methods in high-dimensional setting. Finally, we set $\beta_m = 0.2$ for $m = 1, \ldots, s$, where $s \in \{5, 10\}$; the remaining parameters are equal to zero.

To evaluate the global behaviour of the considered methods, we used the following procedure. In any simulation run, we compute the path associated to dgCox, CoxNet, CoxPath and CoxPen methods, respectively. For each point of a given path, denoted as $\hat{\boldsymbol{\beta}}(\gamma)$, we compute the False and True Positive Rate, defined as

$$\mathrm{FPR}(\gamma) = \frac{\text{number of false predictors selected by } \hat{\boldsymbol{\beta}}(\gamma)}{\text{total number of false predictors}}$$

and

$$\mathrm{TPR}(\gamma) = \frac{\text{number of true predictors selected by } \hat{\boldsymbol{\beta}}(\gamma)}{\text{total number of true predictors}}$$

These quantities are used to compute the Receiver Operating Characteristic (ROC) curve and the corresponding Area Under the Curve (AUC).

Figures 1, 2, 3 and 4 show the results for our simulation study. In scenarios where $\rho = 0.3$, CoxNet, CoxPath and CoxPen exhibit a similar performance, having overlapping curves for both levels of censoring, whereas dgCox method appears to be consistently better with the largest AUC. A similar performance of the methods

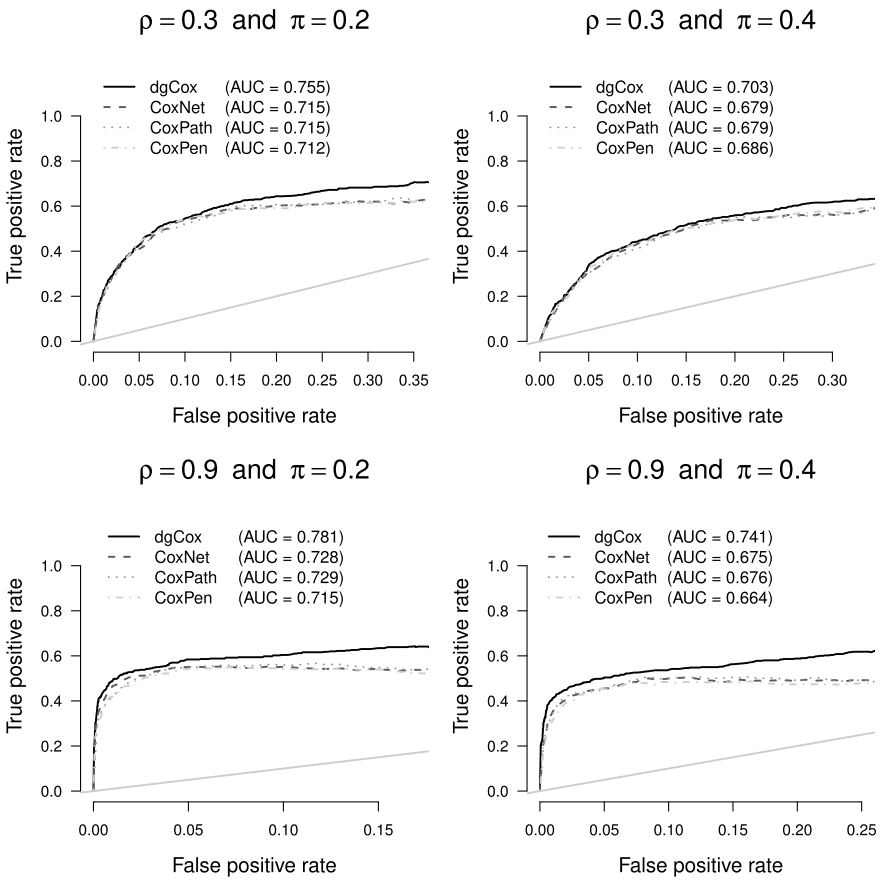


Fig. 1 Results from the simulation study with $s = 5$ and sample size equal to 50; for each scenario we show the averaged ROC curve for dgCox, CoxNet, CoxPath and CoxPen algorithm. The average Area Under the Curve (AUC) is also reported. The 45-degree diagonal is also included in the plots

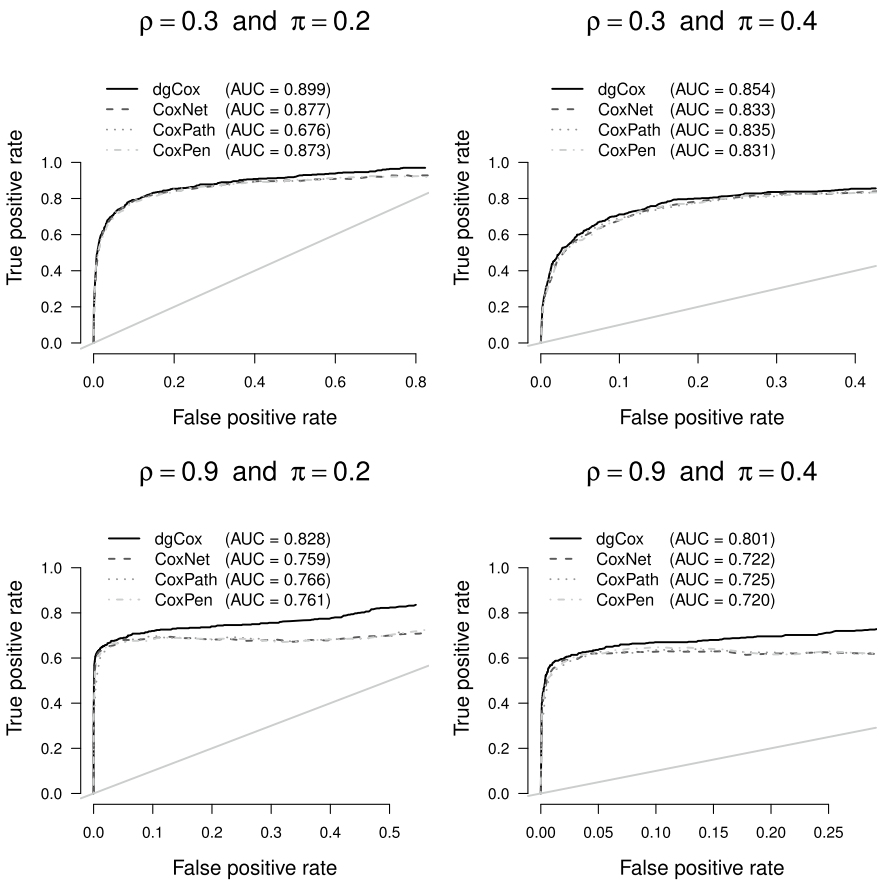


Fig. 2 Results from the simulation study with $s = 5$ and sample size equal to 150; for each scenario we show the averaged ROC curve for dgCox, CoxNet, CoxPath and CoxPen algorithm. The average Area Under the Curve (AUC) is also reported. The 45-degree diagonal is also included in the plots

has been also observed for the other combinations of the ρ and π values. In scenarios where the correlation among neighbouring predictors is high, i.e., $\rho = 0.9$, the dgCox method is clearly the superior approach for all levels of censoring. Similar results are obtained when we increase the sample size.

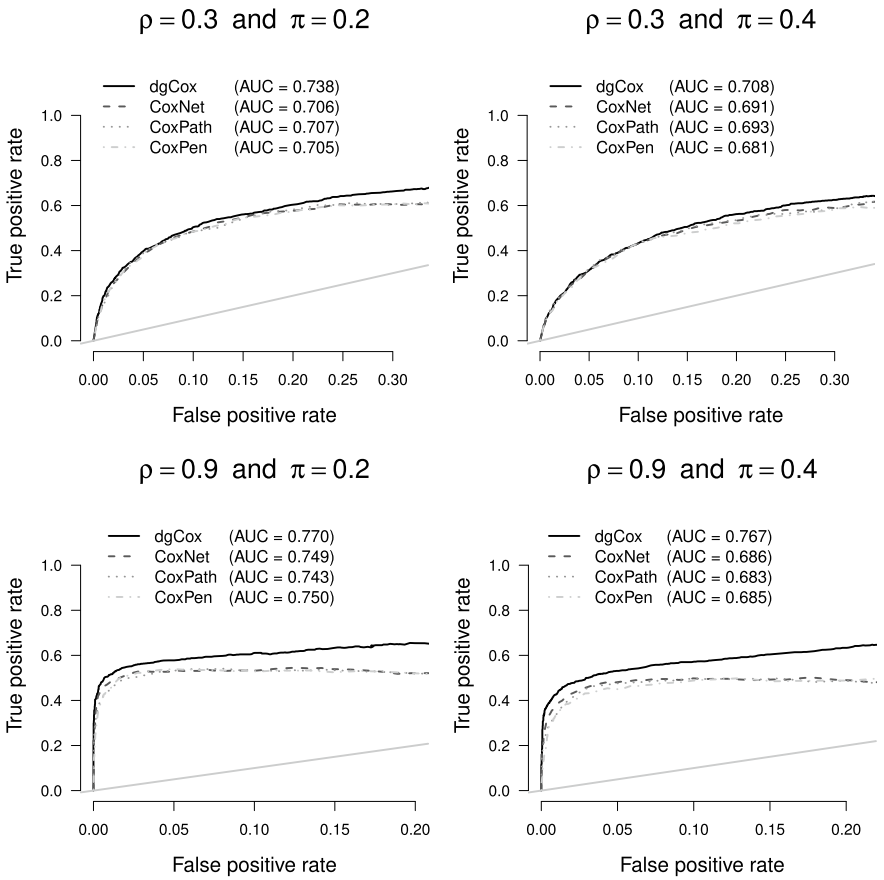


Fig. 3 Results from the simulation study with $s = 10$ and sample size equal to 50; for each scenario we show the averaged ROC curve for dgCox, CoxNet, CoxPath and CoxPen algorithm. The average Area Under the Curve (AUC) is also reported. The 45-degree diagonal is also included in the plots

6 Application to Survival Datasets

In this section we test the predictive power of the proposed method in four recent studies. In particular, we focus on the identification of genes involved in the regulation of colon cancer [19], prostate cancer [31], ovarian cancer [13] and skin cancer [17]. The set-up of the four studies was similar. In the patient a cancer was detected and treated. When treatment was complete a follow-up started. In all cases, the expression of several genes were measured in the affected tissue together with the survival times of the patients, which may be censored if the patients were alive when they left the study.

Table 2 containing a brief description of the datasets used in this section together with some results. In the four scenarios the number of predictors p is larger than the

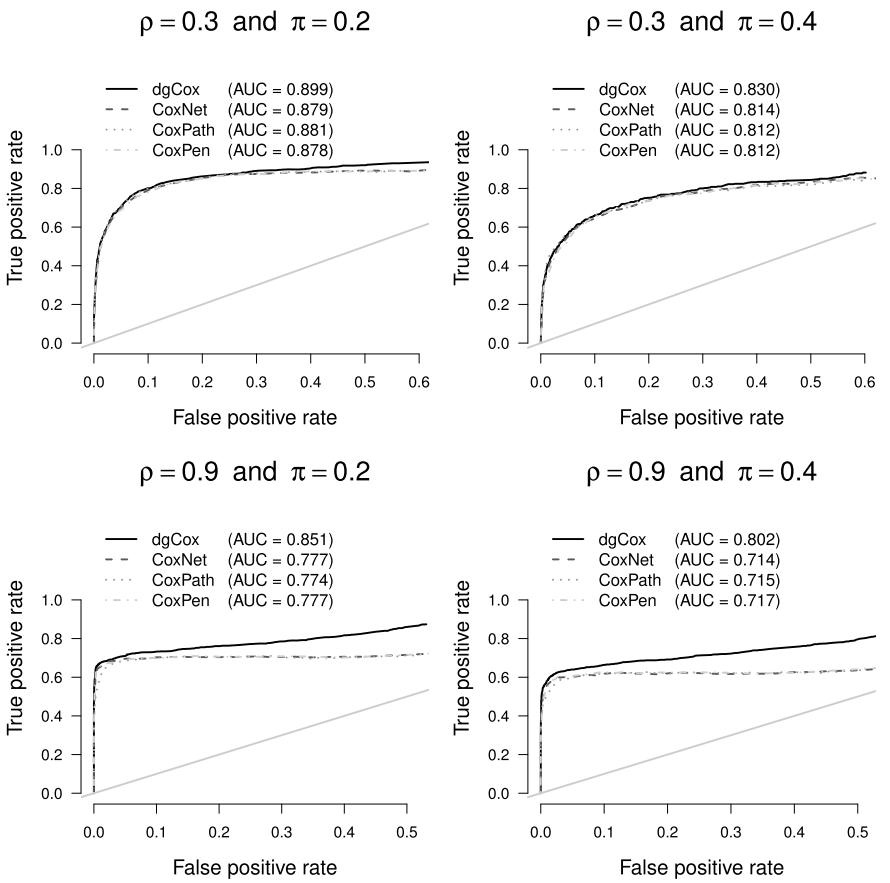


Fig. 4 Results from the simulation study with $s = 10$ and sample size equal to 150; for each scenario we show the averaged ROC curve for dgCox, CoxNet, CoxPath and CoxPen algorithm. The average Area Under the Curve (AUC) is also reported. The 45-degree diagonal is also included in the plots

Table 2 Description of the four cancer datasets

Dataset	n	n. uncensored	p	n. genes selected	p-value	quartiles of $\hat{\boldsymbol{\beta}}_{\hat{A}}(\gamma)$ 1st and 3rd
Colon	125	70	23698	38	0.023	$\{-0.654; 0.922\}$
Prostate	61	24	162	24	0.033	$\{-0.665; 0.629\}$
Ovarian	103	57	306	43	0.004	$\{-0.332; 0.286\}$
Skin	54	47	30807	23	0.025	$\{0.000; 0.001\}$

number of patients n. The dimensionality is especially high in the cases of the colon and skin cancer where the expression of several thousands of genes were measured. In the prostate and ovarian cancer the number of genes is 162 and 306, which will also help us to study the performance of the dgLARS method when the number of variables is just a few orders of magnitude larger than the number of observations.

Although the number of predictors exceeds the sample size, in genomics it is common to assume that just a moderate number of genes affects the phenotype of interest. To identify such genes in this survival context, we use dgCox model. To this end, we randomly select a training sample containing the 60% of the patients and we save the remaining data to test the models. We calculate the four paths of solutions, that is a path for each dataset, and we select the optimal number of components by

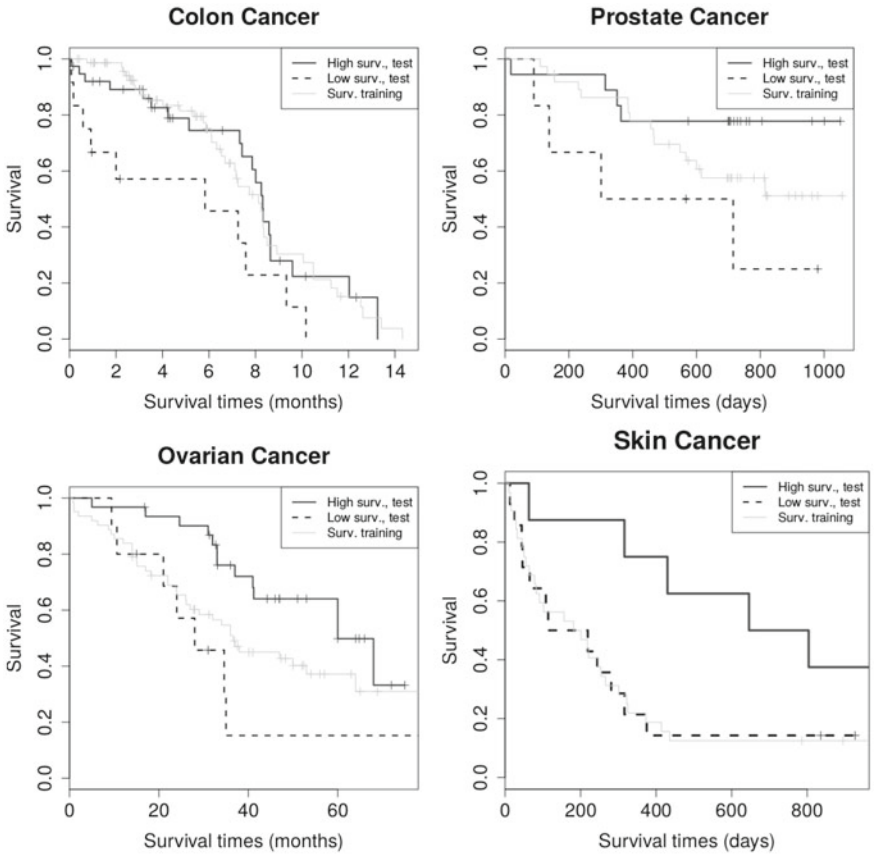


Fig. 5 The Kaplan-Meier survival curves estimates for training data are shown together with the curves associated to the two groups obtained in the test sample by means of the estimated predicted excess risk. In the four cases, the two groups in the test sample show a significant separation according to the Peto & Peto modification of the Gehan-Wilcoxon test

means of the BIC criterion. For the colon, prostate, ovarian and skin cancer studies we find gene profiles consisting of, respectively, 38, 24, 43 and 23 genes.

In order to illustrate the prediction performance of our method we classify the test patients into a low-risk group and a high-risk groups by splitting the test sample into two subsets of equal size according to the estimated individual predicted excess risk. Figure 5 shows the Kaplan-Maier survival curves estimates for the two groups together with the original training survival curve. The differences are significant in the four cases showing the predictive power of the survival function provided by the selected genes. To test the groups separation we use the non-parametric Peto & Peto modification of the Gehan-Wilcoxon test [26]. For all four studies the differences between the low and high-risk groups are significant at the traditional 0.05 significance level.

7 Conclusions

In this paper, we have proposed an extension of the differential geometric least angle regression method to relative risk regression model using the relationship existing between the partial likelihood function and a specific generalized linear model. The advantage of this approach is that the estimates are invariant to arbitrary changes in the measurement scale of the predictors. Unlike SCAD or ℓ_1 sparse regression methods, no prior rescaling of the predictors is therefore needed. The proposed method can be used for a large class of survival models, the so called relative risk models. We have implementations for the Cox proportional hazards model and the excess relative risk model.

By an extensive simulation study, we have shown that proposed method outperforms the considered competitors both in high-dimensional and classical setting. Results show that dgCox model is able to find out the true predictors in more challenging settings, that is when the response variables are highly correlated.

Finally, our method was used to study four recent cancer survival datasets, where we look for a genetic "survival signature". Due to the large number of predictors, the studies are unsuitable for traditional survival regression methods.

In order to illustrate the prediction performance of our method we classified the test patients into a low-risk group and a high-risk groups by splitting the test sample into two subsets of equal size according to the estimated individual predicted excess risk. The predictive power of the survival function provided by the selected genes is clearly shown in the estiamated Kaplan-Maier survival curves, Finally, to test the groups separation we used the non-parametric Peto & Peto modification of the Gehan-Wilcoxon test. For all four studies the differences between the low and high-risk groups were significant at the traditional 0.05 significance level demonstrating that dgCox model is an effective tool in medical analysis for massive gene screening studies.

References

1. Augugliaro, L., Mineo, A.M., Wit, E.C.: Differential geometric least angle regression: a differential geometric approach to sparse generalized linear models. J. R. Stat. Soc. Ser. B **75**(3), 471–498 (2013)
2. Augugliaro, L., Mineo, A.M., Wit, E.C.: dglars: an R package to estimate sparse generalized linear models. J. Stat. Softw. **59**(8), 1–40 (2014)
3. Augugliaro, L., Mineo, A.M., Wit, E.C.: A differential geometric approach to generalized linear models with grouped predictors. Biometrika **103**(3), 563–577 (2016)
4. Bao, L., Kimzey, A., Sauter, G., Sowadski, J.M., Lu, K.P., Wang, D.G.: Prevalent overexpression of prolyl isomerase Pin1 in human cancers. Am. J. Pathol. **164**(5), 1727–1737 (2004)
5. Boldrini, L., Pistolesi, S., Gisfredi, S., Ursino, S., Ali, G., Pieracci, N., Basolo, F., Parenti, G., Fontanini, G.: Telomerase activity and hTERT mRNA expression in glial tumors. Int. J. Oncol. **28**(6), 1555–1560 (2006)
6. Cox, D.R.: Regression models and life-tables. J. R. Stat. Soc. Ser. B **34**(2), 187–220 (1972)
7. Cox, D.R.: Partial likelihood. Biometrika **62**(2), 269–276 (1975)
8. Cox DR (1981) Discussion of paper by D. Oakse entitled "survival times: aspects of partial likelihood". Int. Stat. Rev. **49**(3), 258
9. Cox, D.R., Oakes, D.: Analysis of Survival Data. Monographs on Statistics and Applied Probability. Chapman and Hall, London (1984)
10. Efron, B., Hastie, T., Johnstone, I., Tibshirani, R.: Least angle regression. Ann. Stat. **32**(2), 407–499 (2004)
11. Fan, J., Li, R.: Variable selection via nonconcave penalized likelihood and its oracle properties. J. Am. Stat. Assoc. **96**(456), 1348–1360 (2001)
12. Fan, Y., Tang, C.Y.: Tuning parameter selection in high dimensional penalized likelihood. J R. Stat. Soc.: Ser. B **75**(3), 531–552 (2013)
13. Gillet, J.P., Calcagno, A.M., Varma, S., Davidson, B., Elstrand, M.B., Ganapathi, R., Kamat, A.A., Sood, A.K., Ambudkar, S.V., Seiden, M.V., Rueda, B.R., Gottesman, M.M.: Multidrug resistance-linked gene signature predicts overall survival of patients with primary ovarian serous carcinoma. Clin. Cancer Res. **18**(11), 3197–3206 (2012)
14. Goeman, J.J.: L1 penalized estimation in the Cox proportional hazards model. Biometr. J. **52**(1), 70–84 (2010)
15. Gui, J., Li, H.: Penalized Cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data. Bioinformatics **21**(13), 3001–3008 (2005)
16. Heagerty, P.J., Lumley, T., Pepe, M.S.: Time-dependent roc curves for censored survival data and a diagnostic marker. Biometrics **56**(2), 337–344 (2000)
17. Jönsson, G., Busch, C., Knappskog, S., Geisler, J., Miletic, H., Ringnér, Lillehaug JR., Borg, A., Lønning, P.E.: Gene expression profiling-based identification of molecular subtypes in stage IV melanomas with different clinical outcome. Clin. Cancer Res. **16**(13), 3356–67 (2010)
18. Konishi, S., Kitagawa, G.: Generalised information criteria in model selection. Biometrika **83**(4), 875–890 (1996)
19. Loboda, A., Nebozhyn, M.V., Watters, J.W., Buser, C.A., Shaw, P.M., Huang, P.S., Van't Veer, L.R.A.T., Jackson, D.B, Agrawal, D., Dai, H., Yeatman, T.J.: EMT is the dominant program in human colon cancer. BMC Medical Genomics, pp. 4–9 (2011)
20. McCullagh, P., Nelder, J.A.: Generalized Linear Models, 2nd edn. Chapman & Hall, London (1989)
21. Moolgavkar, S.H., Venzon, D.J.: Confidence regions in curved exponential families: application to matched case-control and survival studies with general relative risk function. Ann. Stat. **15**(1), 346–359 (1987)
22. Nagel, G., Bjørge, T., Stocks, T., Manjer, J., Hallmans, G., Edlinger, M., Häggström, C., Engeland, A., Johansen, D., Kleiner, A., Selmer, R., Ulmer, H., Tretli, S., Jonsson, H., Concin, H., Stattin, P., Lukanova, A.: Metabolic risk factors and skin cancer in the metabolic syndrome and cancer project (Me-Can). Brit. J. Dermatol. **167**(1), 59–67 (2012)

23. Oakes, D.: Survival times: aspects of partial likelihood. Int. Stat. Rev. **49**(3), 235–252 (1981)
24. Park, M.Y., Hastie, T.: L1-regularization path algorithm for generalized linear models. J. R. Stat. Soc.: Ser. B **69**(4), 659–677 (2007)
25. Pazira, H., Augugliaro, L., Wit, E.C.: Extended differential geometric lars for high-dimensional glms with general dispersion parameter. Stat. Comput. **28**(4), 753–774 (2018)
26. Peto, R., Peto, J.: Asymptotically efficient rank invariant test procedures. J. R. Stat. Soc. Ser. A **135**(2), 185–207 (1972)
27. Prentice, R.L., Mason, M.W.: On the application of linear relative risk regression models. Biometrics **42**(1), 109–120 (1996)
28. Prentice, R.L., Yoshimoto, Y., Mason, M.: Relationship of cigarette smoking and radiation exposure to cancer mortality in Hiroshima and Nagasaki. J. Nat. Cancer Inst. **70**(4), 611–622 (1983)
29. Rao, C.R.: On the distance between two populations. Sankhyā **9**, 246–248 (1949)
30. Rippe, R.C.A., Meulman, J.J., Eilers, P.H.C.: Visualization of genomic changes by segmented smoothing using an L_0 penalty. PLoS One **7**(6), e38230 (2012)
31. Ross, R.W., Galsky, M.D., Scher, H.I., Magidson, J., Wassmann, K., Lee, G.S.M., Katz, L., Subudhi, S.K., Anand, A., Fleisher, M., Kantoff, P.W., Oh, W.K.: A whole-blood RNA transcript-based prognostic model in men with castration-resistant prostate cancer: a prospective study. Lancet Oncol **13**(11), 1105–13 (2012)
32. Simon, N., Friedman, J.H., Hastie, T., Tibshirani, R.: Regularization paths for Cox's proportional hazards model via coordinate descent. J. Stat. Softw. **39**(5), 1–13 (2011)
33. Sohn, I., Kim, J., Jung, S.H., Park, C.: Gradient lasso for Cox proportional hazards model. Bioinformatics **25**(14), 1775–1781 (2009)
34. Thomas, D.C.: Addendum to the paper by Liddell, McDonald, Thomas and Cunliffe. J. R. Stat. Soc. Ser. A **140**(4), 483–485 (1977)
35. Thomas, D.C.: General relative-risk models for survival time and matched case-control analysis. Biometrics **37**(4), 673–686 (1981)
36. Tibshirani, R.: Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B **58**(1), 267–288 (1996)
37. Tibshirani, R.: The lasso method for variable selection in the Cox model. Stat. Med. **16**, 385–395 (1997)
38. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B **67**(2), 301–320 (2005)

Perceived Benefits and Individual Characteristics of Internationally Mobile Students: A Discrete Latent Variable Analysis

Silvia Bacci, Valeria Caviezel, and Anna Maria Falzoni

Abstract In this paper, the potential benefits of studying abroad during higher education are analysed on the basis of an online survey administered to about 1600 students of a medium-sized university located in the North of Italy (the University of Bergamo) who spent a credit mobility experience abroad during the academic years from 2008/09 to 2014/15. Two dimensions are specifically investigated: the impact of the international experience on the students' skills, and the fulfilment of expectations concerning the international experience. A two-dimensional latent class Item Response Theory model under a concomitant variable approach is used to assess the item responses. The results show that international mobility positively affects students' perception of personality development and improves their language and soft skills. Mobile students report gains in terms of personal growth and enhanced employability at home and abroad. However, individual characteristics influence the latent class membership probability.

Keywords Concomitant variables approach · Erasmus programme · Item response theory · Latent class model

1 Introduction

The internationalization of higher education has become a priority in European education policy. According to the strategic objectives of Europe 2020, "an EU average of at least 20% of higher education graduates should have had a period of higher

S. Bacci
Department of Statistics, Computer Science, Applications G. Parenti, University of Florence, Viale Morgagni 59, 50134 Florence, Italy
e-mail: silvia.bacci@unifi.it

V. Caviezel (✉) · A. M. Falzoni
Department of Economics, University of Bergamo, Via dei Caniana 2, 24127 Bergamo, Italy
e-mail: valeria.caviezel@unibg.it

A. M. Falzoni
e-mail: anna-maria.falzoni@unibg.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_3

education-related study or training abroad, representing a minimum of 15 ECTS credits or lasting a minimum of three months" (EU Council of Ministers of Education, November 29, 2011). Since it began in 1987/1988, the world's most successful student mobility programme, the Erasmus programme, has provided over three million European students with the opportunity to go abroad and study at a higher education institution or train in a company [10].

Research in this area shows that participation in study abroad is generally highly beneficial to students in terms of personal development and academic and employment career. The studies on the matter may be divided into two broad categories. On the one hand, some studies have investigated the motivations for and potential benefits of international mobility by means of surveys conducted on target groups of students; on the other hand, other studies have tried to test empirically the causal effect of studying abroad on different outcomes, distinguishing between mobile and non-mobile students. Belonging to the first group of studies are the reports funded by the European Commission, The Erasmus Impact Study [5–7], based on large-scale surveys which examine the impact of Erasmus+ programme on students' individual skills enhancement, employability, and sense of shared European identity. A similar approach is followed by the Eurostudent surveys providing a broad cross country comparison of data on the social dimension in European higher education. The most recent survey covers the years 2016–2018 and illustrates various characteristics of internationally mobile students [9].

The second group of studies, in which the selection bias is taken into account to measure the impact of studying abroad, is mainly focused on employability and earnings outcomes. The results show that mobility generally enhances employability. However, a large variation emerges in the impact across different countries (see, among others, [8, 14, 16, 17, 19]).

Our contribution belongs to the first group of studies because it is based on an online survey conducted on about 1600 students enrolled at the University of Bergamo—a medium-sized university located in the North of Italy—who spent a credit mobility experience abroad during the academic years from 2008/09 to 2014/15. Administrative information at enrolment (e.g., age, gender, field of study, etc.) is also available for each student. The aim of the survey is to analyse students' motivations, the perceived impact of the international experience, and the fulfilment of expectations in its regard. Nevertheless, our methodological approach differs from the one adopted by the first group of studies. In fact, the international study experience may be framed as a not directly observable construct composed of multiple dimensions. In what follows we focus on two main aspects related to the international experience: its impact on the student's skills, and the fulfilment of expectations about this experience. Both these elements are latent variables which are measured through a set of polytomously-scored items. Our aim is to measure these latent variables, providing evidence of one or more problematic items. Moreover, we intend to detect individual characteristics that significantly explain the level of latent variables. For these purposes, we formulate a bidimensional Latent Class Item Response Theory (LC-IRT) model [2, 4] with a concomitant variable approach [11].

The rest of the contribution is organised as follows. Section 2 describes the questionnaire and data, whereas Sect. 3 provides details about the statistical model. The main results are illustrated and discussed in Sect. 4. The contribution concludes with some final remarks.

2 Data on International Study Mobility

To evaluate the impact of the international experience on students' skills and the fulfilment of their expectations, we conducted an online survey involving a sample of about 1600 students enrolled at the University of Bergamo. Our sample consists of all the credit mobility experiences recorded by the students during the seven academic years from 2008/2009 to 2014/2015.

In the first part of this section we present the questionnaire prepared for the survey, and then we describe the characteristics of the respondents and the percentages for each response category of items.

2.1 The Questionnaire

Our questionnaire consists of three sections: Decision to study abroad (first section), International experience (second section) and Coming back to Bergamo (third section). Before these three parts, we ask some questions regarding the students' personal details: parents' level of education, parents' employment status, the current employment status of the respondent, the academic year and the number of semesters spent abroad, the type of internationalization programme (Erasmus or Extra EU). In addition, we ask if the respondents and/or their family's members had an experience abroad in the past.

In the first section—Decision to study abroad—we ask the students to explain their decision to study abroad, i.e. to enhance future employability, to improve their curriculum studiorum (CV) to live a new experience with new challenges, to improve foreign language skills and knowledge about the host country. Furthermore, we analyse the factors that direct the choice towards a given host country and university: i.e. alignment of study programmes and availability of scholarships, prestige of host city and reputation of the university, knowledge of the language and culture of the host country, living costs.

In the second section—International experience—we ask the students to compare the linguistic skills, the teaching and assessment methods, the average exam evaluation between the experience abroad and the previous experience in Bergamo. Moreover, we investigate the main funding sources before and during the experience abroad.

Finally, in the third section—Coming back to Bergamo—the respondents are asked to evaluate the impact of the experience abroad on their communication skills, learning ability, foreign language skills, team working skills, adaptability, problem

solving skills and motivation to study. We ask also about the fulfilment of expectations in relation to enhance the future employability in Italy and abroad, personality development, foreign language skills, and ability to interact with foreign people.

In this contribution we consider the twelve items of the third section to measure two latent traits: the impact of the international experience on the students' skills and the fulfilment of expectations concerning the international experience; other items are covariates.

2.2 Data Description

The survey covers all the 1,576 students who have spent one/two semesters abroad for an Erasmus or Extra EU programme from a.y. 2008/09 to 2014/15; the response rate was 48.6% (766 students).

The sample (30.6% males and 69.4% females) comprises students from all the five fields of study of the University of Bergamo: Foreign Languages (41.6%), Economics (28.3%), Human and Social Sciences (17%), Engineering (9.8%), and Law (3.3%). They are enrolled on programmes for a bachelor degree (62.5%), a master degree (35.1%) and a five-year degree of study (2.4%). As far as international experience is concerned, about half of the students spent the fall semester abroad (48.4%) and 21.2% of students the spring semester; for just less than a third of students (30.4%) the experience lasted for the whole academic year. The preferred destination was Spain (27.2%), followed by Germany (17.9%), United Kingdom (15.7%), and France (14.3%); other European countries were overall chosen by 18.6% of the students, whereas the Extra-UE destinations were USA (2.5%), China (2.4%) and Australia (1.3%). 97.7% of the students went abroad through the Erasmus programme and the rest through international programmes.

In regard to the items of the third section of the questionnaire used to describe the two latent traits (items from I1 to I7 measure the impact of the international experience on the students' skills and items from I8 to I12 the fulfilment of expectations about the international experience), Table 1 shows the percentages for each response category. Most of the students think that the study abroad experience has had a significant impact on their adaptability to a new situation (78.6%), foreign language skills (72.8%) and communication skills (63.4%); the same students are very satisfied about the personality development (88.5%), foreign language skills (75.5%) and ability to interact with foreign people (69.8%).

With reference to the covariates used to detect the effect of individual characteristics on the latent class membership probability, Table 2 shows the percentages for each response category. The important factors in the decision to apply for a period of study abroad are curiosity about new challenges (88.5%) and improving language skills (87.7%); whereas students considering knowledge of the host country culture an important factor in the choice of the host university amount to only 42.5% and the cost of living is not important (56.6%). During the international experience, 42.1% of the respondents thinks that active participation is important for the final exam's evaluation and 58.6% that it is important to work in a team. Attending courses and

Table 1 Questionnaire items: distribution of response categories (%)

Item	Not at all/a little	Quite a lot	A great deal
I1. Communication skills	3.6	33.0	63.4
I2. Learning ability	26.0	45.3	28.7
I3. Foreign language skills	1.8	25.4	72.8
I4. Team working skills	14.1	42.8	43.1
I5. Adaptability	1.3	20.1	78.6
I6. Problem solving skills	6.1	39.5	54.4
I7. Motivation to study	30.6	42.5	26.9
I8. Enhance future employability in Italy	31.6	45.6	22.8
I9. Enhance future employability abroad	14.4	52.0	33.6
I10. Personality development	0.3	11.2	88.5
I11. Foreign language skills	2.0	22.5	75.5
I12. Ability to interact with foreign people	3.1	27.1	69.8

taking exams in a foreign language did not create difficulties for 86.3% of the students. The data also show that the mothers of 68.8% of respondents are employed, 20.2% are housewives, and the remaining 11.2% are retired, inactive or no longer alive. As regards the last two covariates considered, 93.6% of respondents carry out the experience abroad through the Erasmus programme, while the rest (6.4%) through the Extra-EU programme. Finally, for 87.6% of respondents the mobility scholarship is one of the three main funding sources.

3 Methods

In this section we first describe the statistical model used to assess the item responses and then we provide some hints for the model selection.

Table 2 Covariates: distribution of response categories (%)

Covariate	Not at all / A little	Quite a lot	A great deal
Motivations to study abroad			
Improve CV	10.1	39.9	50
Improve language skills	1.5	10.8	87.7
Improve host country knowledge	13.2	38.9	47.9
Curiosity about new challenges	1.1	10.4	88.5
Enhance future employability	22.5	40.5	37
Host university choice			
Knowledge host country culture	24.6	32.9	42.5
Living cost	56.6	34	9.4
Exams evaluation			
Active participation	18	39.9	42.1
Team working	11.2	30.2	58.6
Difficulty with foreign language	86.3	12.9	0.8
	Yes		No
Mother's details			
Housewife	20.2		79.8
Employed	68.6		32.4
Erasmus	93.6		6.4
Mobility scholarship	87.6		12.4

3.1 The Latent Class Item Response Model

The international study experience may be framed as a not directly observable construct composed of multiple dimensions. For this reason, we rely on the class of item response theory (IRT) models that allow to explain the probability of observed responses to multiple items (i.e., the items composing the questionnaire) as a function of one or more latent variables (i.e., the unobservable construct of our interest).

Let Y_{ij} be the response of individual i to item j ($i = 1, \ldots, n$; $j = 1, \ldots, J$), with y denoting the observable response category that may assume ordered values $y = 0, 1, \ldots, l_j - 1$. Let also $\boldsymbol{\Theta} = (\Theta_1, \Theta_2, \ldots, \Theta_s)'$ be the vector of latent variables that drive the response process, with $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_s)'$ realization of $\boldsymbol{\Theta}$. Often, a parametric assumption about vector $\boldsymbol{\Theta}$ is introduced (typically, the normal-

ity is assumed). However, this approach has some disadvantages. First, the estimation procedure may be computationally demanding because of the multidimensional integral characterizing the likelihood function. Second, a clustering of n individuals in a reduced number of groups instead of a ranking along a latent continuum may be of interest for the research aims. Hence, a valid alternative to the parametric approach is represented by a semi-parametric version of the IRT models, named Latent Class IRT (LC-IRT) models [1, 20]).

In the class of LC-IRT models, $\boldsymbol{\Theta}$ is assumed to have a discrete distribution with k support points, denoted by $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_k$, and mass probabilities (or weights) $\pi_{i1}, \ldots, \pi_{ik}$ $(i = 1, \ldots, n)$. Each support point identifies a latent class, that is a subpopulation of individuals having the same level of the latent construct measured by the questionnaire. Each weight π_{iu} $(u = 1, \ldots, k)$ denotes the probability of belonging to a certain latent class and is assumed to depend on a vector $\mathbf{X}_i = (X_{i1}, \ldots, X_{ip})'$ of observable individual characteristics, that is, $\pi_{iu} = p(\boldsymbol{\Theta} = \boldsymbol{\xi}_u | \mathbf{X}_i = \mathbf{x}_i)$.

The relation between latent variable $\boldsymbol{\Theta}$ and item response Y_{ij} is modeled as

$$\log \frac{p(Y_{ij} \geq y | \boldsymbol{\Theta} = \boldsymbol{\xi}_u)}{p(Y_{ij} < y | \boldsymbol{\Theta} = \boldsymbol{\xi}_u)} = \gamma_j \sum_{d=1}^{s} (z_{jd}\theta_d - \beta_{jy}), \; j = 1, \ldots, J; \; y = 1, \ldots, l_j - 1,$$

with z_{jd} dummy variable equal to 1 if item j measures latent trait θ_d and 0 otherwise, and γ_j and β_{jy} item discriminating and difficulty parameters, respectively, having the usual interpretation as in the item response theory context. The model at issue is named multidimensional Latent Class Graded Response Model (LC-GRM; see [2, 4] and references therein), because it resembles the global logit link of the GRM proposed in the parametric framework [18].

The manifest distribution of $\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{iJ})'$ follows as in a standard LC model [12, 15]

$$p(\mathbf{Y}_i = \mathbf{y}_i) = \sum_{u=1}^{k} p(\mathbf{Y}_i = \mathbf{y}_i | \boldsymbol{\Theta} = \boldsymbol{\xi}_u) \pi_{iu}$$

where the conditional probability of the response is specified as

$$p(\mathbf{Y}_i = \mathbf{y}_i | \boldsymbol{\Theta} = \boldsymbol{\xi}_u) = \prod_{j=1}^{J} p(Y_{ij} = y | \boldsymbol{\Theta} = \boldsymbol{\xi}_u),$$

in virtue of the local independence assumption, and the mass probability is expressed as a function of the individual covariates through a multinomial logit formulation

$$\log \frac{\pi_{iu}}{\pi_{i1}} = \delta_{0u} + \mathbf{x}_i' \boldsymbol{\delta}_{1u}, \quad i = 1, \ldots, n; \; u = 2, \ldots, k, \tag{1}$$

with δ_{0u} and $\boldsymbol{\delta}_{1u}$ denoting, respectively, the class-specific intercept term and the class-specific vector of regression coefficients.

The model at issue may be estimated by the maximum likelihood approach, through the EM algorithm; details are provided in [2]. Functions for the model estimation are implemented in the R package MultiLCIRT [3].

3.2 Model Selection

An important aspect that cannot be ignored in the class of models at issue, and, more in general, in the LC models, is the choice of the number k of latent classes. The main stream literature agrees on using the Bayesian Information Criterion (BIC) to select the value of k. In practice, a common approach consists in estimating the model for increasing values of k, being constant all the other elements of the model, and selecting that value of k corresponding to the minimum BIC. However, it has also to be taken into account that, in the presence of several covariates, the likelihood of the model and the value of BIC are affected by the selection process of covariates. In other words, if one or more covariates are eliminated from the model because not significant, the sequence of BIC values is likely to be different with consequences for the optimal k. Moreover, a strategy based on the minimization of BIC index (or other information criteria) often leads to the selection of a high number of latent classes with small weights and the profiles of individuals are not simple to interpret.

To solve the two above problems, we propose an iterative procedure consisting in the following steps:

Step 1: estimate the model without covariates and for increasing values of k ($k \geq 1$) until the first increasing BIC is obtained. Select that value of k corresponding to the first relative decrease of BIC smaller than a given (small) prespecified threshold, say 0.01;

Step 2: given k selected at Step 1, estimate the multidimensional LC-GRM with all the plausible covariates of interest. If all the covariates are significant at a given significance level (e.g., 5%), stop; otherwise go on to the next step;

Step 3: estimate the model with statistically significant covariates selected at Step 2 and for increasing values of k ($k \geq 1$) until the first increasing BIC is obtained. Select the value of k as in Step 1. If all the covariates are significant, stop; otherwise repeat Step 3.

The procedure finishes when there are no more changes in the value of k and in the set of significant covariates.

4 Results and Discussion

In this section we illustrate the main results obtained from estimation of the LC-GRM described in the previous section. We first provide details on the model specification and the student profiles. Then, we outline what are the most problematic aspects of

the international study experience. Finally, we focus on the individual covariates that most contribute to characterizing the student profiles.

4.1 Model Specification and Student Profiles

The LC-GRM we estimated is characterized by $s = 2$ latent dimensions, Θ_1 and Θ_2. Θ_1 denotes the impact of the international study experience on the student's skills, whereas Θ_2 describes the fulfilment of expectations concerning the international experience. These two latent variables are measured by $J = 1, 2$ items with three ordered response categories ($l_j = 0, 1, 2$) denoting increasing levels of positive impact or satisfaction. More in detail, Θ_1 is measured by the following 7 items:

- I1. Communication skills
- I2. Learning ability
- I3. Foreign language skills
- I4. Team working skills
- I5. Adaptability
- I6. Problem solving skills
- I7. Motivation to study

and Θ_2 by the remaining 5 items:

- I8. Enhance future employability in Italy
- I9. Enhance future employability abroad
- I10. Personality development
- I11. Foreign language skills
- I12. Ability to interact with foreign people

According to the above-described iterative procedure setting the significance level for the covariates at 5% and the reduction rate of BIC at 1%, we selected in seven steps a LC-GRM with $k = 3$ components. A summary of the selection procedure is displayed in Table 3, where the selected number of latent classes, the number of covariates, and the number of not significant covariates are indicated for each step of the procedure.

The proposed procedure leads to distinguish three profiles of students. As shown in Table 4, class 1 collects students that benefited just a little from the international experience (average weight equal to 23.0%), whereas students with significant advantage from this type of experience belong to class 3 (average weight equal to 29.4%). The remaining part of students (47.5%) is allocated in class 2, showing intermediate levels on both the latent variables. We also note a wide discrepancy between the estimated levels of Θ_1 and Θ_2 (both of them measured on the logit scale). Impact of skills assumes high values with a large variability, whereas fulfilling expectations about the international experience reaches lower levels with a smaller variability.

Table 3 Iterative model selection procedure: number of components selected (k), number of covariates (# cov.), number of Not Significant covariates (# NS cov.)

Step	k	# cov.	# NS cov.
1	9	None	–
2	4	53	32
3	4	21	1
4	3	20	1
5	4	16	4
6	3	15	1
7	3	14	0

Table 4 Estimates of support points $\hat{\xi}_{ud}$ and averages of weights $\hat{\pi}_{iu}$ ($d = 1, 2$; $u = 1, 2, 3$)

Latent variable d	Class $u = 1$	Class $u = 2$	Class $u = 3$
Impact on skills (Θ_1, $d = 1$)	1.849	4.349	6.578
Fulfilling expectations (Θ_2, $d = 2$)	0.036	0.808	1.820
Avg. weights	0.230	0.475	0.294

4.2 Item Parameters

An important advantage of an IRT is that it can be used to model parameters that are item-specific, thus making it possible to detect the most problematic aspects related with the latent construct. Estimates of difficulty parameters $\hat{\beta}_{jy}$ and discriminating parameters $\hat{\gamma}_j$ are displayed in Table 5 and plotted, together with the estimated support points, in Fig. 1.

To make interpretation of the model parameters easier, we recall that β_{j1} describes the difficulty of answering $y = 1, 2$ with respect to $y = 0$ and β_{j2} refers to answering $y = 2$ with respect to $y = 0, 1$. Thus, difficulty parameters shifted leftwards with respect to the latent variable level (i.e., in Fig. 1 horizontal segments on the left of the vertical segments) characterize items that are perceived as "easy" by the respondents. On the contrary, difficulty parameters shifted rightwards with respect to the latent variable level (i.e., in Fig. 1 horizontal segments on the right of the vertical segments) characterize items that are perceived as "difficult" by the respondents. In other words, "easy" items (left-shifted along the x-axis) are items for which low levels of the latent construct (impact on skills or satisfaction) are sufficient to receive substantial agreement from the respondents; conversely, "difficult" items (right-shifted along the x-axis) denote "critical" aspects of the latent construct under study that do not encounter the agreement of individuals, up to individuals belonging to a latent class with high support points.

Table 5 Estimates of item difficulty parameters $\hat{\beta}_{jy}$ and item discriminating parameters $\hat{\gamma}_j$ ($j = 1, \ldots, 12$; $y = 1, 2$)

Item	$\hat{\beta}_{j1}$	$\hat{\beta}_{j2}$	$\hat{\gamma}_j$
Θ_1: *Impact of the international experience on the student's skills*			
I1. Communication skills ($j = 1$)	0.000	3.687	1.000
I2. Learning ability ($j = 2$)	2.635	6.712	0.677
I3. Foreign language skills ($j = 3$)	−1.321	2.871	0.862
I4. Team working skills ($j = 4$)	0.783	5.408	0.561
I5. Adaptability ($j = 5$)	−1.916	2.287	0.804
I6. Problem solving skills ($j = 6$)	−0.202	4.376	0.704
I7. Motivation to study ($j = 7$)	2.752	7.416	0.477
Θ_2: *Fulfilling the expectations of the international experience*			
I8. Enhance future employability in Italy ($j = 8$)	0.000	2.218	1.000
I9. Enhance future employability abroad ($j = 9$)	−0.492	1.455	1.488
I10. Personality development ($j = 10$)	−1.277	−0.109	3.240
I11. Foreign language skills ($j = 11$)	−0.670	0.317	3.529
I12. Ability to interact with foreign people ($j = 12$)	−0.565	0.444	3.311

The most problematic aspects concerning the impact of international experience on skills (Θ_1) are represented by the learning ability (I2) and by the motivation to study (I7): in both cases, the difficulty parameters $\hat{\beta}_{j1}$ and $\hat{\beta}_{j2}$ ($j = 2, 7$) are definitely greater with respect to the other items as well as with respect to the support point of class 3 ($\hat{\xi}_{31} = 6.578$; Table 4). Conversely, acquisition of foreign language skills (I3) and adaptability (I5) are not problematic at all, because they likely find favour also with individuals belonging to class 1. The result concerning easy items entirely matches the evidence shown in The Erasmus Impact Study [5], in which nine in ten students reported gains in adaptability and improvements in foreign language skills. By contrast, the difficult items seem more related to mobile students' personal traits.

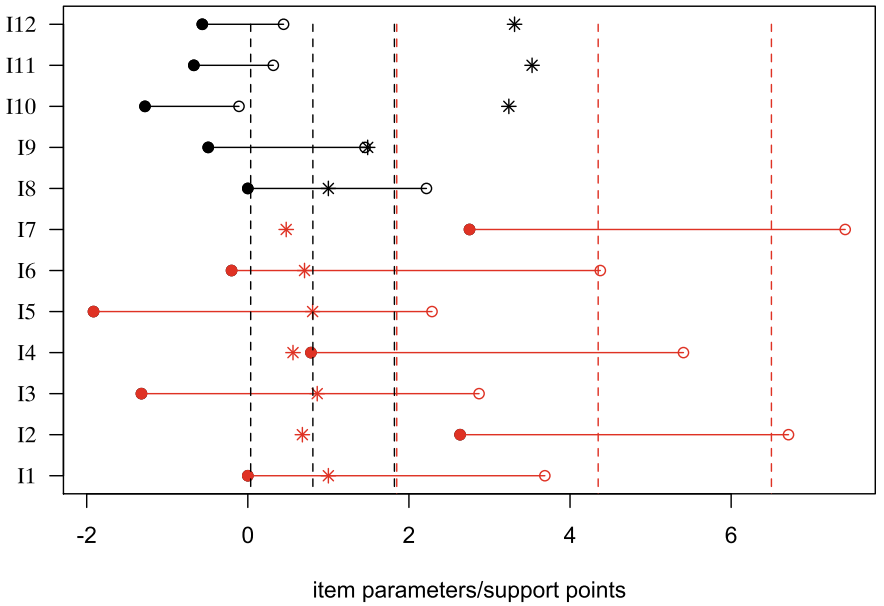


Fig. 1 Estimated item parameters: $\hat{\beta}_{j1}$ (solid circle), $\hat{\beta}_{j2}$ (empty circle), $\hat{\gamma}_j$ (star); estimated support points (vertical segments). Values that refer to latent variable Θ_1 are red colored and values that refer to latent variable Θ_2 are black colored

As regards the fulfilment of expectations (Θ_2), the utility of international experience for the Italian labour market (I8) is perceived as the most critical aspect, whereas the students' expectations in terms of personal growth (I10) are usually completely satisfied independently of the latent class. In the literature, results concerning the impact of international experience on labour market outcomes are quite heterogeneous and depend on the research approach. In general, studies based on surveys, such as [5], show that the experience abroad is beneficial for finding the first job and in short time. A positive impact on employability of studying abroad has also been found by studies like [8], but interestingly the effect is mainly driven by the impact that study abroad programmes have on the employment prospects of graduates from disadvantaged backgrounds. Other studies on earnings and the labour market return to international experience, such as [14, 19], show large variations in their results.

Finally, we note that the ability of the items to discriminate among similar levels of the latent construct is high for items I10, I11, I12, whereas it is quite low for the other items. In particular, items contributing to measurement of latent variable Θ_1 show a weak discriminating power.

4.3 Effect of Individual Characteristics

The effect of individual characteristics on the latent class membership probability is detected by the analysis of the regression coefficients $\hat{\delta}_{1u}$ $(u = 2, 3)$ of Eq. 1. Table 6 displays $\hat{\delta}_{12}$ and $\hat{\delta}_{13}$ for the covariates that resulted globally statistically significant at 5% according to the model selection procedure described in Sect. 3.2. We recall that $\hat{\delta}_{12}$ denotes how the covariate affects the odds of belonging to class 2 with respect to the odds of belonging to class 1, whereas $\hat{\delta}_{13}$ compares class 3 to class 1.

To facilitate interpretation of the regression coefficients Table 6 displays how significant covariates are distributed across the three latent classes. For each latent class, Table 7 shows the proportion of students selecting that characteristic. The allocation of individuals to the classes was based on the maximum a posteriori approach [13],

Table 6 Multinomial logit sub-model: estimates of regression coefficients $\hat{\delta}_{1u}$ $(u = 2, 3)$, standard errors, p-values

	$\hat{\delta}_{12}$	$\hat{\delta}_{13}$	$se_{\hat{\delta}_{12}}$	$se_{\hat{\delta}_{13}}$	$p_{\hat{\delta}_{12}}$	$p_{\hat{\delta}_{13}}$
Intercept ($\hat{\delta}_{02}$, $\hat{\delta}_{03}$)	−1.004	−4.480	0.877	1.095	0.252	0.000
Erasmus	−0.295	−1.332	0.589	0.596	0.617	0.026
Housewife mother	−1.199	−0.951	0.542	0.598	0.027	0.111
Employed mother	−1.223	−1.252	0.499	0.543	0.014	0.021
Mobility scholarship	0.938	0.959	0.342	0.412	0.006	0.020
Difficulty with foreign lang.	−0.457	−0.819	0.331	0.408	0.167	0.045
Motivations						
Improve CV	0.222	0.541	0.258	0.297	0.388	0.069
Improve language skills	1.139	1.837	0.332	0.52	0.001	0.000
Improve host country know.	0.623	1.333	0.273	0.304	0.022	0.000
Curiosity about new challenges	1.219	1.472	0.338	0.451	0.000	0.001
Enhance future empl.	1.028	1.837	0.318	0.341	0.001	0.000
Host university choice						
Know. host country culture	0.412	1.117	0.272	0.348	0.130	0.001
Living cost	0.364	0.618	0.261	0.294	0.164	0.036
Exams evaluation						
Active participation	0.498	0.821	0.305	0.328	0.102	0.012
Team working	−0.144	0.755	0.28	0.336	0.608	0.025

Table 7 Distribution of students in the latent classes, for each significant covariate (proportions)

Covariate	Class $u = 1$	Class $u = 2$	Class $u = 3$
Erasmus	0.947	0.955	0.900
Housewife mother	0.259	0.184	0.196
Employed mother	0.693	0.689	0.674
Mobility scholarship	0.767	0.899	0.913
Difficulty with foreign lang.	0.185	0.128	0.104
Motivations			
Improve CV	0.370	0.487	0.643
Improve language skills	0.693	0.918	0.965
Improve host country knowledge	0.233	0.484	0.670
Curiosity about new challenges	0.772	0.910	0.943
Enhance future employability	0.153	0.346	0.583
Host university choice			
Knowledge host country culture	0.598	0.774	0.852
Living cost	0.302	0.447	0.522
Exams evaluation			
Active participation	0.265	0.383	0.622
Team working	0.481	0.508	0.787

which consists of computing the conditional probabilities of belonging to the classes given the observed item responses, and, then, selecting the latent class with the highest posterior probability.

From Table 7 covariates related to motivations for studying abroad furnish a clear characterization of the three classes: improving own foreign language skills and curiosity about new challenges strongly characterize class 2 and, to a greater extent, class 3, according to an increasing trend (more than 90% of students selected these motivations). Then, they follow improving knowledge about the host country and improving own CV, chosen by more than 60% of individuals allocated to class 3 and about one half of individuals to class 2. The use of interactive teaching methods, such as team working, characterized the experience of 78.7% of students in class 3 and only 50% or less of students in classes 2 and 1. As regards the host country, the specific type of country is not included in the subset of significant covariates. However, the possibility of knowing the host country's culture was deemed important by 85.2% of students in class 3 versus 59.8% of students in class 1. Similarly, living costs are a relevant aspect for more than one half of students in class 3 versus only 30.2% of students in class 1. Finally, the type of international mobility programme

(Erasmus versus other programmes) and the mother's employment status do not play any specific role.

5 Conclusions

In higher education, participation in study abroad is perceived as highly beneficial to students in terms of personal development and academic and employment career. The results of our study confirm this perception. Mobile students report perceived benefits in terms of personality development (in particular, adaptability and communication and interpersonal skills) and improvement of their language knowledge and soft skills (such as problem-solving and decision-making skills). As regards the fulfilment of expectations about the international experience, students report gains in terms of personal growth and enhanced employability at home and abroad. The results indicate that individual characteristics have to be taken into account to explain differences among students in the latent class membership probability.

References

1. Bartolucci, F.: A class of multidimensional IRT models for testing unidimensionality and clustering items. Psychometrika **72**, 141–157 (2007)
2. Bacci, S., Bartolucci, F., Gnaldi, M.: A class of multidimensional latent class IRT models for ordinal polytomous item responses. Commun. Stat .Theory Methods **43**, 787–800 (2014)
3. Bartolucci, F., Bacci, S., Gnaldi, M.: MultiLCIRT: an R package for multidimensional latent class item response models. Comput. Stat. Data Anal. **71**, 971–985 (2014)
4. Bartolucci, F., Bacci, S., Gnaldi, M.: Statistical Analysis of Questionnaires: A Unified Approach Based on R and Stata. Chapman & Hall/CRC Press, Boca Raton (2015)
5. Consult, C.H.E.: Erasmus+ Higher Education Impact Study. European Commission, Brussels (2019)
6. Consult, C.H.E.: Brussels Education Services, Centrum fur Hochschulentwicklung, Compostela Group of Universities, Erasmus Student Network: The Erasmus impact study. European Commission, Brussels (2014)
7. CHE Consult GmbH and CHE Consult Prague s.r.o.: The Erasmus impact study. Regional Analysis. European Commission, Brussels (2016)
8. Di Pietro, G.: Do study abroad programs enhance the employability of graduates? Educ. Financ. Policy **10**, 223–243 (2015)
9. DZHW (Ed.): Social and Economic Conditions of Student Life in Europe: Synopsis of indicators. Final report. Eurostudent VI 2016–2018. W. Bertelsmann Verlag, Bielefeld (2018)
10. Commission, European: Erasmus—Facts Figures & Trends. Publications Office of the European Union, Luxembourg (2018)
11. Formann, A.K.: Mixture analysis of multivariate categorical data with covariates and missing entries. Comput. Stat. Data Anal. **51**, 5236–5246 (2007)
12. Goodman, L.A.: Exploratory latent structure analysis using both identiable and unidentiable models. Biometrika **61**, 215–231 (1974)
13. Goodman, L.A.: On the assignment of individuals to latent classes. Sociol Methodol. **37**, 1–22 (2007)

14. Kratz, F., Netz, N.: Which mechanisms explain monetary returns to international student mobility? Stud. High Educ. **43**, 375–400 (2018)
15. Lazarsfeld, P.F., Henry, N.W.: Latent Structure Analysis. Houghton Mifflin, Boston (1968)
16. Parey, M., Waldinger, F.: Studying abroad and effect on international labour market mobility: evidence from the introduction of Erasmus. Econ. J. **121**, 194–222 (2010)
17. Rodrigues, M.: Does student mobility during higher education pay? Evidence from 16 European Countries. JRC Scientific and Policy Reports, European Commission (2013)
18. Samejima, F.: Estimation of ability using a response pattern of graded scores. Psychometrika Monogr. **17**, Psychometrik Society, Richmond VA (1969)
19. Schnepf, S.V., D'Hombres, B.: International Mobility of Students in Italy and the UK: does It Pay off and for Whom? IZA DP No. 12033 (2018)
20. von Davier, M.: A general diagnostic model applied to language testing data. Brit. J. Math. Stat. Psy. **61**, 287–307 (2008)

Consumers' Preferences for Coffee Consumption: A Choice Experiment Integrated with Tasting and Chemical Analyses

Rossella Berni, Nedka D. Nikiforova, and Patrizia Pinelli

Abstract This study proposes an innovative approach for analysing consumer preferences for coffee by integrating a choice experiment with a guided tasting and chemical analysis. Firstly, two types of coffee were chosen from the mass market retailers with different sensorial profiles (100% Arabica, and Arabica and Robusta blends); subsequently, a guided tasting has been included to analyse the role of the sensory descriptors. An optimal design for the choice experiment was planned in order to achieve the joint purpose of the efficient estimation of the attributes, and the assessment of the information obtained from the guided tasting. The same choice experiment was administered twice, e.g. before and after the guided tasting. Random Utility Models were applied for better evaluating the consumers' behaviour.

Keywords Choice experiments · Optimal design · Random Utility Models-RUMs · Coffee tasting

1 Introduction

Coffee, one of the most frequently consumed beverages in the world, is a crucial commodity in the global economy. At the consumer level, the coffee quality has a "credence" dimension, related to convenience, product origin and process characteristics (organic production, social aspects), whereas the "experience" dimension of quality is mainly related to taste and flavour, as well as healthy properties. In the coffee sector, the presence of an eco-friendly label is particularly important, as the

R. Berni (✉) · N. D. Nikiforova · P. Pinelli
Department of Statistics Computer Science Applications "G. Parenti", University of Florence, Viale Morgagni 59, Florence, Italy
e-mail: rossella.berni@unifi.it

N. D. Nikiforova
e-mail: n.nikiforova@unifi.it

P. Pinelli
e-mail: patrizia.pinelli@unifi.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_4

environmental and social standards of this product have implications for the sustainability of viable tropical ecosystems and the well-being of agricultural producers and farm labourers. Hence, coffee is one of the first internationally traded products where collective efforts have been made to develop standards on processes that address socio-economic and environmental concerns.

Choice experiments (CEs) are one of the most applied preference evaluation methods for studying consumers' behaviour for a new product or service with applications in various fields. For instance, in the field of food and drinks, Kallas et al. [8] applied CEs for evaluating Catalan wine including the role of an advertisement campaign. Bi et al. [5] used a sensory evaluation method, and CEs, to assess to what extent the sensory characteristics affected the consumers' preferences for organic orange juice. Asioli et al. [1] studied consumers' preferences for iced coffee in Norway by comparing the results obtained through a CE and a rating-based conjoint experiment.

In this manuscript we propose an innovative approach to analyse consumers' preferences for coffee consumption by integrating a CE with consumer sensory tests and chemical analysis [4]. Two types of Italian brands of coffee ground for moka were chosen from the mass market retailers with different sensorial profiles: a soft, velvety and aromatic blend (100% Arabica), and an intense aromatic coffee (Arabica and Robusta blend). The caffeine content of coffee was acquired through a High-Performance Liquid Chromatography (HPLC) method with Diode Array Detector (DAD). A consumer sensory test was also planned, and to this end, a scoring card, where consumers have to give a score to the sensory descriptors of coffee, was developed. Furthermore, an optimal design was built for the CE through a compound design criterion [3]. The same CE was administered on two consecutive occasions [10], e.g. before and after the tasting, in order to analyse the role of the sensory evaluation in determining the consumers' behaviour. All this information was analysed through Random Utility models (RUMs), i.e. the Mixed Logit (MXL) model, allowing us to also study the respondents' heterogeneity.

The manuscript is organised as follows: Sect. 2 illustrates the theory related to the experimental design and the RUMs; the case-study and the results are described in Sects. 3 and 4 respectively; final remarks and discussion follow.

2 Outlined Theory for the Case-Study

In this Section we illustrate the fundamental theoretical elements of the CE and the RUMs applied (Sects. 2.1 and 2.2).

2.1 An Optimal Design for the Choice Experiment

A fundamental issue when dealing with the method of CEs is the underlying experimental design, consisting of the choice-sets to be administered to each respondent. With the aim of taking into account the information derived from the guided tasting, in the planned experimental design we built a variable that measures and summarises the quantitative information obtained through the two scorecards supplied during the tasting. More precisely, the optimal design for the CE was planned in order to build choice-sets for the joint purpose of: (i) an efficient estimation of the attributes, and (ii) the detection of the effect of the information obtained through the guided tasting and the chemical analysis. To this end, the following compound D-optimality design criterion Φ_C was applied [2, 3, 13]:

$$\Phi_C(\xi) = \frac{\alpha}{k_1}[\log|I_{11}(\xi)|] + \frac{(1-\alpha)}{k_2}[\log|I(\xi)| - \log|I_{11}(\xi)|] \quad (1)$$

where I_{11} is the Fisher Information Matrix (FIM) corresponding to the k_1 coefficients related to the attributes of the CE, while I is the FIM that contains both the k_1 coefficients for the attributes of the CE, and the k_2 coefficient for the variable related to the tasting, with constraint $k_1 + k_2 = k$. Furthermore, the coefficient α $(0 \leq \alpha \leq 1)$ in formula (1) reflects the relative interest in both objectives in the design criterion Φ_C. When $\alpha = 1$, we obtain a D-optimal choice design for the k_1 model coefficients, while for $\alpha = 0$ we obtain a D_s-optimal design for the k_2 coefficients. It must be noted that when considering the MXL model (Sect. 2.2), the FIMs (formula (1)) depend on the unknown coefficients' values. In order to deal with this issue we utilized Bayesian design techniques by integrating the design criterion Φ_C over a prior distribution of likely parameters values [6, 9]. Therefore, for obtaining the compound D-optimal design for the MXL model (Sect. 2.2) we assumed a Normal prior distribution for the mean vector β, and nominal values for the vector of heterogeneity coefficients. Thus, for obtaining the final optimal design for the CE, the compound design criterion is expressed as follows:

$$\Phi_{CB}(\xi) = \int_{\mathbb{R}^k} \frac{\alpha}{k_1}[\log|I_{11}(\xi)|] + \frac{(1-\alpha)}{k_2}[\log|I(\xi)| - \log|I_{11}(\xi)|]\pi(\beta)d\beta \quad (2)$$

The compound D-optimal design for the CE is carried out through the SAS/IML software, IML-Interactive Matrix Language. The Coordinate-Exchange algorithm is used to obtain the final optimal design [9, 12], and the k-dimensional integral in formula (2) is approximated through Quasi-Monte Carlo method and Halton sequencies [14]. The final optimal design is then composed of 24 binary choice sets. These choice-sets were randomly subdivided into groups composed of 8 choice-sets; then each respondent was asked to give his/her preference to 8 choice-sets.

2.2 The RUMs Applied for the Case-Study

In order to define the class of RUMs, we indicate every alternative by j; the choice-set is then formed by J alternatives $(1, \ldots, j, \ldots, J)$, while i denotes the respondent $(i = 1, \ldots, I)$. Moreover, the individual i who chooses the alternative j has a random utility U_{ij} expressed for each unit i, as a linear function $U_{ij} = x_{ij}^{'}\beta + \epsilon_{ij}$ formed by: (i) a deterministic part (the vector x_{ij}) containing the characteristics of respondent i and alternative j, while β is the vector of unknown coefficients, and (ii) a stochastic part ϵ_{ij}, that is the random component, generally supposed to be independent and also Gumbel or type I extreme value-distributed. Furthermore, we assume that the respondent i expresses his/her preferences according to the maximization of his/her utility.

The MXL model [11] belongs to the RUM class. It allows to evaluate the respondents' heterogeneity by considering the attributes' coefficients as random variables and not as fixed ones. A general formulation for a single decision when the choice is binary $(J = 2)$ is based on the assumption of a general continuous distribution for the $\psi_{ij,}$ called mixing term, so that the utility index U_{ij} becomes:

$$U_{ij} = V_{ij} + \psi_{ij} + \epsilon_{ij} \tag{3}$$

A density for ψ_{ij} is defined as $g(\psi|\Phi)$ where Φ contains the parameters of the error distribution, such as Normal, Uniform, Log-Normal. If ψ is not evaluated, then the MXL reduces to the simple Conditional Logit model. The expression for the MXL model, e.g. the unconditional probability related to the choice of a consumer i and a choice-set C_i, is equal to:

$$P(y_i = j) = P_i(j) = \int_{\psi} L_i\big(j|\psi_{ij}\big)g\big(\psi_{ij}|\phi\big)d\psi_{ij} \tag{4}$$

$$L_i\big(j|\psi_{ij}\big) = \frac{\exp(x_{ij}^{'}\beta + \psi_{ij})}{\sum_{k \in C_i}\exp(x_{ik}^{'}\beta + \psi_{ik})} \tag{5}$$

where $C_i = (1, \ldots, j, \ldots, J)$ is the general choice-set formed by J alternatives supplied to a consumer i; ψ is the random term for evaluating the respondents' heterogeneity. Formulas (4) and (5) are the general expressions of the MXL model, where the $g(\psi|\Phi)$ is the mixing component and the term $L_i(.)$, formula (5), is related to the deterministic part conditioned to the respondent's characteristics.

3 The Case-Study

3.1 The Choice Experiment Planning

The optimal design for the case-study has been outlined in Sect. 2 (formula (2)). The attributes of the CE were chosen after a thorough investigation of the coffee available from mass market retailers by especially considering type of coffee, sensory properties and price. More precisely, we identified six attributes for the CE reported in Table 1 (the coded levels of the attributes are reported within bracket): the type of coffee, and two attributes related to the coffees' taste (a soft and velvety and an intense and aromatic taste), both at two levels.

The Intense and Aromatic attribute defines the whole taste for the Arabica and Robusta blend, while the Soft and Velvety attribute completely defines the 100% Arabica blend. Therefore, a preference towards the Intense and Aromatic taste implies a preference *versus* the Arabica and Robusta blend, the same for the Soft and Velvety taste when considering the 100% Arabica blend.

Another two attributes refer to: (i) the packaging, e.g. a soft bag and a jar both in a modified atmosphere, and (ii) the "Label Indication", e.g. the indication of the geographical origin of the coffee, and the presence of any type of certification of sustainability (economic, social and/or environmental). The last attribute was the price at three levels: € 4.50, € 6.00 and €7.50, corresponding to a quantity of 250 g.

It must be noted that the study follows this order: background questionnaire, Choice 1 session, tasting of the two coffees (i.e. the guided tasting session), and Choice 2 session. Moreover, in both the choice experiment and the guided tasting sessions, the brands of the two coffee were blinded.

Table 1 Attributes (and levels) involved in the choice experiment

Attribute	Original Levels and coded levels between brackets
Coffee type	Blend of Arabica and Robusta (−1) Blend of 100% Arabica (+1)
Packaging	Soft bag in a modified atmosphere (−1) Jar in a modified atmosphere (+1)
Label indication	Geographical origin (−1) Certification of sustainability (+1)
Intense and aromatic taste	Fairly present (−1) Highly present (+1)
Soft and velvety taste	Fairly present (−1) Highly present (+1)
Price	4.50–6.00–7.50 euro

3.2 The Consumers' Guided Tasting

A brief description of the sensory and chemical characteristics of coffee was firstly illustrated, and followed by an explanation of the scorecard, consisting of a 7-point numerical scale, similar to those used by coffee firms for the sensorial quality assessment of coffee samples. Since our experiment was addressed to non-trained subjects, the guided tasting consisted of an explication on how to fill in the score-card, explaining that the lowest score of the scale corresponds to the worst perception category. Thus, the consumer tastes each type of coffee and then he/she assigns a score from 1 to 7 for colour, smell (aroma), taste and tactile sensations (body).

In order to include the information related to the tasting in both the experimental design and in the modelling estimation, we built a new variable, which is a synthesis of the scores assigned by each respondent during the tasting session. This variable, denoted by "Tasting", is evaluated within the experimental design in the range [−1, 1], where the extreme value (−1) is related to Arabica and Robusta blend, while the extreme value (+1) is related to 100% Arabica blend. Thus, the more the "Tasting" goes towards −1, the more the consumer's preferences tends towards the Arabica and Robusta blend, and *vice-versa.*

4 Model Results

Participants of this study were recruited by using mailing lists of cultural associations in Tuscany. The response rate was 15%; it was not a probabilistic sample; the respondents were N = 107. A background questionnaire was initially administered for gathering information on the respondents' socioeconomic characteristics as well as their concerns about sustainability and their purchasing behaviour. In Table 2 the description related to baseline variables is reported. Each session of the guided tasting was made for about 15–20 people; within each session, the consumer/taster assigned a different score for colour, smell (aroma), taste and tactile sensation (body).

The caffeine content, opportunely standardized, allows us to evaluate the consumers' preferences towards more or less strong coffee. The chemical analysis of the two types of coffee registered an average value of 53.67 and 81.08 mg/cup of caffeine for the 100% Arabica and the Arabica and Robusta blend respectively. These results indicate that more caffeine means a blend of Arabica and Robusta varieties.

The choice modelling results for Choice1 (C1) and Choice2 (C2) sessions are illustrated in Tables 3 and 4 respectively. The estimated MXL model for Choice 1 session is expressed as follows:

Table 2 Baseline variables description (N = 107 = 100%)

Variable	Percent	Variable	Percent
Gender		*Employee*	
Male	47.66	Student, retired, homemaker	56.07
Female	52.34	Office worker	28.04
Age		Manager	15.89
19–22	7.48	*Family situation*	
22–35	60.74	Poor	1.87
35–60	28.04	Below average	11.21
> 60	3.74	Average	40.19
Marital status		Above average	45.79
Unmarried	77.57	Affluent	0.93
Married	21.50	*Smoker*	
Other	0.93	Yes	23.36
Education		No	76.64
Up to middle school	7.48		
High school	37.38		
Degree	60.75		

$$U_{ij}^{C1} = \sum_{k'=1}^{3} x_{ijk'}\beta_{k'} + \sum_{k''=1}^{3} x_{ijk''}[\beta_{k''} + \sigma_{k''}Z_{k''}] + \sum_{l=1}^{L-1} x_{ijl}[\beta_l + \sigma_l Z_l] + \varepsilon_{\mathrm{ij}} \quad (6)$$

where $\beta_{k'}$ are the coefficients for the attributes of Coffee Type, Packaging and Label Indication; $\beta_{k''}$ and $\sigma_{k''}$ are the main effects and heterogeneity coefficients for the intense and aromatic taste, the soft and velvety taste, and the caffeine; β_l and σ_l are the main effects and heterogeneity coefficients for the price attribute ($l = 1, \ldots, 3$; $L = 3$); $Z_{k''}$ and Z_l are $k'' \times k''$ and $l \times l$ diagonal matrices containing the Standard Normal random draws, while ϵ_{ij} is the random component.

In Table 3, the positive sign of the estimated coefficient related to the type of coffee indicates a preference for the 100% Arabica blend, while the estimated coefficient of caffeine is positive, and tends towards a blend of coffee containing more caffeine. The positive sign of the estimated coefficient of the soft and velvety taste means a preference towards the highest level, while for the intense and aromatic taste, the negative sign of the estimated coefficient indicates a preference for its fair presence. Still observing Table 3, the consumers prefer a label indication of sustainability, and soft-bag packaging. The two levels of the price attributes are both negative with

Table 3 Model results for Choice1 (C1) session

Attribute	Estimate	St. error	p-value
Coffee type	0.9731	0.1829	< 0.0001
Packaging	− 0.2232	0.1168	0.0560
Label indication	0.7961	0.1751	< 0.0001
Intense and aromatic taste	− 0.8586	0.1752	< 0.0001
Soft and velvety taste	1.2066	0.2341	< 0.0001
Caffeine	1.5448	0.3882	< 0.0001
Price-1st level	0.0000	.	.
Price-2nd level	− 0.6303	0.3090	0.0413
Price-3rd level	− 1.0360	0.4622	0.0250
Heterogeneity effect			
Intense and aromatic taste	− 0.0177	2.0455	0.9931
Soft and velvety taste	0.0544	2.3370	0.9814
Caffeine	2.0324	0.5791	0.0004
Price-1st level	0.0000	.	.
Price-2nd level	1.7428	0.6422	0.0067
Price-3rd level	0.9136	1.1406	0.4231
AIC = 776.7836 BIC = 838.5630 McFadden's LRI = 0.3673			

respect to the reference level (price1): therefore, the consumers' willingness-to-pay (WTP) decreases when the price rises. It must be noted that we report the estimated heterogeneity effects in terms of σ rather than σ^2. Only two attributes show a highly significant heterogeneity, e.g. the caffeine and the second price level.

Regarding the C2 model results (Table 4), the estimated MXL model is similar to those for Choice 1 modelling (formula (6)), where the only difference is the inclusion of the main effect and the heterogeneity coefficient for the variable "Tasting". When observing Table 4, it can be noted how all the estimated coefficients are significant or almost significant. The coffee type coefficient changes the sign, from positive to negative, indicating a preference for the Arabica and Robusta blend, while the estimated caffeine coefficient still remains positive, indicating a preference for a coffee containing more caffeine. The negative sign of the coefficient for soft and velvety taste indicates a preference for its fair presence level, and the positive sign for intense and aromatic taste indicates a preference for its high presence level, both results confirming the tendency to choose the Arabica and Robusta blend. By considering the packaging, the soft bag preference is confirmed as well as the label indication with regard to sustainability. In this second modelling step, the guided tasting information is evaluated through the variable "Tasting"; its estimated coefficient is negative and this is the further confirmation of the respondents' preferences towards the Arabica and Robusta blend. Lastly, the two estimated coefficients related to the price attributes show the same pattern obtained in C1 modelling, confirming

Table 4 Model results for Choice2 (C2) session

Attribute	Estimate	St. error	p-value
Coffee type	− 0.7403	0.1688	< 0.0001
Packaging	− 0.1954	0.1140	0.0866
Label indication	0.9999	0.1961	< 0.0001
Intense and aromatic taste	0.7102	0.1958	< 0.0003
Soft and velvety taste	− 1.0427	0.2163	< 0.0001
Caffeine	1.5973	0.4471	< 0.0004
Tasting	− 2.4315	1.2361	0.0492
Price-1st level	0.0000	.	.
Price-2nd level	− 0.6668	0.2665	0.0124
Price-3rd level	− 2.2212	0.5658	< 0.0001
Heterogeneity effect			
Intense and Aromatic Taste	0.8485	0.3821	0.0264
Soft and Velvety Taste	0.5964	0.3761	0.1127
Caffeine	2.5790	1.0094	0.0106
Tasting	1.8391	7.6155	0.8092
Price-1st level	0.0000	.	.
Price-2nd level	− 0.0181	4.7156	0.9969
Price-3rd level	− 1.8915	0.8677	0.0293
AIC = 886.1176 BIC = 957.4017 McFadden's LRI = 0.2786			

the previous WTP results. When considering the heterogeneity coefficients (Table 4, bottom), the results show a very low heterogeneity effect. In fact, we can observe how only three heterogeneity coefficients are significant: the intense and aromatic taste, the caffeine, and the 3rd price level. The soft and velvety taste coefficient shows a relevant, even though non-significant p-value ($p = 0.1127$). Nevertheless, two heterogeneity coefficients show very large standard errors, e.g. the variable "Tasting" and the 2nd price level. For the variable "Tasting" this is probably due to a larger variability because it is an overall synthesis of the various taste descriptors.

Furthermore, we also evaluated the relative importance of each attribute calculated as the ratio between the part-worth utility range for a given attribute and the sum of the part-worth utility ranges for all the attributes [7]. The results are reported in Table 5 for both C1 and C2 modelling steps. When considering C1 modelling, the most important variable turns out to be the caffeine, while the less important one is the packaging. A relatively high importance is also obtained for the price and the soft and velvety taste, as well as for the coffee type and the intense and aromatic taste. In C2 modelling, the most important attribute is the "Tasting"; this result further confirms the relevant role of the guided tasting for better defining the consumers' preferences. The caffeine is still highly important for the respondents, while the remaining attributes confirm

Table 5 Attributes relative importance in C1 and C2 modelling

Attribute	Attribute relative importance in C1 modelling	Attribute relative importance in C2 modelling
Coffee type	0.1094	0.0474
Packaging	0.0251	0.0125
Label indication	0.0895	0.0639
Intense and aromatic taste	0.0986	0.0997
Soft and velvety taste	0.1418	0.1048
Price	0.1334	0.1315
Caffeine	0.4022	0.2671
Tasting		0.2731

pretty much the same importance as in C1 modelling. Lastly, when considering C2 modelling, we can note that the inclusion of the variable "Tasting" explains part of variability and importance previously explained by the attributes of Coffee Type, Label Indication and Soft and Velvety taste, which obtained higher importance in C1 modelling.

Moreover, for each estimated MXL model (Tables 3 and 4) we also report some goodness-of-fit measures, i.e. the Akaike's (AIC) and the Schwarz's (BIC) indexes, and the McFadden's Likelihood Ratio Index (LRI or also called Pseudo-R^2). The goodness-of-fit is satisfactory for both models. The application is performed by using the procedure Multinomial Discrete Choice-MDC (SAS System; Windows Platform vs. 9.4).

5 Final Remarks and Discussion

The Choice modelling results (Sect. 4) clearly indicate that the guided tasting session, together with the information provided on the chemical composition of coffee, plays a relevant role in defining the consumers' behaviour. In fact, in C1 session, the respondents expressed their preferences considering their previous knowledge about coffee and their lifestyle, deriving quality expectations from credence attributes, rather than experienced attributes. This key aspect probably explains the controversial results obtained in C1 modelling, related to the choice of 100% Arabica blend jointly with a preference for more caffeine, by also considering the relevant heterogeneity effect of the caffeine with a highly significant p-value. Instead, in C2 session, both the information provided by an expert about coffee and the guided tasting session contributed to obtain more informative and accurate preferences expressed by the respondents. In fact, as a further confirmation of this result, the changes in the consumers' preferences between C1 and C2 sessions are only related to the type of coffee and the two attributes related to its taste properties, e.g. the soft and velvety and the intense and aromatic tastes. Moreover, the results obtained in C2 modelling are in coherence with

the quantity of caffeine because the respondents chose the Arabica and Robusta blend, and in line with this, they showed a preference for more caffeine. The same pattern is verified for the attributes of soft and velvety and intense and aromatic tastes; both attributes showed a low heterogeneity effect by also considering the corresponding lower standard errors achieved in C2 with respect to C1 modelling step. Moreover, in C2 modelling the "Tasting" coefficient confirms the consumers' preferences for the Arabica and Robusta blend. When considering the price attribute, the estimated coefficients show the same signs in both sessions, confirming that the consumers' WTP decreases when the price rises. However, the two price levels show a different pattern. More precisely: (i) the 2nd price level is almost steady in both C1 and C2 model results; (ii) the 3rd price level changed its importance in C2 with respect to C1 modelling. Moreover in C2 session, the 3rd price level attribute shows a significant heterogeneity effect; this issue could be a sign of an increasing uncertainty of the respondents WTP after the guided tasting. Further analyses could be performed by evaluating socio-economic and respondent characteristics within the modelling step, also developing and improving specific issues related to the integration among CEs, chemical analyses and scorecards.

References

1. Asioli, D., Almli, V., Næs, T.: Comparison of two different strategies for investigating individual differences among consumers in choice experiments. A case study based on preferences for iced coffee in Norway. Food Qual. Prefer. **48**, 174–184 (2016)
2. Atkinson, A.C., Bogacka, B.: Compound D- and Ds-Optimum designs for determining the order of a chemical reaction. Technometrics **39**, 347–356 (1997)
3. Atkinson, A.C., Donev, A.N., Tobias, R.D.: Optimum Experimental Designs, with SAS. Oxford University Press, Oxford (2007)
4. Berni, R., Nikiforova, N.D., Pinelli, P.: Consumers' preferences for coffee consumption: a choice experiment including organoleptic characteristics and chemical analyses. In: Abbruzzo, A., Brentari, E., Chiodi, M., Piacentino, D. (eds.) Book of Short Papers SIS 2018, pp. 1442–1448, Pearson, Italia
5. Bi, X., Gao, Z., House, L.A., Hausmann, D.S.: Tradeoffs between sensory attributes and organic labels: the case of orange juice international. Int. J. Consum. Stud. **39**, 162–171 (2015)
6. Chaloner, K., Verdinelli, I.: Bayesian experimental design: a review. Stat. Sci. **10**, 273–304 (1995)
7. Green, P.E.: Conjoint analysis in consumer research: Issues and outlook. J. Consum. Res. **5**, 103–123 (1978)
8. Kallas, Z., Escobar, C., Gil, J.M.: Assessing the impact of a Christmas advertisement campaign on Catalan wine preference using choice experiments. Appetite **58**, 285–298 (2012)
9. Kessels, R., Jones, B., Goos, P., Vandebroek, M.: An efficient algorithm for constructing bayesian optimal choice designs. J. Bus. Econ. Stat. **27**, 279–291 (2009)
10. Lombardi, G.V., Berni, R., Rocchi, B.: Environmental friendly food. Choice Experiment to assess consumer's attitude toward climate neutral milk: the role of communication. J. Clean. Prod. **142**, 257–262 (2017)
11. McFadden, D., Train, K.: Mixed MNL models for discrete response. J. Appl. Econom. **15**, 447–470 (2000)

12. Meyer, R.K., Nachtsheim, C.J.: The coordinate-exchange algorithm for constructing exact optimal experimental designs. Technometrics **37**, 60–69 (1995)
13. Wynn, H.P.: The sequential generation of D-optimal expemental designs. Ann. Math. Stat. **41**, 1055–1064 (1970)
14. Yu, J., Goos, P., Vandebroek, M.: Efficient conjoint choice designs in the presence of respondent heterogeneity. Mark. Sci. **28**, 122–135 (2009)

Urban Transformations and the Spatial Distribution of Foreign Immigrants in Messina

Francesca Bitonti, Angelo Mazza, Massimo Mucciardi, and Luigi Scrofani

Abstract Messina exhibits a fragmented urban structure, a consequence of past historical events, mainly the 1908 earthquake. After this tragic event, Messina experienced economic downturns and nowadays it passively suffers rather than managing its considerable mercantile traffics. The fragmented urban fabric affects the residential location of foreign migrants. Related literature distinguishes between two sources of spatial segregation: apparent contagion (i.e. economic inhomogeneities affecting the urban context) and true contagion (individual preference to live close to ethnically similar neighbors). We use point pattern analysis to assess residual clustering of migrant households while adjusting for economic inhomogeneity. We implement a case–control approach to avoid confounding between the two sources: migrant households represent cases, while a random sample of natives constitutes the controls. Results show that Sri Lankans, Filipinos (exceeding one kilometer), and Romanians exhibit the highest voluntary segregation, contributing to the creation of spatial clusters that boost the polycentric structure of Messina.

Keywords Territorial organization · Spatial analysis · Foreign immigration · Voluntary segregation

F. Bitonti · A. Mazza · L. Scrofani (✉)
Dipartimento di Economia e Impresa, Università Degli Studi di Catania, Catania, Italy
e-mail: scrofani@unict.it

F. Bitonti
e-mail: francesca.bitonti@phd.unict.it

A. Mazza
e-mail: amazza@unict.it

M. Mucciardi
Dipartimento di Scienze Cognitive, Psicologiche, Università Degli Studi di Messina, Pedagogiche e Degli Studi Culturali, Messina, Italy
e-mail: massimo.mucciardi@unime.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_5

1 Introduction

In the present work, the authors study the settlements of foreign immigrants which have contributed to shape the urban dynamics of the metropolitan city of Messina during the last decades. The authors briefly investigate the history of Messina and its territorial and economic organization in relation to the transport infrastructure network and the population settlements. The historical analysis reveals the crucial role of the commercial areas and the old town, which also hosts the harbor, a symbol and flywheel of Messina. The historical borough corresponds also to the place of residence for the majority of foreign immigrants, relatively to the other neighborhoods. Subsequently, a spatial analysis of the distribution of foreign immigrants is carried out. The empirical application provides insights into how the presence of foreign immigrants can profoundly affect the urban development and the identity of Messina itself.

Spatial models have been implemented to study the flows and residential location of foreign immigrants in Messina, in particular those of the most numerous six ethnic groups, in order to highlight the phenomenon of voluntary segregation. Traditional studies in social research distinguish between two different determinants of residential segregation: the economic inhomogeneities characterizing the urban context (i.e. rent costs and job availability) and the spatial attraction between individuals who share the same identity and culture. Nevertheless, conventional indices of spatial segregation do not consider the different influences exerted by the two sources of spatial clustering. These indices generally rely on census tracts, i.e. aggregated scales, which do not allow for analyses at the individual household level. Therefore, the authors employ individual household data from the city's Population Register to measure the spatial distribution of foreign immigrants in Messina. The inhomogeneous K-function [1] is used to evaluate spatial attraction, ruling out the effects of spatial inhomogeneity. The distribution considers also group classification by work specialization, e.g. housework and caregiving, which are becoming increasingly widespread among migrants compared to business and retailing. The application of the models to the data shows a significant voluntary spatial segregation for Sri Lankans, Romanians, Moroccans, and Filipinos. Conversely, other ethnic groups are consistent with random allocation. The analysis of the spatial distribution of foreign immigrants highlights the presence of work specialization among migrants belonging to different nationalities. Some ethnic groups concentrate within the urban areas to comply with the work demand exhibited by the native population, as the demand for domestic work and for retailing. This mechanism finds explanation in the multipolar identity of Messina and its metropolitan role.

2 Urban Activities and Migrant Settlement Patterns

The current urban fabric of the municipality of Messina is highly heterogeneous and articulated in a substantial set of hill towns, hamlets, and seaside villages which extend for about sixty kilometers along the coast. Its development on several poles, connected by a complex network of roads, highways, railways and port areas, could not fail to affect its dynamic, multicultural, in transition, and therefore nuanced identity, which is of uncertain definition, open, like that of many crossroads (see the strategic plan in [2, 3], p. 20).

At the beginning of the fifteenth century, the city flourished thanks to the port area which sustained substantial commercial traffic, like the one related to the export of silk. The port also sustained portable maintenance activities to the numerous ships that landed there [4]. The period of prosperity granted the city even to issue its own coin [5, p. 43]. Messina experienced an economic ferment which was accompanied by a cultural dynamism and an urban development. The mid-August fair, which gathered the European and Asian producers and traders, the foundation of the Silk Consulate, the Greek School, and the University (1548) are only some pieces of evidence of the economic and cultural growth the city was experiencing. On the other hand, the construction of the San Salvatore fort at the entrance to the port, of the Palazzata (a magnificent build on the access to port: see [6], of a new arsenal, and new fortifications testified the urban expansion of the city. Messina became one of the most prominent crossroads of commerce between Europe, Asia, and Africa [7].

At the end of the seventeenth century, the Spanish repression against the Messina turmoil reversed the flourishing condition of the city leading to an economic and political, but also demographic downturn. As a consequence, Messina barely reached 40,000 inhabitants during the second half of the eighteenth century. During the nineteenth century, the renovated role of the city harbor as a strategic point of connection to the mainland proved to be insufficient to compete with the opening of the Suez Canal and the construction of roads to Palermo and the South. These events contributed to shaping the actual fragmented and disconnected structure of Messina [5, p. 89].

During the late sixties of the nineteenth century, a city plan known as the Spadaro plan constituted the first attempt to rationalize the urban fabric, supporting the development of the city activities along the coastline and mountain roads [8]. According to the 1861 census, the city of Messina had a population of 103,324 inhabitants distributed among the urban center (62,024) and other 48 villages included within the urban perimeter [9]. The renovated and incipient role of the city as a commercial and financial epicenter was undermined by the expansion of nearby ports (Milazzo to west, Riposto to east and Calabrian ports to the north), the abolition of the free port, the citrus monoculture practice, and the loss of the function of transit port for coal ships, replaced by new vessels fueled with liquid fuel which allowed for longer routes [10, pp. 63–65].

At the beginning of the twentieth century, Messina became port of departure of many migratory waves: firstly, to the Americas, later to Northern Italy and Europe. Messina was almost destroyed by an earthquake in 1908, after which the

city underwent a serious economic crisis [11]. Caminiti [12] reports that, above all, the forced displacement of the surviving population and its next return had disastrous implications on the identity not only of the city but also of the Unitary State.

The reconstruction city plan proposed by Borzì aimed at reorganizing the devastated urban fabric and boosting the urban extension to the north and south. It is clear, however, that in the urban reorganization of the anthropic activities natural elements, such as streams, acted as reference points and sometimes as limits [13, p. 21].

Cairoli square and the tangent Cannizzaro street became the northern urban boundary of the new expansion towards the south, where the new buildings of the areas adjacent to the San Filippo and Zafferia streams gathered. Other areas of expansion, starting in the late 1960s, affected the northern hilly areas along the ring road. Borzì's urban plan of 1911 remained largely unfulfilled due to the war events, favoring—when possible and in the absence of central spaces—the construction of an ultra-popular building in the poor neighborhoods of Gazzi and Camaro, where the wooden slums have not always been replaced by brick houses [14].

On one side, the earthquake allowed reviewing the social and economic hierarchies and their spatial organization, while it fostered the urban speculation on the other. Different centers benefited from the crisis of Messina, in particular the city of Catania which became the reference center of the Eastern Sicily [15, p. 424]. After the seismic event, Messina developed a polycentric spatial structure (confirmed also by Benito Mussolini during his visit to the city on August 10, 1937). The multipolar urban organization has persisted still today and has fragmented the identity of the city. The intentions of the current urban planners are those to regenerate the urban context to convert Messina from an emergency city to a center of hospitality and environmental protection, avoiding further soil exploitation [3, p. 42].

The aftermath of the Second World War compelled the city to rise again and to reconvert the economic activities from the production sector to the services one. Administrative activity, health service (General Hospital since the 1950s), tourism (cruise traffic), cultural and educational service, transport, and trade replaced the industrial and handicraft works. The production plants have been located indeed outside the urban center, except for the Industrial Area of Larderia. In particular, the transport network has become pain and pleasure for the current municipal economy. The contemporary urban fabric is a maze of streets which includes also two highways (A20 and A18) and the port arterial roads (relative to the ports of Messina Zona Falcata, Messina Tremestieri, and Messina Rada San Francesco). Today the Zona Falcata old harbor receives a significant inflow of cruise ships, besides the ferry trains, ferryboats, and wheeled transports managed by maritime carriers.

Today, the realization of the Tremestieri port contributes to lighten the city center from the heavy traffic heading and departing from Zona Falcata [16]. In particular, Tremestieri should convey the heavy Ro-Ro traffic, while the Rada San Francesco port should be reserved for the light Ro-Ro traffic [3, p. 9]. Nevertheless, Tremestieri harbor is frequently unserviceable because the seafloor requires continuous cleaning from the debris. As a consequence, during the closure periods, the traffic begins to congest again in the urban center.

To date, Messina represents a strategic hub of international and national trade, and people flows. As a consequence many thousand cars, trucks, and people pass through the districts, paralyzing the entire urban area [16]. During the last four decades, the unbearable urban congestion, the increasing youth emigration, and the population movement toward the neighboring villages are leading Messina to a demographic decline (for an analysis of the urbanization of Messina and other southern cities see [17]). Also, the Zona Falcata port keeps congesting the road traffic in the 4th district (see Fig. 1): the most critical urban area which hosts the traveler influx, holding also residential, touristic, and commercial roles [3, p. 11]. The 80% of the workforce of Messina, which has 232,555 individuals (Dec 31, 2018) and has been steadily decreasing for the last twenty years, is employed in trade, public administration, and other tertiary activities.

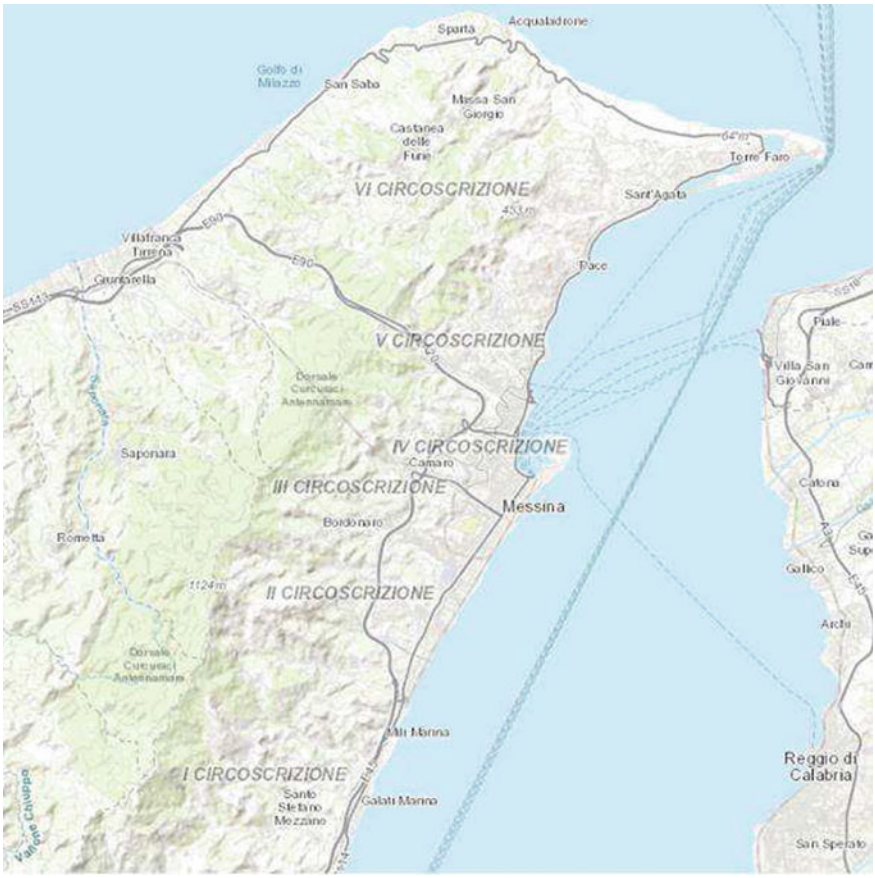


Fig. 1 Messina, the districts and the major transport connections by sea and land (*Source* http://www.comune.messina.sitr.it/)

The population structure has undergone a radical change during the last forty years. The city is inhabited by a large number of elderly and foreign migrants [18]. While in 1971 there was one old person per child, in 2011 the ratio amounted to four to one. Today the most populous district is the third one, with 55,450 residents, whereas the least is the first one, inhabited by 22,363 citizens. At the end of 2018, Messina counted 12,265 foreign inhabitants (equal to 5.27% of the total population) coming from 116 different countries, of which 13.67% born in the city (1,702 units). The majority of the foreign migrants come from Asia (57.4%), while the largest ethnic groups are Sri Lankans (4,048 units), Filipinos (2,365 units), and Romanians (1,629 units) according to official data, but irregular migrants account for a relevant share [19].

Messina attracted multiple waves of immigration due to economic, political, or family reasons pertaining to the various communities. The first wave dates back to the 1930s when an important influx of migrants came from Australia. The countries of departure of the second massive arrival occurred between the 1970s and the 1980s, were instead the European ones (especially Greek[1]). Starting from the 1980s, Messina has attracted Sri Lankans,

Filipinos, Chinese but also Polish and many Romanians.

The foreign residents of Messina are mostly unmarried because they are minors and young people, or cohabitants (http://www.strettoweb.com/foto/2020/01/interculturalfest-presentato-lo-studio-sugli-stranieria-messina/38895/).

The 4th district has the highest incidence of foreign citizens, equal to 10%, while the 2nd district has the lowest one, about 2% [20, p. 5]. The majority of foreign-born residents (80%) is allocated in the central areas of the city (3rd and 5th district). This particular concentration is attributable to the migrants' employment in domestic services or assistance to older people. Not surprisingly, the 4th district is characterized by the highest residents' mean age of Messina, despite the presence of young foreign inhabitants.

The housing emergency deriving from the post-earthquake reconstruction affects the city of Messina still today. A multitude of huts (3,000 units) dots the landscape of many neighborhoods (e.g. Giostra, Camaro, Gonzaga, and Fondo Fucile districts: see [21]). In 2011 the economical dwellings amounted to 47%, while the valuable ones accounted for slightly more than 3% [3, p. 31]. The housing issue has conditioned the public building programs and urban development but has also contributed to the birth of slums, both within the center and periphery [22]. For a long period, foreign immigrants have been massively supporting the demand for housing in the city [3, p. 31].

The heavy presence of foreign-born residents in the central urban areas depends on their work as housekeepers and caregivers, as mentioned before, but also as peddlers in the daily or weekly markets.

[1] The Greek community is particularly active and in 2010 has instituted the Hellenic Community of the Strait, seating in Messina; an association gathering Greeks and individuals with Greek origins, permanently or provisionally resident in the province of Messina, Reggio Calabria, Vibo Valentia, Catanzaro and Crotone, known also with the acronym C.E.D.S. The association is apolitical, nondenominational and non-profit; it organizes cultural events and free Greek language courses.

3 Data and Methods

A minority ethnic group is considered spatially clustered when its spatial arrangement diverges from the one expected under a random spatial distribution [23]. Broadly speaking, it is possible to identify two sources of spatial clustering: spatial inhomogeneity or apparent contagion, and spatial attraction or true contagion. The first one concerns inhomogeneities that might yield economic-based segregation: the accessibility to low-cost public facilities, the large variability in housing prices depending on the different urban areas, and the availability of specific work activities. In this sense, one might argue that the different ethnic groups would not be randomly allocated, even if ruling out the influence of ethnic discrimination on residential choices [24].

As to the second source of spatial segregation, Clark and Fosset [25] reported that each of the ethnic groups in the USA prefers living in neighborhoods where their own group represents the majority or near-majority. Foreign immigrants can benefit from positive spillovers in dwelling close to their compatriots, as reciprocal acceptance, common language, and support in general. Nevertheless, despite the reasons underlying individual preferences toward ethnic attraction, the two sources yield comparable configurations of residential segregation [25]. Therefore, it is relevant to disentangle true from apparent contagion. The "economy–induced" segregation provides some insights to understand the general level of segregation and raises issues of social equity. Conversely, the Schelling model shows how the individual voluntary intentions to live close to ethnically similar neighbors is a prominent cause of residential segregation [25, 26].

The spatial distribution of households may be represented by a point pattern on a map [27–29]. In the context of spatial analysis, the simplest theoretical model for spatial point pattern is that of complete randomness, called homogeneous Poisson process (HPP). In HPP the expected value of events (in our case resident households) occurring within a unitary region $u \in R$ follows a Poisson distribution, whose intensity λ is uniformly distributed over R [26]. This implies that the region is completely homogeneous in terms of its ability to attract households and that their residence choices are random and independent among them. Most often, the HPP is the standard null hypothesis of spatial analyses; however, within urban areas, the assumption of constant intensity λ of HPP is not suitable: potential residential density, for instance, renting prices or the availability of specific occupations may considerably vary across the different areas of the city. The inhomogeneous Poisson process (IPP) is a generalization of the HPP obtained replacing the constant intensity λ with a spatially varying intensity function $\lambda(u)$, $u \in R$.

Ripley's K-function [1], usually denoted by $K(d)$, is used to detect the presence of clustering (events gathering in specific areas, as shops in a mall) or dispersion (events following a regular pattern, as the gas stations along the highways or drugstores and tobacco shops in a city) with respect to a completely random allocation of events in the region of interest. The K-function counts the number of additional events occurring within a circle of radius d surrounding an arbitrary event; this number

is then compared with that expected in an HPP: if the observed number is similar, higher or lower than to an HPP, it is possible to infer random distribution, clustering or dispersion respectively. Letting d varying it is then possible to repeat the analysis at different geographic scales. Baddeley et al. [30] generalized the K-function to the inhomogeneous point process. The inhomogeneous K-function proposed denoted by $K_{inhom}(d)$ hereafter, allows measuring the concentration (or dispersion) surplus compared to the configuration yielded by random allocation, netting out the spatial inhomogeneities identified by variations in $\lambda(u)$.

When the analysis of a spatial point pattern detects a high level of concentration (or dispersion) it is challenging to distinguish whether it stems from variations in the characteristics of the territory (the cost of rents, for instance, may induce phenomena of economic segregation) or it depends on attractive (or repulsive) spatial dynamics [31]. This issue has been addressed by Mazza and Punzo [32], implementing a case–control approach. The data of a case–control study consist of the realizations of two distinct spatial point processes: the first representing cases of a condition of interest (e.g. possessing foreign citizenship), the second representing controls, a sample randomly drawn from the total population. Assuming that the sampling fraction of controls tends to zero compared to the population at risk, controls represent the realization of an IPP with intensity $\lambda^*(u)$. Where the population at risk tends to cluster, the number of observed controls will be higher on average. Cases, in turn, constitute a second independent point process with intensity $\lambda(u)$. Here, it is of relevance to verify whether cases form an IPP with intensity proportional to that of the controls, that is $\lambda(u) = \rho\lambda^*(u)$ (with $\rho = 1$ when the number of cases equals that of controls), or whether they present a different spatial structure. Diggle et al. [33] show that the ratio between the two intensity functions may be modeled depending on a vector of m explanatory spatial variables $z(u) = (z_1(u), ..., z_m(u))^j$, that is

$$\lambda(u) = \lambda^*(u) f(z(u); \theta) \tag{1}$$

where $f(\cdot)$ is any nonnegative function with parameters θ.

Here controls are a random sample of households drawn from all households with Italian citizenship, the explanatory variables employed are related to the price of rents.

Spatial intensity $\lambda(u)$ of cases is estimated by the model

$$\lambda(u) = \lambda^*(u)\exp\{\alpha + z(u)^j\beta\} \tag{2}$$

where $\beta = (\beta_1, \ldots, \beta_m)$ is the vector of regression coefficients and λ^* is the intensity of controls estimated applying kernel smoothing (see [34]). Let Y_j be equal to 1 or 0 according to whether the j-th household is a case or a control; under the assumption that both cases and controls are mutually independent IPP, conditionally to their location, we have

$$P\left(Y_j = 1|z_j\right) = \frac{\exp(\alpha + z_j{}'\beta)}{1 + \exp(\alpha + z_j{}'\beta)} \tag{3}$$

where z_j is the vector of spatial covariates on the location of the j-th household.

We consider the following spatial covariates:

- the overall population density, estimated using Tobler's pycnophylactic interpolation algorithm on census data [35]
- minimum rent cost per square meter for private residential properties

Data at hand come from the Population Register Office of Messina, as recorded on June 30, 2017. The dataset includes immigrants and their children/nephews born in their country of origin and who at birth acquired only the (foreign) citizenship of their parents. All addresses have been geocoded using the Google Maps Geocoding API. Minimum rent cost per square meter for private residential properties has been provided by the Italian Revenue Agency—OMI database (available at https://wwwt.agenziaentrate.gov.it/servizi/Consultazione/ricerca.htm). Computations have been executed using the *spatstat* package [34] for the R computing environment [36].

4 Results

Table 1 reports the distribution by nationality of foreign immigrant households in the city of Messina, corresponding to 5% of the global resident population (foreign citizens are individuals holding non-Italian citizenship and with habitual abode within Italy: https://www.tuttitalia.it/sicilia/38-messina/statistiche/cittadini-stranieri-2017/; accessed 23 Apr 2020). Asia is the prevalent continent of origin (59%), while the most numerous groups come from Sri Lanka, Philippines, and Romania. Foreign migrants satisfy the local labor demand for cheap, flexible, unqualified, and unprotected manpower: a type of work that even unemployed locals tend to avoid.[2] As several studies on other cities have demonstrated, a work specialization among migrants of different nationalities has developed also in Messina (e.g., [38, 39], Mucciardi et al. [40]) for the city of Catania; [41] for Palermo, [42] for Naples, [43] for a comparison of Sri Lankans' settlement patterns in the cities of Naples, Messina, Palermo and Catania). Sri Lankans, Filipinos, Romanians, and Poles are for the most part domestic workers in Italian families' residences. Romanians and Poles are almost entirely made up of women (see Table 1) who are mainly employed as caregivers for the elderly. Other groups coming from the Maghreb and sub-Saharan Africa, especially males coming from Morocco and Senegal, perceive housework as

[2] Individuals coming from non-EU countries have to acquire the visa. Once legally entered Italy, they have to apply for the residency permit for the same reason they received the visa. About 80% of the Italian visas is granted for tourism (after 3 months of legal permanence, the immigrants enter the world of illegality if they do not regularize their position). Business, family and subordinate work are the residuals motivations for visa [37, p. 46].

Table 1 Foreign immigrants residing in the Municipality of Messina on 01/01/2017 (*Source* Population Registry of the Municipality of Messina)

Country of origin	Males	Females	Total	% of total foreign immigrants (%)
Sri Lanka	2.072	1.824	3.896	32.9
Philippines	1.142	1.277	2.419	20.4
Romania	485	1.026	1.480	12.8
Morocco	712	420	1.132	10
Greece	335	59	394	3.3
People's Republic of China	189	193	382	3.2
Poland	50	259	309	2.7
Others			1835	14.7
Total			11.847	100

feminine and thus degrading, and prefer to work as peddlers. They usually sell counterfeit products, such as handbags, sunglasses, jewelry, and other seasonal items in the streets of the central shopping district (4th district), at open-air markets, at the main crossroads and, during summer, along the nearby beaches. Finally, the Chinese are mostly employed in import–export activities and retail. These specializations, considering the different logistic needs related to the occupational categories and considering the integration possibilities, have fostered the differentiation of the spatial distribution.

Figure 2 illustrates the distribution of foreign immigrant households in the city of Messina. Sub-figure (a) shows the smooth pycnophylactic interpolation of population counts for census tracts [35], while sub-figure (b) shows the minimum rent cost per square meter for private residential properties. As illustrated in Fig. 3, the six most numerous nationalities share a spatial trend, exhibiting high values in the central areas of the city (the central northern part of the maps) and low values in the peripheral areas. However, this spatial dynamics is stronger for the vendors' group than for the domestic workers' group. After estimating the intensities $\lambda(u)$ for each of the considered nationalities, we can estimate the values of the inhomogeneous K-function and compare the concentration level of the residential patterns for each nationality. The charts in Fig. 4 display the K-functions estimated using the intensities $\lambda(\mathrm{u}) = \lambda^*(\mathrm{u})\exp\{\alpha + \mathbf{z}(\mathrm{u})^j\beta\}$ (continuous blue lines) and the correspondent 95% confidence intervals under the null hypothesis of absence of interaction, based on 999 Monte Carlo simulations (dashed blue lines). The distances d are expressed in meters. Significant spatial concentration or dispersion emerges at the distances corresponding to the peaks of $\hat{K}_{inhom,i}(d)$ outside the envelopes. To highlight the relevance of the spatial covariates employed for the estimation of the intensities, the chart reports also the K-functions estimated by $\lambda(u) = \rho\lambda^*(u)$ (continuous black lines). Sri Lankans and Romanians exhibit significant spatial attraction. Conversely, the settlement models of Greeks and Poles prove to be coherent with a random allocation at all the considered distances. Further studies are required to evaluate the

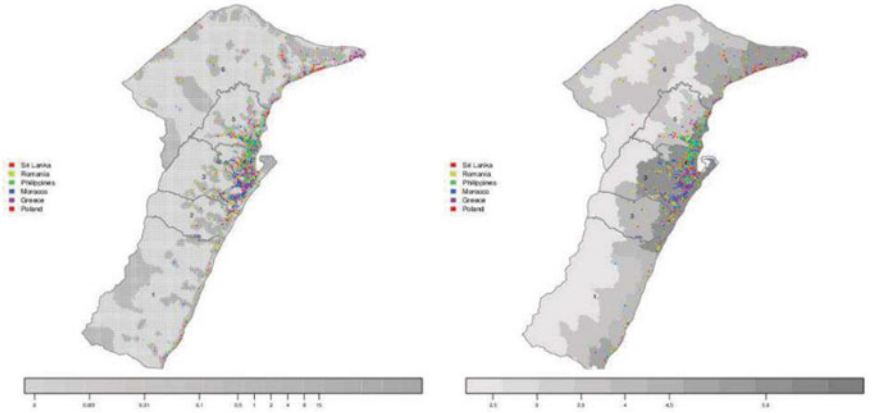


Fig. 2 Spatial distribution of foreign immigrant households in Messina

reasons underlying the different settlement patterns yielded by the various nationalities. Nevertheless, our results appear consistent with the implications of chain migration on the spatial attraction among foreign immigrants [44].

5 Conclusions

The picture emerging is that of a city that has become over time gateway to and from mainland Italy and hub of maritime traffics to some of the major Eurasian routes. An articulated infrastructural network, including the railway junction, the harbor areas, and the roadways to west and southwest, has profoundly shaped the urban arrangement of Messina. The historic reference facilities have been almost destroyed during the earthquake in 1908. The tragic event negatively affected the economic and demographic dynamics of the city since it fostered the development of the southern Etnean area and the region between Milazzo and Barcellona Pozzo di Gotto.

A significant concentration has been detected among the communities of Sri Lankans and Romanians, but also of Filipinos for distances greater than one kilometer. On the other end, the settlement patterns of Poles and Greeks appear compatible with random allocation. These results support the hypothesis of a link between chain migration and spatial attraction. A similar attraction results in clusters of foreign immigrant communities, especially within the 4th district, affecting the definition of a polycentric, heterogeneous urban identity. In addition to villages, hamlets, hill, and maritime centers that already shaped the morphology of Messina for a few centuries, clusters of foreign migrant communities currently add further elements of cultural and social pluralism. Our study corroborates the hypothesis that the most numerous foreign communities (Sri Lankans and Romanians) prefer to voluntarily segregate

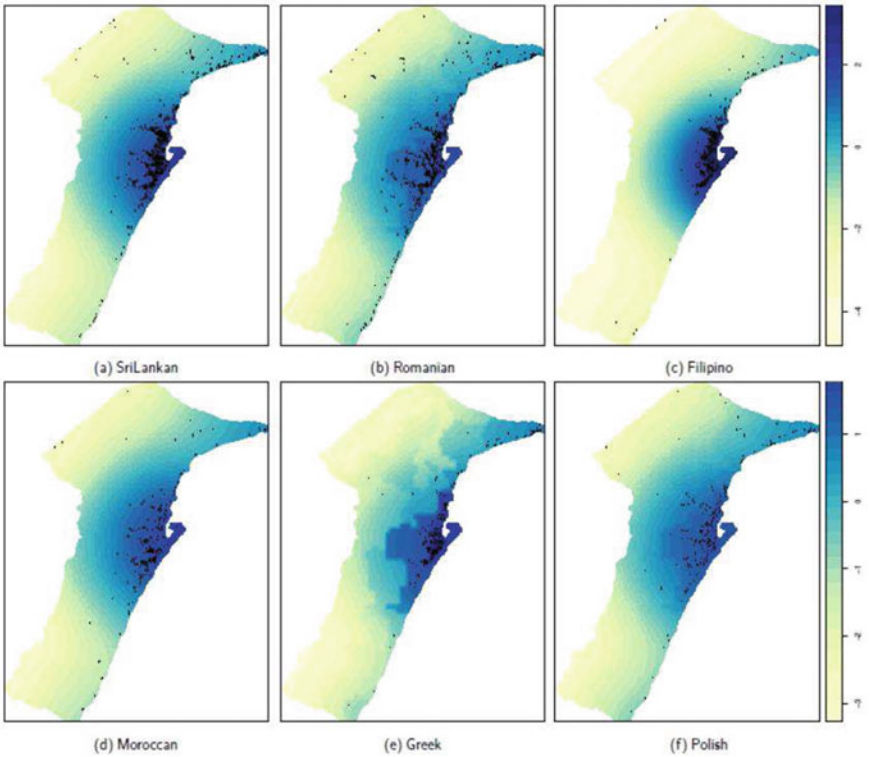


Fig. 3 Estimated spatial intensity $\widehat{\lambda}(u) = \lambda^*(u)\exp\left\{\alpha + z(u)'\beta\right\}$ for each nationality, in logarithmic scale (dots represent the household's location of the considered nationalities)

themselves within the central historic district (4th) because it represents their work area. They prefer indeed living in decrepit buildings, so to benefit from the proximity to the decent apartment blocks occupied by the middle class, where they as domestic workers or caregivers.

Messina is an intersection of stories and flows, open to the transit of travelers and goods carried by boats crossing the Strait. The city of the Strait lacks a plan aiming at realizing a new, cohesive version as a coordinating center of balanced metropolitan development. This is the reason why Messina passively suffers from its own flows and traffics instead of coordinating them. In this sense, the reinforcement of the role of Messina seems of primary interest: the real interlocutors of the city are not the neighboring hamlets and villages, but the city of Reggio Calabria to which the Sicilian city is particularly close also for the dialect having a typically Grecanic–Calabrian accent, rather than Sicilian.

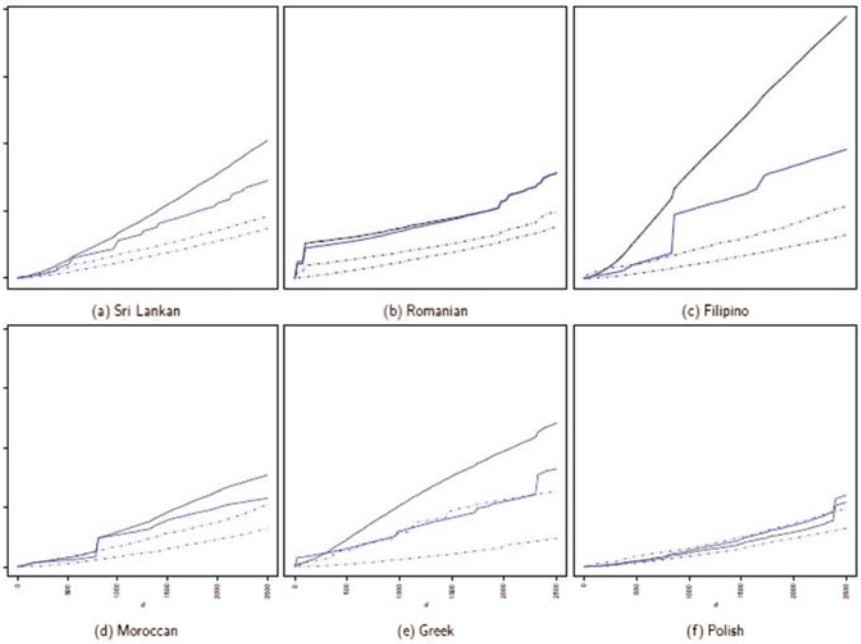


Fig. 4 Estimation of the inhomogeneous K-function

References

1. Ripley, B.D.: Spatial statistics. Wiley, New York (1981)
2. Campione, G.: Messina... qui comincia la Sicilia: "topografia della memoria". Edizioni Studium, Messina (2017)
3. Città di Messina: Messina Città Metropolitana. Connettere sogni—generare futuro. Messina (2015). https://www.cittametropolitana.me.it/in-evidenza/masterplan/
4. Gigante, A.I.: Lineamenti di morfologia urbana di Messina nel XVII secolo. In: Di Bella, S. (ed.), La rivolta di Messina (1674–1678) e il mondo mediterraneo nella seconda metà del Seicento. Proceeding of Convegno storico internazionale 10–12 ottobre 1975. Luigi Pellegrini Editore, pp. 445–456 (1979)
5. Provincia Regionale di Messina: Piano Territoriale Provinciale—PTP, approvato con delibera del consiglio provinciale n. 19 del 13/02/2008 (2008). http://www.provincia.messina.sitr.it/ptp.html
6. Galletta, F.: Messina, la persistenza del segno. Le tracce della città del 1908 dentro il piano di ricostruzione: il caso di via Risorgimento. In: Fiandaca, O., Lione, R. (eds.), Il sisma. Ricordare, prevenire, progettare. Alinea Editrice, Firenze, pp. 237–248 (2009)
7. Fumia, A.: Messina la capitale dimenticata. Addictions-Magenes Editoriale (2018)
8. Gigante, A.I.: Le città nella storia d'Italia, Messina. Laterza, Bari (1980)
9. Alibrandi, F., Salemi A (1988) Dall'Unità al terremoto. In: Giuseppe Campione (ed) Il progetto urbano di Messina: documenti per l'identità 1860–1988. Roma, Gangemi Editore, pp 93–198 (1988)
10. Battaglia, R.: Messina tra crisi e tentativi di rilancio dal "lungo Ottocento" al terremoto. In: Baglio, A., Bottari, S. (eds.), Messina dalla vigilia del terremoto del 1908 all'avvio della ricostruzione. Istituto di studi storici Gaetano Salvemini, Messina, pp. 57–70 (2010)

11. Campione, G., Puglisi, G., Callegari, P.: La furia di Poseidon. Messina 1908 e dintorni. Silvana Editoriale 1 **360** (2009)
12. Caminiti, L.: La grande diaspora: 28 dicembre 1908 la politica dei soccorsi tra carità e bilanci. GBM, Messina (2009)
13. Di Leo, L.G.: Messina, una città ricostruita. Materiali per lo studio di una realtà urbana. In: Lo Curzio, M. (ed.), Atti del seminario del 15/5/1982. Edizioni Dedalo, Bari (1985)
14. Gigante, A.I.: Messina. Storia della città tra processi urbani e rappresentazioni iconografiche. Laterza, Messina (2010)
15. Barone, G.: Stato, società e gerarchie urbane nel terremoto del 1908. In: Baglio, A., Bottari, S. (eds.), Messina dalla vigilia del terremoto del 1908 all'avvio della ricostruzione. Istituto di studi storici Gaetano Salvemini, Messina, pp. 411–424 (2010)
16. Casablanca, A.: Impatto ambientale e inquinamento a Messina. Armando Siciliano Editore, Messina (2001)
17. Scrofani, L. (2018). Le aree urbane nei processi di periferizzazione e di sviluppo del Mezzogiorno, in Sussidiarietà e… giovani al Sud. Rapporto sulla Sussidiarietà 2017/2018, Fondazione Sussidiarietà, Milano, pp. 167–198
18. Comune di Messina (2016) Messina dal 1861 ad oggi. L'evoluzione della popolazione in città attraverso i censimenti generali. Messina. Available via http://srvwebced.comune.messina.it/index.php
19. Cusumano, E.: The future of irregular migration to Europe. In: Villa, M. (ed.), The future of migration to Europe, Ledizioni LediPublishing, Milano (2020)
20. Comune di Messina (2018) Bilancio demografico comunale. Report del Servizio Statistica Dipartimento Servizi al Cittadino del Comune di Messina, Anno 2018
21. Saitta, P. (2013). Quota zero. Messina dopo il terremoto: la ricostruzione infinita. Donzelli, Roma
22. Zampieri, P.P.: Esplorazioni urbane. Urban art, patrimoni culturali e beni comuni. Rimozioni, implicazioni e prospettive della prima ricostruzione italiana (Messina 1908–2018). Il Mulino, Bologna (2018)
23. Freeman, L., Pilger, J., Alexander, W.: A Measure of Segregation Based Upon Spatial Arrangements. University of Hawaii (1971)
24. Schelling, T.C.: Dynamic models of segregation. J. Math. Sociol. **1**(2), 143–186 (1971)
25. Clark, W.A.V., Fossett, M.: Understanding the social context of the Schelling segregation model. Proc. Natl. Acad. Sci. **105**(11), 4109–4114 (2008)
26. Diggle, P.J.: Statistical analysis of spatial point patterns, 2nd edn. Edward Arnold, London (2003)
27. Almquist, Z.W., Butts, C.T.: Point process models for household distributions within small areal units. Demogr. Res. **S13**(22), 593–632 (2012)
28. Kumar, N.: Spatial sampling design for a demographic and health survey. Popul. Res. Policy Rev. **26**(5–6), 581–599 (2007)
29. Matthews, S.A., Parker, D.M.: Progress in spatial demography. Demogr. Res. **28**(10), 271–312 (2013)
30. Baddeley, A., Møller, J., Waagepetersen, R.: Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. Stat. Neerl. **54**(3), 329–350 (2000)
31. Harvey, D.: Geographic processes and the analysis of point patterns: testing models of diffusion by quadrat sampling. Trans. Inst. Br. Geogr. **40**, 81–95 (1966)
32. Mazza, A., Punzo, A.: Spatial attraction in migrants' settlement patterns in the city of Catania. Demogr. Res. **35**(5), 117–138 (2016)
33. Diggle, P.J., Gomez-Rubio, V., Brown, P.E., Chetwynd, A., Gooding, S.: Second-order analysis of inhomogeneous spatial point processes using case-control data. Biometrics **63**(2), 550–557 (2007)
34. Baddeley, A., Turner, R.: Spatstat: An R package for analyzing spatial point patterns. J. Stat. Soft. **12**(6), 1–42 (2005). http://www.jstatsoft.org/v12/i06/
35. Tobler, W.R.: Smooth pycnophylactic interpolation for geographical regions. J. Am. Stat. Assoc. **74**(367), 519–530 (1979)

36. R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna (2019). http://www.R-project.org/
37. Caritas e Migrantes: XXV Rapporto Immigrazione 2015. La cultura dell'incontro. Tau, Todi (2015)
38. Altavilla, A.M., Mazza, A.: On the analysis of immigrant settlement patterns using quadrat counts: the case of the city of Catania (Italy). Adv. Appl. Stat. **29**(2), 111–123 (2012)
39. Licciardello, O., Damigella, D.: Immigrants and natives: ways of constructing new neighbourhoods in Catania, Sicily. In: Armbruster, H., Meinhof, U.H. (eds.) Negotiating Multicultural Europe: Borders, Networks, Neighbourhoods, pp. 141–158. Palgrave Macmillan, London (2011)
40. Mucciardi, M., Altavilla, A., Mazza, A.: AM: Analysis of the residential pattern of foreign immigrants in Catania using GWR modelling. Rivista Italiana di Economia Demografia e Statistica **71**(2) (2017)
41. Busetta, A., Mazza, A., Stranges, M.: Residential segregation of foreigners: an analysis of the Italian city of Palermo. Genus **71**(2–3), 177–198 (2016)
42. Mazza, A., Gabrielli, G., Strozza, S.: Residential segregation of foreign immigrants in Naples. Spat. Demogr. **6**(1), 71–87 (2018)
43. Benassi, F., Bitonti, F., Mazza, A., Strozza, S.: Sri Lankans' residential segregation and spatial inequalities in Southern Italy: an empirical analysis using fine-scale data on regular lattice geographies. Qual. Quant. (2022). https://doi.org/10.1007/s11135-022-01434-5
44. MacDonald, J.S., MacDonald, L.D.: Chain migration, ethnic neighbourhood formation and social networks. Milbank Mem. Fund Q. **42**(1), 82–97 (1964)

Cultural Participation and Social Inequality in the Digital Age: A Multilevel Cross-National Analysis in Europe

Laura Bocci and Isabella Mingo

Abstract Cultural participation is considered as a necessary element of social equity, able to generate positive effects on individual opportunities and quality of life as a whole. Adopting a cross-national perspective, this study considers both traditional and new cultural practices deriving from the rise of new technologies, aiming at analyse how social inequality affects cultural participation in the European countries in the digital age. The main specific goals are the following: (1) elaboration of a synthetic index of Cultural Participation at European level; (2) identification of the determinants of participation at both individual and country level; (3) test of the interactions between some country features and individual characteristics, indicators of social differences, to verify their effects on cultural participation. The empirical analysis is based on 26,053 respondents aged 15 years and over, collected by the Special Eurobarometer survey n. 399 containing comparable data on cultural participation. To take into account country characteristics, some variables have been taken from other statistical sources (Eurostat). Data analysis resorted to Nonlinear Principal Component Analysis and multilevel regression models.

Keywords Cultural participation · Social inequality · Digital age · Multilevel analysis

1 Cultural Participation and Social Inequality in a Cross-National Perspective

The determinants of cultural participation have been broadly investigated since the work of [1]. Moreover, social inequality in cultural consumption has been studied by several scholars [2–5].

L. Bocci (✉)
Department of Social and Economic Sciences, Sapienza University of Rome, Rome, Italy
e-mail: laura.bocci@uniroma1.it

I. Mingo
Department of Communication and Social Research, Sapienza University of Rome, Rome, Italy
e-mail: isabella.mingo@uniroma1.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_6

However, in the current scenario in which new technologies have altered the consumption of cultural goods [12] and have allowed an increased diversification of leisure activities [9] while reducing public cultural expenditure in many European countries, it becomes crucial identifying both the characteristics of cultural consumers and old and new forms of social inequality. This information can be used by public agencies to inform about cultural policy decisions as well as by cultural professionals to encourage cultural participation.

Many studies from various countries confirm the link between an individual's socio-economic position and cultural behaviour, though adequate comparisons of cultural inequality between countries are still scarce [17]. In summarizing recent empirical literature on the determinants of cultural participation, [17] concludes that higher education, income, occupational prestige, occupational status and class position tend to be associated with more frequent cultural participation.

With respect to highbrow cultural consumption, studies of social inequality across countries often compare only small numbers of countries and employ separate regression models instead of more appropriate multi-level models [3, 15, 16, 20]. These comparative studies show that determinants of cultural consumption differ between countries [30].

Previous research on cross-national comparative patterns of association between the socio-demographic characteristics of individuals and their cultural participation has often relied on separate nationally representative surveys. Studies based on national surveys are typically difficult to compare because the survey questions are phrased differently, variables are operationalized differently, and samples are not always nationally representative [19, 25].

Cross-national comparative studies of cultural participation based on European-level data sets include works of [10, 26, 31]. More recent studies are those of: Van Hek and Kraaykamp [29], who analyse differences in cultural consumption of highbrow culture using data from the Eurobarometer 2007; Katz-Gerro and Lopez-Sintas [18], who examine individuals' tastes for music in several European countries with Eurobarometer 2001 data; Falk and Katz-Gerro [9], who use the European Union Statistics on Income and Living Conditions (EU-SILC) 2006 data to analyse the characteristics of visitors to museums and historical sites in 24 European countries; Perez-Villadoniga and Suarez-Fernandez [23], who use the EU-SILC 2005 data to study the determinants of cultural participation in four European countries, particularly focusing on education and income.

However, there are still open questions such as the extent to which the determinants of cultural participation differ across countries, considering also cultural practices deriving from the rise of new technologies, and what the underlying factors could be. More specifically, the cultural participation should be explored taking into account both individuals' characteristics and constraints and social context.

The aim of our research, therefore, is to address the cultural participation in a multidimensional view taking into account a broad spectrum of cultural activities and describing how both individual characteristics and contextual features affect it.

2 Outline of the Study

Cultural participation is not only a right enshrined in the United Nations Declaration of Human Rights, but it is also considered a key element for the quality of individual and collective life. This is widely recognized by both literature and international Agencies that have contributed to the definition of its concept and to the setting-up of some indicators for its empirical analysis [6, 8, 28].

At European level, the relevance of this issue is highlighted by the Agenda for Culture adopted in 2007 by the Council of the European Union and the European Council, as well as a number of policy actions set out in the Work Plan for Culture for 2015–2018, adopted by EU Culture Ministers in December 2014.

The vision of cultural participation as a phenomenon strongly linked to the development and to the well-being of advanced economies is increasingly widespread. Cultural participation is considered as a necessary element of equity, an ingredient of the development process, able to generate positive effects on individual opportunities, quality of life and wellness as a whole.

The concept of cultural participation can be made operative on the basis of the main international and European projects: Unesco's Framework for Cultural Statistics in 1986 [28], LEG (Leadership Group on Cultural Statistics) in 2000 [8] and ESSNet-Culture (European Statistical System Network on Culture) in 2011 by Eurostat [6]. These projects, albeit with some differences, agree to adopt a pragmatic definition based on the identification of the so-called cultural domains and to include cultural practices that fall into those domains without any distinction in terms of quality (highbrow/lowbrow) and including different types of participation.

Furthermore, these projects emphasize the changes in cultural practices deriving from the rise of ICT (Information and Communication Technology) and especially from the new possibilities offered by the Internet that make cultural participation more complex. Media users have more and more control over the selection of cultural contents offered via different channels, including mobile media. This is the *convergence culture* where patterns of media use are merging, moving from medium specific content toward content flowing across multiple media channels [14].

In this study, we consider a broad spectrum of cultural activities. The main goal is to deepening European country differences in cultural participation taking into account micro, meso and macro determinants.

Accordingly, three research questions are formulated:

(1) To what extent do European countries differ in cultural participation, and to what extent are these differences affected by a country's wealth, level of cultural funding and supply and spread of urbanization?
(2) To what extent are the differences in cultural participation between European countries affected by individual characteristics?
(3) What are the effects of country's characteristics on social differences in cultural participation?

To answer these research questions, a strategy of analysis in two steps is employed:

(1) a synthetic index of Cultural Participation is defined taking into account a broad spectrum of cultural activities;
(2) the differences in the determinants of cultural participation of European citizens across countries are analysed by using a multilevel approach.

The contribute is organized as follows. The dataset used is described in Sect. 3, while the methods of analysis are presented in Sect. 4. Section 5 discusses the main results. Lastly, Sect. 6 concludes.

3 Data

The data used in our analysis come from the Special Eurobarometer survey n. 399 [7] collected in the 28 member states of the European Union in 2013.

This survey detects the attitudes of the European public towards a range of cultural activities, looking at their participation as both consumers and performers of culture.

The dataset contains comparable data on cultural participation from 27,563 respondents. Representative stratified probability samples of about 1,000 respondents per country were collected using a face-to-face, computer-assisted interviewing mode. Eurobarometer surveys do not provide response rates.

For our analysis only 25 European countries, the ones with a sample size greater than 1,000, are considered. They are: France, Belgium, Netherlands, German, Italy, Denmark, Ireland, Great Britain, Greece, Spain, Portugal, Finland, Sweden, Austria, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, Slovenia, Bulgaria, Romania, Croatia. The sample has a total size of $N = 26{,}053$ individuals aged 15 years and over.

In order to address cultural participation in a multidimensional view, the choice of indicators was driven by the conviction that nowadays it is carried out through traditional activities as well as using the Internet.[1]

Therefore, the following 9 variables dealing with traditional cultural activities in the Eurobarometer survey are considered: 1-seen a ballet, a dance performance or an opera; 2-been to the cinema; 3-been to the theatre; 4-been to a concert; 5-visited a public library; 6-visited a historical monument or site (palaces, castles, churches, gardens, etc.); 7-visited a museum or gallery; 8-watched or listened to a cultural program on TV or on the radio; 9-read a book. Respondents indicated how many times in the past 12 months they had participated in these activities. Possible answers were: 1- not in the last 12 months; 2–1-5 times in the last 12 months; 3- more than 5 times in the last 12 months. Responses range from 1 (no participation) to 3 (very frequent participation).

Moreover, other 8 variables, which can supplement the previous ones, allow us to take into account some cultural activities performed by the Internet: 1-visiting

[1] The set of variables concern only with one form of cultural participation such as attending and receiving. The two other forms of amateur practice and social participation are not considered in this contribute.

museum or library websites or other specialized websites; 2-downloading movies, radio programs (podcasts) or TV programs; 3-watching streamed or on demand movies or TV programs; 4-reading newspaper articles online; 5-downloading music; 6-listening to radio or music; 7-reading or looking at cultural blogs; 8-searching for information on cultural products or events. Response modes to these variables were: 1-not mentioned; 2-mentioned.

To analyse how social inequality affects cultural participation, some individual-specific characteristics are considered. Since many studies indicates both individual's educational attainment and family status as possible influencing factors of cultural participation [2, 16], socio-demographic and cultural characteristics of people, along with some aspects of the community where they live, are included in the analysis. Specifically, the following variables are considered: gender (Male/Female), age (15–29; 30–49; 50–65; Over 65), educational attainment[2] (Still studying; 20 years and older; 16–19 years; Up to 15 years; No full-time education), labour market status (Self-employed, Employed, Not working), family social class (High level, Middle level, Low level), type of residence area (Large town, Small/middle town, Rural area).

Furthermore, the existing differences in cultural participation between European countries are of interest. Actually, it is reasonable to suppose that specific country characteristics significantly affect cultural participation. Therefore, the following macro features are also accounted for in the analysis: economic features in terms of country's level of wealth, cultural policies, cultural supply and level of urbanization.[3]

With regard to the economic features, the variable Gross Disposable Income of Households (source: Eurostat 2017) is considered as an indicator of purchasing power of households and their ability to invest in goods and services by accounting for taxes and social contributions too. This proxy of country's level of wealth, directly available for families, is assumed to have a positive effect on the levels of cultural participation. Therefore, we expect that the higher the level of family disposable income in a country, the higher an individual's level of cultural participation.

About the role of cultural policies, operationalized as the variable "Government Expenditure in Cultural Services" (source: Eurostat 2013), two different approaches can be found in the literature. The first one argues that public funding of culture expands the public's ability to access culture despite barriers related to socio-economic characteristics of individuals. Instead, the second approach emphasizes the reproduction of distinction by preserving inequality between social groups [9]. The results of our study could bring a new empiric element to settle these two contrasting positions.

[2] The variable "educational attainment" in the Eurobarometer is measured by the age at which a person has left full-time education by asking the following question: "How old were you when you stopped full-time education?".

[3] In addition to these variables (source: Eurostat), the following variables were also considered in some preliminary analyses: Government expenditure as % GDP in Education, Cultural enterprises as % of total of services, Tertiary education (levels 5–8) of the aged population 25–64 (%), Participation rate in education and training (%), Households Broadband coverage (%). Nonetheless, these variables resulted not significant.

As regards the cultural supply, its amount in a country is operationalized as the "Employment in the cultural sector" as the proportion of total employment (source: Eurostat 2013) because it can be assumed that the increase in cultural products and services requires an increase in employment in cultural sectors. A positive relationship between cultural supply and demand is expected, so that in countries where there is a larger cultural supply, there are higher levels of cultural participation too.

Finally, we control for the spread of urbanization in the countries using the proportion of Population in the cities (source: Eurostat 2013). The aim is to assess the existence of a positive relationship between urbanization and cultural participation, even in the presence of several ways to enjoy cultural products and services mediated by new technologies.

The interactions between these aforementioned country- and individual-specific characteristics are also tested to verify their effects on cultural participation.

4 Methods

To study cross-national differences in cultural participation, multilevel modelling is employed. With this method, it is possible to simultaneously estimate differences between countries and between individual respondents.

The selected variables described in Sect. 3 were the input for a strategy of analysis that consists of the following two steps: (1) elaboration of a synthetic index of Cultural Participation at European level; (2) identification of the determinants of participation at both individual and country level by using multilevel models with the synthetic index as dependent variable.

4.1 A Synthetic Index of Cultural Participation at European Level

In the definition of a synthetic index of Cultural Participation, all the variables dealing with cultural activities, including those carried out via the Internet, are taken into account. Since these variables are all categorical (both nominal and ordinal), Nonlinear Principal Component Analysis (NPCA) was applied to calculate the index.

As it is known, in NPCA, optimal quantification replaces the category labels with category quantifications in such a way that as much as possible of the variance in the quantified variables is accounted for. Specifically, the method maximizes the first p eigenvalues of the correlation matrix of the quantified variables, where p indicates the number of components chosen in the analysis. The aim of optimal quantification is the maximization of the Variance Accounted For (VAF) in the quantified variables [11, 22].

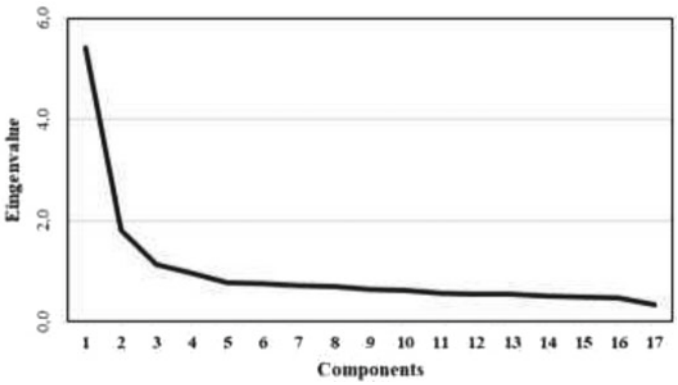


Fig. 1 NPCA scree plot

The scree criterion was used to decide the adequate number of components to be retained in the solution. In the present analysis, different dimensionalities consistently place the elbow at the second component, as shown in Fig. 1.

Therefore, at first, a solution with 2 dimensions was considered. However, inspection of the two-dimensional solution revealed that the second component was difficult to interpret, suggesting that this solution is not theoretically sensible. This lack of interpretability suggested that the mono-dimensional solution seems more appropriate.

The first dimension (alpha = 0.87, VAF = 32.12%) has positive correlations (between 0.5 and 0.7) with all variables (Table 1).

The choice of ordinal analysis levels for the variables has been evaluated by examining their transformation plots: for ordinal variables, they indicate that the categories are in the right order and the difference between categories 1 and 2 is slightly larger than that between categories 2 and 3 (Fig. 2). The transformation plots for nominal variables result in straight lines which indicate that these variables are linearly related to the other variables (Fig. 2).

The stability of results is confirmed according to the 95% bootstrap confidence regions for eigenvalues, component loadings, person scores and category quantifications [21] estimated using balanced bootstrap with 1,000 bootstrap samples (Table 1).

The first component can be interpreted as a Cultural Participation Synthetic (CPS) index. CPS index (ranging from −1.2 to 3.6; mean = 0; median = −0.19; standard deviation = 1; asymmetry = 0.72; kurtosis = −0.23) assumes different mean values by gender, age, education (Table 2) and country (Fig. 3). It is highest among women, decreases as the age increases, it steps up with higher levels of education and is quite different among European countries: its minimum pertains to Portugal (−0.59), the maximum to Sweden (0.90) with different distributions inside each country (Fig. 3).

Table 1 First component loadings and bootstrap estimations

Variable	Variable label	Loading	Bootstrap Mean	Confidence limit	
				Lower	Upper
RQB1_1	Seen a ballet, a dance performance or an opera	0.509	0.509	0.497	0.520
RQB1_2	Been to the cinema	0.618	0.618	0.609	0.626
RQB1_3	Been to the theatre	0.596	0.596	0.586	0.606
RQB1_4	Been to a concert	0.616	0.616	0.607	0.625
RQB1_5	Visited a public library	0.551	0.551	0.540	0.561
RQB1_6	Visited a historical monument or site	0.686	0.686	0.678	0.693
RQB1_7	Visited a museum/gallery	0.696	0.696	0.689	0.704
RQB1_8	Watched or listened to a cultural program on TV or on the radio	0.480	0.480	0.469	0.490
RQB1_9	Read a book	0.586	0.586	0.578	0.595
RBQ6_1	Internet for cultural purposes: visiting museum or library websites	0.540	0.540	0.529	0.551
RBQ6_3	Internet for cultural purposes: downloading movies, radio programs (podcasts) or TV programs	0.470	0.470	0.457	0.482
RBQ6_4	Internet for cultural purposes: watching streamed or on demand movies or TV	0.501	0.501	0.490	0.513
RBQ6_5	Internet for cultural purposes: reading newspaper articles online	0.606	0.606	0.597	0.615
RBQ6_7	Internet for cultural purposes: downloading music	0.450	0.450	0.437	0.462
RBQ6_8	Internet for cultural purposes: listening radio/music	0.568	0.568	0.557	0.578
RBQ6_9	Internet for cultural purposes: reading or looking at cultural blogs	0.459	0.459	0.446	0.472

(continued)

Table 1 (continued)

Variable	Variable label	Loading	Bootstrap Mean	Confidence limit	
				Lower	Upper
RBQ6_12	Internet for cultural purposes: searching for information on cultural products or events	0.621	0.621	0.612	0.629

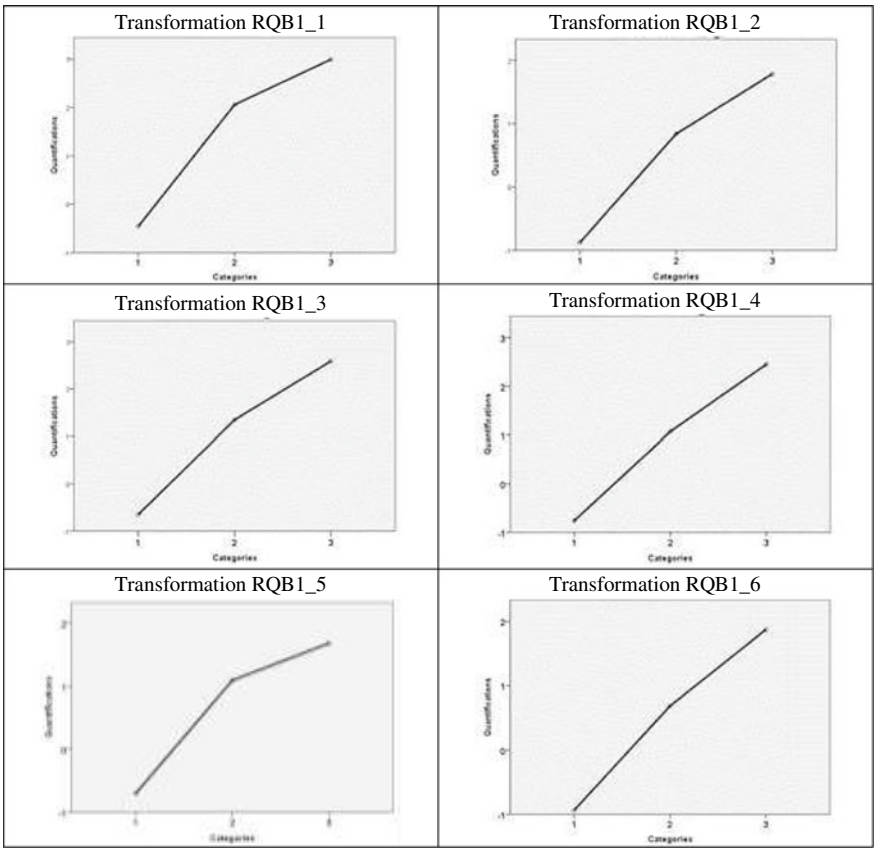


Fig. 2 NPCA Transformation Plots – ordinal and nominal levels (the variable labels are explained in Table 1)

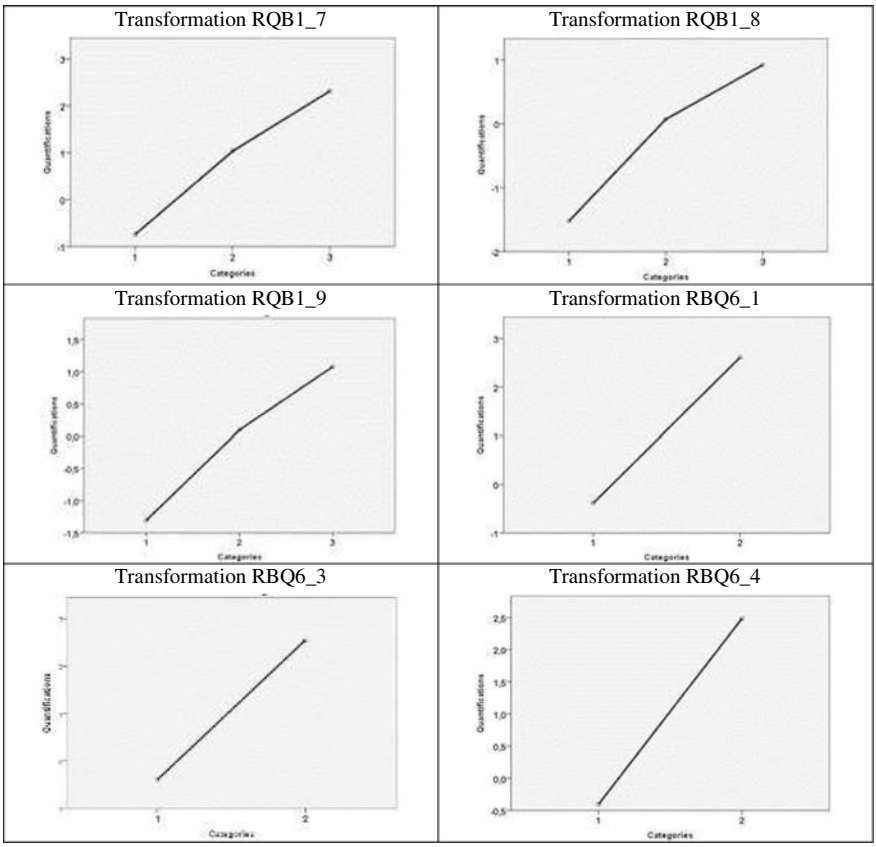


Fig. 2 (continued)

4.2 The Determinants of Cultural Participation in the European Countries: A Multilevel Approach

The structure of the data to be analysed is clearly hierarchical, since individuals are nested within countries. Given the intrinsically hierarchical nature of the data set and hypothesizing that the variability of the CPS index can depend on both individual characteristics and the different contexts in which people live, a multilevel approach is used [24].

In our case, data consist of the values of CPS index (dependent variable) and several explanatory variables, both social–demographic individual features and country variables, referred to the i-th respondent in the j-th country ($i = 1, \ldots, N_j$, $j = 1, \ldots, 25$ and $\sum_j N_j = 26,053$). Therefore, there are two levels of analysis: level two, the highest, is that of countries, and level one, the lowest and nested within the upper level, is that of individuals.

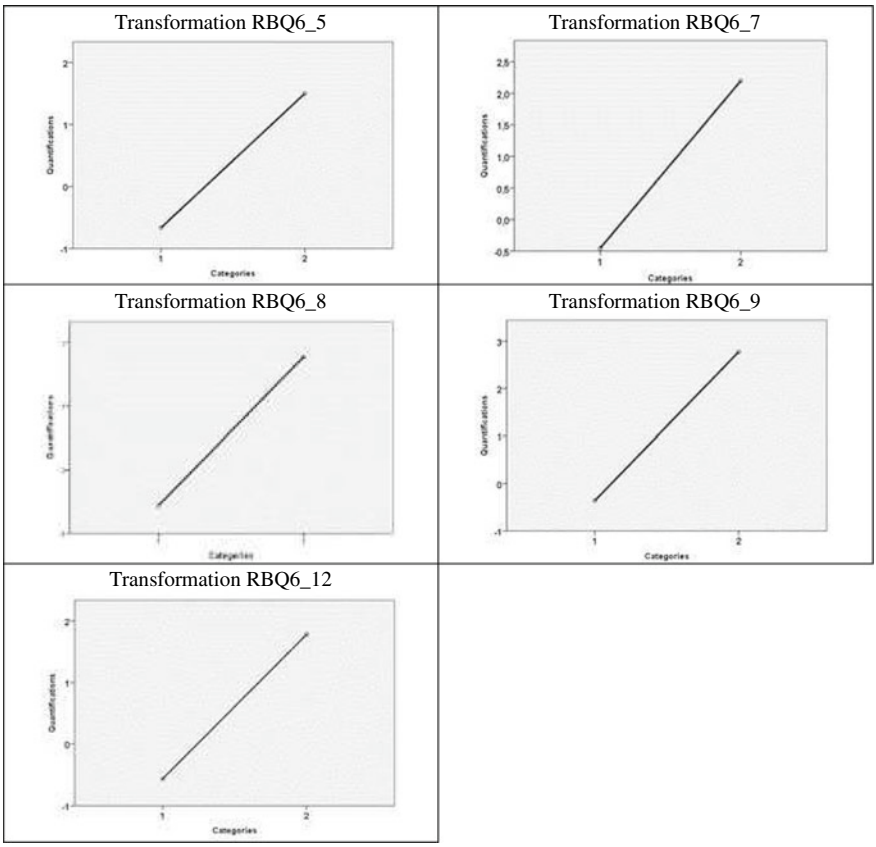


Fig. 2 (continued)

Since it is reasonable to assume that countries can have a systematic effect on the Cultural Participation of individuals, CPS index values within the same country are dependent or correlated.

In this context a multilevel analysis with no explanatory variables at all, the so called *random intercept-only model*, was firstly applied

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij}, \quad i = 1, \ldots, N_j \quad \text{and} \quad j = 1, \ldots, 25, \tag{1}$$

where Y_{ij} is the value of CPS index on individual i in country j, e_{ij} represents some individual-dependent residual, while u_{0j} is a random country-dependent deviation.

Model (1) provides a partitioning of the variance between the first (e_{ij}) and the second (u_{0j}) level (individual- and country-level variance, respectively) and allows us to evaluate the Intraclass Correlation Coefficient (ICC) for the country effect. ICC can be considered both a measure of the variability between countries and the degree of the non-independence of individuals nested into countries. In model (1), intercept

Table 2 Cultural Participation Synthetic (CPS) index by people categories

		Mean	N	Standard deviation
Gender	Male	−0.025	11830	1.002
	Female	0.021	14223	0.998
Age	15–29	0.397	4650	1.009
	30–49	0.129	8660	0.997
	50–65	−0.141	7209	0.962
	Over 65	−0.352	5534	0.885
Educational attainment	Still studying	0.755	1978	0.956
	20 years and older	0.547	7943	0.996
	16–19 years	−0.220	11260	0.837
	Up to 15 years	−0.740	4298	0.577
	No full-time education	−0.864	196	0.665
Total		0.000	26053	1.000

All differences between group means are significant (Welch test $p < 0.001$)

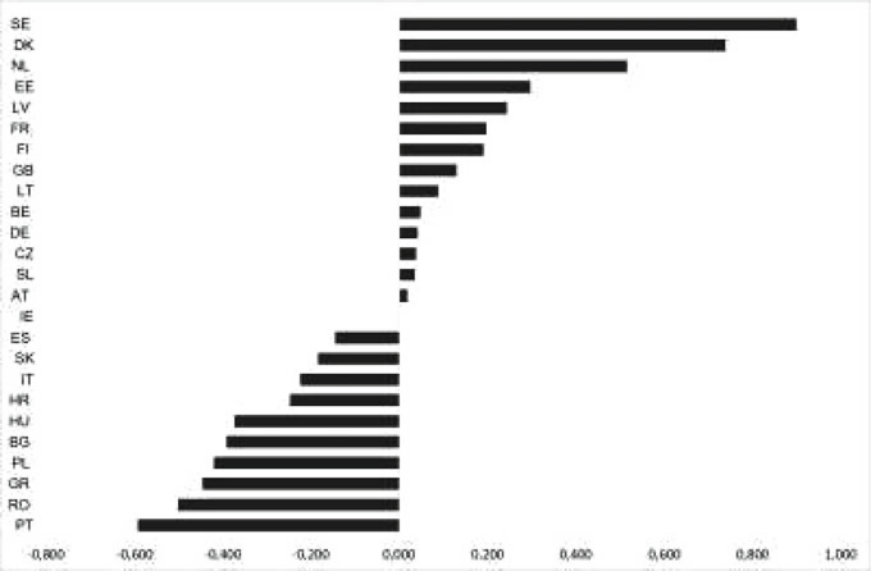


Fig. 3 CPS index by country

varies significantly between countries ($Wald\ Z = 3.44$, $p-value = 0.001$; LR test vs. linear regression = 3435.50, one-tailed $p - value = 0.0000$) and ICC = 0.136 (Table 3) shows that about 13.6% of the CPS index variability is due to the variability between countries.

Table 3 Model results: Individual-level variance, Country-level variance and fit statistics

	Model 1		Model 2		Model 3		Model 4	
	Estimate	Sig.	Estimate	Sig.	Estimate	Sig.	Estimate	Sig.
Individual-level variance	0.872	0.000	0.618	0.000	0.618	0.000	0.617	0.000
Country-level variance	0.137	0.001	0.079	0.001	0.020	0.001	0.019	0.002
ICC	0.136	–	0.113	–	0.031	–	0.029	–
R^2 I level			0.291		0.291		0.292	
R^2 II level			0.423		0.854		0.861	
Fit statistic								
Akaike Corrected (AICC)	70507.10		59427.70		59410.96		59396.57	
Bayesian (BIC)	70523.43		59443.96		59427.22		59412.83	

Since these previous results justify a multilevel approach [13, 27], an analysis in three steps was then performed.

At first, $T = 6$ level-one (individual) explanatory variables X_t ($t = 1,...,T$) were introduced in the multilevel model

$$Y_{ij} = \gamma_{00} + \sum_{t=1}^{T} \gamma_{t0} X_{tij} + u_{0j} + e_{ij}, \quad i = 1, \ldots, N_j \text{ and } j = 1, \ldots, 25, \tag{2}$$

and then $P = 4$ level-two (country) explanatory variables Z_p ($p = 1,...,P$) were also put in defining the following model

$$Y_{ij} = \gamma_{00} + \sum_{t=1}^{T} \gamma_{t0} X_{tij} + \sum_{p=1}^{P} \gamma_{0p} Z_{pj} + u_{0j} + e_{ij},$$
$$i = 1, \ldots, N_j \text{ and } j = 1, \ldots, 25. \tag{3}$$

Both individual- and country-level variables included in models (2) and (3) are those introduced in Sect. 3. Specifically, the $T = 6$ individual variables X_t are: gender, age, educational attainment, labour market status, family social class, type of residence area; while the $P = 4$ country variables are: Gross Disposable Income of Households (GDI), Government Expenditure in Cultural Services as % GDP (GECS), Employment in the Cultural Sector as % of total employment (ECS), Distribution of Population in the Cities (DPC).

In model (2), taking into account only the individual covariates, the variability between countries is still high (ICC = 0.113), even if less than that obtained from model (1), so there is still a fair amount of variation (11.3%) across countries that can

be explained by level-two covariates. Actually, in model (3) ICC drastically decreases to 0.031 (Table 3), a value below the cut-off of 0.05 set by most researchers [13, 27].

Lastly, the cross-level interaction terms were also included in the final multilevel model

$$Y_{ij} = \gamma_{00} + \sum_{t=1}^{T} \gamma_{t0} X_{tij} + \sum_{p=1}^{P} \gamma_{0p} Z_{pj} + \sum_{p=1}^{P} \sum_{t=1}^{T} \gamma_{tp} X_{tij} Z_{pj} + u_{0j} + e_{ij},$$
$$i = 1, \ldots, N_j \quad \text{and} \quad j = 1, \ldots, 25. \tag{4}$$

In order to test how country features affect the relationship between individual characteristics and cultural participation, in model (4) some significant cross-level interactions are considered: that between GDI and educational attainment, and that between cultural funding and family social class.[4]

5 Results

The results from model (4), which is the best of the four models according to ICC, AICC and BIC (Table 3), show that all predictors are significant.

Compared to the reference value, CPS index increases in the youngest aged 15–29 people, who are still studying, are self-employed, live in cities and belong to high social class (Table 4).

Nonetheless, the results show that differences at individual level are not sufficient to explain all the determinants of Cultural Participation. It is necessary to take into account some characteristics of the countries and some cross-level interactions: all the predictors – economic, cultural, and political – included in model (4) have a positive influence on the CPS index (Table 4).

Analyzing Table 4, at individual level, the results show that women participate more than men in culture activities and the elderly less than the young. As expected, the effects of educational attainment are large and positive: higher educated individuals in Europe have a higher probability to participate in culture activities, but the students are the most active. The effects of family affluence are also significant: participation increases with the level of family social class. Regarding the labour market status, it seems that self-employed participate more to culture activities as compared to the employed and not working people.

Interestingly, it is shown that community size matters: level of cultural participation is higher for the respondents living in large cities than those living in rural area or middle towns.

Considering the influence of the country characteristics, the results show that the level of cultural supply in a country (ECS) significantly and positively affects

[4] In some preliminary analyses, all the cross-level interactions between country variables and individual variables were tested, but only those introduced in model (4) resulted significant.

Table 4 Estimates for random intercept Model (4): fixed effects and cross-level interactions

Model term		Coefficient	Std. Error	t	Sig
Intercept		−1.056	0.0755	−13.985	0.000
Gender (reference: Female)	Male	−0.115	0.0189	−6.062	0.000
Age (reference: Over 65)	15–29	0.293	0.0407	7.198	0.000
	30–49	0.214	0.0284	7.536	0.000
	50–65	0.079	0.0205	3.843	0.000
Educational attainment (reference: No full-time education)	Still studying	1.282	0.0866	14.803	0.000
	20 years and older	0.956	0.0782	12.231	0.000
	16–19 years	0.406	0.0768	5.284	0.000
	Up to 15 years	0.102	0.0686	1.481	0.139
Labour Market Status (reference: Not working)	Self-employed	0.305	0.0325	9.396	0.000
	Employed	0.218	0.0199	10.947	0.000
Residence Area (reference: Rural area)	Large town	0.263	0.0253	10.395	0.000
	Small/middle town	0.102	0.0186	5.483	0.000
Family Social Class (reference: Low level)	High level	0.329	0.0257	12.777	0.000
	Middle level	0.159	0.0141	11.271	0.000
GDI (Gross Disposable Income)		0.167	0.0495	3.386	0.001
GECS (Government Expenditure in Cultural Service % GDP)		0.074	0.0208	3.539	0.000
ECS (Employment in cultural sector as % of total employment)		0.131	0.0356	3.689	0.000
DPC (Distribution of population in the cities (%))		0.049	0.0220	2.229	0.026
Family Class* GECS (reference: Low)	High level	0.066	0.0147	4.480	0.000
	Middle level	0.012	0.0076	1.597	0.110
Educational attainment * GDP (reference: No full-time education)	Still studying	−0.142	0.0570	−2.494	0.013
	20 years and older	−0.034	0.0528	−0.651	0.515
	16–19 years	−0.053	0.0481	−1.107	0.268
	Up to 15 years	−0.075	0.0491	−1.524	0.127

cultural participation, more than national wealth (GDI), cultural policies (GECS) and the spread of urbanization in the country (DPC).

Referring to whether the effects of family affluence are affected by country cultural policies, cross-level interactions were tested. In Table 4 the results show that in the countries that invest in cultural sectors, the effect of family affluence on cultural participation is much higher than in countries with low levels of cultural funding:

this result seems to corroborate the hypothesis of the persistence of social inequalities. Instead, since the interaction between GDI and educational attainment has negative coefficients, there seems to be a rebalancing effect of inequalities due to education in countries where household disposable income is higher.

The general importance of the selected country characteristics for differentiation in cultural participation is illustrated by the variance explained at the country level: it increases (from 42.3 to 85.4%) when the four country variables are added to the model. Moreover, it rises to 86.1% when cross-level interactions were considered (Table 3).

6 Conclusions

In this study we address the extent to which the determinants of cultural participation differ across European countries taking into account that nowadays cultural practice is carried out both through traditional activities and Internet.

To answer the three research questions, at first a multidimensional vision of cultural participation was adopted calculating a synthetic index of Cultural Participation which takes into account all the cultural activities detected in the Special Eurobarometer survey n. 399 [7].

Afterwards, multilevel models on Eurobarometer data were estimated to analyse the determinants of cultural participation of European citizens deepening the direction of the relationships with socio-demographic individual characteristics and relevant country features.

In addition to looking at the association between certain country characteristics and cultural participation, the so-called cross-level interactions were studied relating the Gross Disposable Income of Households at country level to an individual's educational attainment and the level of country's cultural funding to family social class of individuals.

The results of the study help to understand the differences in the influence of different factors in shaping individuals' participation in cultural activities.

With respect to the research questions, the results confirm that cultural participation differs between European countries.

At macro level, the empirical findings confirm the hypothesized relationship: country's level of wealth, cultural policies, cultural supply and level of urbanization significantly and positively contribute to the explanation of variation in cultural participation. Nonetheless, among these features the ones that more influence the cross-national differences are related to the country's level of cultural supply and the level of wealth.

Moreover, also considering the new forms of cultural activities carried out through the new technologies, the results endorse the influence of individual characteristics on the differences in cultural participation between European countries consistently with literature studies (see for example, 17).

The innovative part of our study refers to the effects of country's characteristics on social differences in cultural participation. The cross-level interactions referring to the influence of country characteristics on the effects of education and family social class highlight that while the country's level of cultural funding strengthens social inequalities [9], the country's level of wealth, in terms of disposable income for families, has a rebalancing effect on social inequalities.

Finally, even if the hybrid use of cultural contents, both through traditional media and new technologies, suggests considering the various activities jointly, it might be worthwhile to carry out separate analyses to better highlight the different relationships with some individual and contextual characteristics.

References

1. Baumol, W.J., Bowen, W.G.: Performing arts-the economic dilemma: a study of problems common to theatre, opera, music and dance. Twentieth Century Fund. (1966)
2. Bennett, T., Savage, M., Silva, E., Warde, A., Gayo-Cal, M., Wright, D.: Culture, Class, Distinction. Routledge, London (2009)
3. Chan, T.W., Goldthorpe, J.: Social Status and Cultural Consumption. Cambridge University Press, Cambridge, UK (2010)
4. DiMaggio, P.: Cultural capital and school success: high-school students. Am. Sociol. Rev. **47**(2), 189–201 (1982)
5. Erickson, B.H.: Culture, class, and connections. Am. J. Sociol. **102**, 217–251 (1996)
6. ESSnet Project: ESSnet Culture Final Report. Luxembourg (2011)
7. European Commission. Special Eurobarometer 399. Cultural access and participation Report (2013)
8. Eurostat. Cultural Statistics in the EU, «Final report of the LeG» – Population and social conditions. Working papers 3/2000/E/n.1. European Commission, Luxemburg (2000)
9. Falk, M., Katz-Gerro, T.: Cultural participation in Europe: can we identify common determinants? J. Cult. Econ. **40**, 127–162 (2016)
10. Gerhards, J.: Die Kulturell Dominierende Klasse in Europa. Eine Vergleichende Analyse der 27 Mitgliedsländer der Europäischen Union im Anschluss an Pierre Bourdieu. Kölner Zeitschrift für Soziologie und Sozialpsychologie **60**, 723–748 (2008)
11. Gifi, A.: Nonlinear Multivariate Analysis. Wiley, New York (1999)
12. Handke, C., Stepan, P., Towse, R.: Cultural economics, the Internet and Participation. In: Ateca-Amestoy, V.M., Ginsburgh, V., Mazza, I., O'Hagan, J., Prieto-Rodriguez, J. (eds.) Enhancing Participation in the Arts in the EU: Challenges and Methods, pp. 295–310. Springer International Publishing, Cham (2017)
13. Hox, J.: Multilevel Analysis: Techniques and Applications, Routledge (2010)
14. Jenkins, H.: Convergence Culture: Where Old and New Media Collide. New York University Press, New York, London (2006)
15. Katz-Gerro, T.: Highbrow cultural consumption and class distinction in Italy, Israel, West Germany, Sweden, and the United States. Soc. Forces **81**(1), 207–229 (2002)
16. Katz-Gerro, T.: Comparative evidence of inequality in cultural preferences: gender, class, and family status. Sociol. Spectr. **26**(1), 63–83 (2006)
17. Katz-Gerro, T.: Cross-national cultural consumption research: inspirations and disillusions. Kölner Zeitschrift für Soziologie und Sozialpsychologie **51**, 339–360 (2011)
18. Katz-Gerro, T., Lopez-Sintas, J.: The breadth of Europeans' musical tastes: disentangling individual and country effects. Adv. Sociol. Res. **14**, 97–122 (2013)

19. Kawashima, N.: Comparing cultural policy: Towards the development of comparative study. Int. J. Cult. Policy. **1**(2), 289–307 (1995)
20. Kraaykamp, G., Nieuwbeerta, P.: Parental background and lifestyle differentiation in Eastern Europe: social, political, and cultural intergenerational transmission in five former socialist societies. Soc. Sci. Res. **29**(1), 92–122 (2000)
21. Linting, M., Van der Kooij, A.J.: Nonlinear principal components analysis with CATPCA: A tutorial. J. Pers. Assess. **94**(1), 12–25 (2012)
22. Meulman, J.J., Van Der Kooij, A.J., Heiser, W.J.: Principal components analysis with nonlinear optimal scaling transformations for ordinal and nominal data. The Sage Handbook of Quantitative Methodology for the Social Sciences. Elsevier Science B.V., Amsterdam (2003)
23. Perez-Villadoniga, M., Suarez-Fernandez, S.: Education, income and cultural participation across Europe. Cuadernos económicos de ICE **98**, 89–103 (2019)
24. Raudenbush, S.W., Bryk, A.S.: Hierarchical Linear Models: Applications and Data Analysis Methods, vol. 1. Sage Publications, Thousand Oaks, London, UK (2002)
25. Schuster, M.J.: Participation studies and cross-national comparison: proliferation, prudence, and possibility in the arts and culture. Cult. Trends. **16**, 99–196 (2007)
26. Sisto, A., Zanola, R.: Cinema attendance in Europe. Appl. Econ. Lett. **17**, 515–517 (2010)
27. Snijder, T., Bosker, R.: Multilevel Analysis. An introduction to basic and advanced modelling. SAGE Publications (1999)
28. Unesco: Framework for Culture Statistics. Paris (1986)
29. Van Hek, M., Kraaykamp, G.: Cultural consumption across countries: a multi-level analysis of social inequality in highbrow culture in Europe. Poetics **41**(4), 323–341 (2013)
30. Virtanen, T.: Dimensions of taste for cultural consumption: across-cultural study on young Europeans. In: Proceedings of the 8th AIMAC Conference, July 3–6, HEC Montreal, Canada (2005)
31. Virtanen, T.: Across and beyond the bounds of taste: On cultural consumption patterns in the European Union. Turku: Turku School of Economics (2007)

Reducing Bias of the Matching Estimator of Treatment Effect in a Nonexperimental Evaluation Procedure

Maria Gabriella Campolo, Antonino Di Pino Incognito, and Edoardo Otranto

Abstract The traditional matching methods for the estimation of treatment parameters are often affected by selectivity bias due to the endogenous joint influence of latent factors on the assignment to treatment and on the outcome, especially in a cross-sectional framework. In this study, we show that the influence of unobserved factors involves a cross-correlation between the endogenous components of propensity scores and causal effects. We propose a correction for the bias effect of this correlation on matching results, adopting a state-space model to identify and estimate the unobserved factors. A Monte Carlo experiment supports this finding.

Keywords Bias in programme evaluation · Endogenous treatment · Propensity score matching · State-space model

1 Introduction

Matching is one of the most widespread methods to infer the causal effects of a treatment using observational data, but several authors have underlined some drawbacks in terms of statistical properties (see, for example, [6, 8, 12]). In particular, cross-sectional methods for estimating causal effects are affected by a considerable bias due to the endogeneity of the treatment. Namely, the decision of a subject to undergo the treatment is endogenous with respect to the potential outcome. As a consequence, matching methods that identify the decision of a subject to undergo treatment using

M. G. Campolo · A. Di Pino Incognito (✉) · E. Otranto
Dipartimento di Economia, University of Messina, Messina, Italy
e-mail: dipino@unime.it

M. G. Campolo
e-mail: mgcampolo@unime.it

E. Otranto
e-mail: eotranto@unime.it

E. Otranto
Centre for North South Economic Research (CRENOS), University of Cagliari, University of Sassari, Cagliari, Italy

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_7

observed covariates could fail because unobservable factors endogenously influence both the propensity of a subject to undergo treatment and the outcomes [9].

There are not standard (and simple) solutions for this problem, especially because the choice of the treatment status by the subject rarely reproduces the typical conditions of a randomized experiment. To approximate these results, most estimation methods of treatment effect impose limitations on the sample and/or adopt exclusion restrictions, including, for example, in the sample only the subjects who, on the basis of their observed characteristics, are identified "at the margin" between the two treatment statuses (e.g. [10]).

In this study we suggest that it is possible, with a propensity score matching approach in a cross-sectional framework, to reduce the bias due to endogeneity not imposing these constraints. We assume that the potentially omitted endogenous factors can be represented by a stochastic component of the propensity score correlated with the causal effects of the treatment. This implies that the causal effect of each subject is correlated with the causal effect of another subject with a similar propensity score; moreover the stochastic component is autocorrelated, as causal effects relative to similar propensity scores will be more similar. For this purpose we adopt a sort of state-space model (see, for example, [4]), where common latent factors are detected in correspondence to the endogenous stochastic component of the propensity score sorted in an ascending (or descending) order. The predictions of these components are used as correction terms in the matching procedure.

We compare the performance of the proposed method with respect to the traditional Propensity Score Matching Estimator (*PSME*, cf., among others, [8, 14]) by Monte Carlo experiments.

The paper is structured as follows. In the next section we describe the new procedure and introduce the new estimator, showing its properties in terms of reduction of selection bias, and comparing its characteristics to those of the more common matching procedures based on propensity score. Section 3 shows the results of the Monte Carlo experiments, while, Sect. 4 concludes with some final remarks.

2 Model Specification

The solution we propose to correct the matching estimator is based on two innovative aspects. The most important is the individuation of an autoregressive process that characterizes, jointly, individual propensity scores and causal effects. Another important novelty, strictly linked to the previous one, is given by the correction term derived as the state variable of a State-Space model that identifies the endogenous components of the causal effects.

To better explain the endogenous relationship between causal effect and propensity to undergo treatment, we start to consider the potential outcome gained by choosing one of the two treatment statuses as a relevant (endogenous) determinant of the decision to undergo treatment. In particular, we specify the model assuming

that the difference between the expected outcomes, y_{1i} and y_{0i}, obtainable, respectively, under the regimes $T_i = 1$ (if the subject belongs to the treatment group) and $T_i = 0$ (if the subject belongs to the comparison group), determines, at least in part, the choice of the regime.

Let us consider a Probit (or Logit) model, where the (latent) propensity to undergo treatment of the i-th subject, T_i^*, depends linearly on the covariates in $\mathbf{Z}$:

$$T_i^* = \mathbf{z}_i' \boldsymbol{\beta} + v_i \tag{1}$$

where $\mathbf{z}_i'$ is the i-th row of the matrix $\mathbf{Z}$, $\boldsymbol{\beta}$ is a vector of unknown coefficients and v_i is a zero mean random disturbance with unit variance. If $T_i^* > 0$, $T_i = 1$ (the subject underwent treatment), otherwise $T_i = 0$ (the subject did not undergo treatment).

As a consequence of the endogeneity of T_i^* with respect to the causal effects, $\Delta_i = y_{1i} - y_{0i}$, we can suppose that the propensities T_i^*, sorted in ascending or descending order, are autocorrelated, and the same holds for the causal effects Δ_i. In practice, subjects i and $i + 1$, with contiguous propensities T_i^* and T_{i+1}^*, show similar causal effects, Δ_i and Δ_{i+1}. We specify our model according to the "Two-Regime" Roy Model (e.g., [3, 7]), adding to the selection Eq. (1) two further equations:

$$y_{1i} = \mu_{1i} + u_{1i} \qquad \text{if } T_i = 1; \quad \text{otherwise latent} \tag{2a}$$

$$y_{0i} = \mu_{0i} + u_{0i} \qquad \text{if } T_i = 0; \quad \text{otherwise latent} \tag{2b}$$

In Eqs. (2a and 2b), μ_{1i} and μ_{0i} are the outcomes obtained by treated and untreated subjects, respectively, depending on the decision to undergo treatment ($T_i = 1$) or not ($T_i = 0$). The error terms u_{1i} and u_{0i} are normally distributed with zero mean and variances equal to σ_1^2 and σ_0^2 respectively. The covariances σ_{1v} and σ_{0v} of the disturbances of both outcome equations, u_{1i} and u_{0i}, with the disturbances of the selection equation (Eq. 1), v_i, are measurements of the endogeneity of the propensity to undergo treatment, T_i^*, with respect to the outcome gained under $T = 1$ and $T = 0$.

Correlation between outcomes and propensity scores, as well as the autocorrelation of the causal effects, may be specified starting from the definition of causal effects, Δ_i, obtained as:

$$\Delta_i = y_{1i} - y_{0i} = \mu_{1i} - \mu_{0i} + (u_{1i} - u_{0i}) \tag{3}$$

We suppose that u_{1i} and u_{0i} are both linearly related to v_i, involving a certain degree of endogeneity. Formally we have:

$$u_{1i} - u_{0i} = \sigma v_i + \varepsilon_i \tag{4}$$

Putting $\mu_{1i} - \mu_{0i} = \mu_i$, Eq. (3) can be written as a measurement equation of a State-Space Model, as follows (cf., among others, [4]):

$$\Delta_i - \mu_i = \sigma v_i + \varepsilon_i \tag{5}$$

In Eq. (5), ε_i is a vector of $n \times 1$ disturbance terms uncorrelated across i. The variable v_i can be considered as the state variable whose elements are not observable, but are assumed to be generated by a first-order Markov process (*transition equation*):

$$v_i = \rho v_{i-1} + \eta_i \tag{6}$$

where η_i is a vector of independent disturbances with zero mean. The dependent variable of Eq. (5), $\Delta_i - \mu_i$, represents the stochastic component of the causal effect Δ_i, endogenous with respect to the decision to undergo treatment. Starting from this result, the selectivity effect due to the endogeneity of the decision to undergo treatment may be corrected by estimating σv_i in Eq. (5), and using the corresponding predicted values, $\sigma \hat{v}_i$, as a correction term in the matching estimation of the causal effects. In doing this, a preliminary estimation of causal effects Δ_i is obtained at a first stage by applying a traditional propensity score matching procedure. Then, at a second stage, matching is replicated using the corrected outcomes $y_{1i} - \sigma \hat{v}_i$ so as to obtain the corrected causal effects $\Delta_i - \sigma \hat{v}_i = \hat{\mu}_i$. We call this estimator the State-Space Corrected Matching (*SSCM*) estimator.

2.1 Assessing Bias Evaluation

In this study we provide an empirical comparison between *SSCM* and *PSME* estimator in order to ascertain the respective effectiveness in reducing bias in the estimation of treatment parameters. However, an important aspect to discuss is given by the nature of bias and how it is related to the characteristics of the estimator. In particular, deriving the bias components allows us to introduce the analysis on the properties of the *SSCM* estimator.

In doing this, let's start to assume a definition of bias (B) as it may result from the estimation of the Average Treatment Effect on Treated (ATT) parameter applying a traditional propensity score matching procedure to estimate the effect of a training program [8]:

$$B = E(y_0|\mathrm{Z},\ T = 1) - E[y_0|P(\mathrm{Z}_0),\ T = 0] = B_1 + B_2 + B_3 \tag{7}$$

where $P(\mathrm{Z}_0)$ is the conditional propensity score to experience the regime of treatment for the untreated. Equation (7) describes bias such as a difference between the outcome achievable in a counterfactual condition by treated subjects and the expected outcome gained by a subject that chooses the untreated regime conditional to the his/her propensity to undergo the treatment. The difference in outcome described in Eq. (7) does not depend on the effect of treatment, but on the (observed and latent) characteristics of the subject not depending on treatment. [8] show how bias can be decomposed in three components, B_1, B_2 and B_3, given, respectively, by

lack of common support, different distribution of covariates, **Z**, between treated and untreated groups, and difference in outcomes due to unobservable factors; that is:

$$B_1 = E[(y_0|\mathrm{Z}, P(\mathrm{Z}|T=1)] - E[(y_0|\mathrm{Z}, P(\mathrm{Z}|T=0)] \tag{8}$$

$$B_2 = E(y_0|\mathrm{Z}, T=0)\,[P(\mathrm{Z}|T=1) - P(\mathrm{Z}|T=0)] \tag{9}$$

$$B_3 = E(y_0|\mathrm{Z}, T=1) - E[(y_0|\mathrm{Z}, T=0)]P(\mathrm{Z}|T=1) \tag{10}$$

The component B_3 is given by the selection on unobservables and is defined in Econometrics as selection bias or unobserved heterogeneity.

The *SSCM* estimation procedure is not affected by the first component of bias, B_1, because of the trimming of units outside the common support applied at the first stage of the procedure. Regarding the effects due to differences in covariates and selection on unobservables, we derive the bias terms B_2 and B_3 by decomposing the measurement Eq. (5) adopting the well-known Blinder-Oaxaca (*BO*) decomposition method (e.g. [2, 13])[1]:

$$\Delta_i - \mu_i = \left[\Delta_i - \mathbf{z}_{0i}^{'}(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) - (\mathbf{z}_{1i} - \mathbf{z}_{0i})^{'}\boldsymbol{\alpha}_1 - (\mathbf{z}_{1i} - \mathbf{z}_{0i})^{'}(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)\right] \tag{11}$$

Taking into account Eqs. (2a, 2b and 3), we obtain, from Eq. (11), a more accurate specification of the stochastic component of the causal effects (3):

$$u_{1i} - u_{0i} = \left[\Delta_i - \mathbf{z}_{0i}^{'}(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) - (\mathbf{z}_{1i} - \mathbf{z}_{0i})^{'}\boldsymbol{\alpha}_1 - (\mathbf{z}_{1i} - \mathbf{z}_{0i})^{'}(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)\right] \tag{12}$$

Introducing the definition of *ATT* and using the Eqs. (3 and 4), we have:

$$ATT = E[y_{1i} - y_{0i}|T=1] = E[\Delta_i|T=1] = \mathrm{E}[\mu_i + (u_{1i} - u_{0i})|T=1] \tag{13}$$

and

$$ATT = E[\mathbf{z}^{'}{}_i(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) + (\delta_{1i} - \delta_{0i}) + (u_{1i} - u_{0i})|T=1] \tag{14}$$

$\delta_{1i} - \delta_{0i}$ is the difference in unobserved factors between the two regimes. Taking in to account that latent factors may affect simultaneously both the "shift" in coefficients and the differences in covariates between the two regimes, we replace the mean difference of $\delta_{1i} - \delta_{0i}$, with the "Interaction Term": $(\boldsymbol{\delta}_{1i} - \boldsymbol{\delta}_{0i}) = (\mathbf{z}_{1i} - \mathbf{z}_{0i})^{'}(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$ as in the "three-fold" version of the Blinder-Oaxaca decomposition.

[1] In practice, by introducing the *BO* decomposition, we can detect three distinct effects on v_i, and, consequently, on the probability to undergo treatment T_i^*: (i) the effect of the "shift" in the coefficients, $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_0$, on the outcome due to the choice of the regime; (ii) the effect of "change" in covariates, $\mathbf{z}_{1i}$ and $\mathbf{z}_{0i}$ between the regimes; and (iii) a component that measures the interaction between the first two effects.

Then, substituting Eq. (12) into Eq. (14), we obtain:

$$ATT = E[\Delta_i + z_i^{'}(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) - [\mathbf{z}_{0i}^{'}(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) + (\mathbf{z}_{1i} - \mathbf{z}_{0i})^{'}\boldsymbol{\alpha}_1]|\mathrm{T} = 1] \quad (15)$$

Equation (15) shows that, in order to ensure the correct estimation of the *ATT* parameter, given by $E[\Delta_i|T = 1]$, the sum of the "shift component", $\mathbf{z}_{0i}^{'}[(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$, measuring the unexplained change of regime, and the "change-in-covariates" (or "covariates") component, $(\mathbf{z}_{1i} - \mathbf{z}_{0i})^{'}\boldsymbol{\alpha}_1$, measuring the effect of difference in covariates between treated and untreated, should equalized the deterministic component of the causal effects of treatment, given by $z_i^{'}(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$. Both shift and covariates components are included in the specification of the regressors set of the measurement equation. Then the estimation of the measurement equation, adopting a Kalman filter, allows us to provide a correction of bias in treatment effect estimation. The covariates component, $(\mathbf{z}_{1i} - \mathbf{z}_{0i})^{'}\boldsymbol{\alpha}_1$, and the shift component $\mathbf{z}_{0i}^{'}(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$, can be here considered analogous, respectively, to the B_2 and B_3 components of bias in matching.

3 Monte Carlo Experiment

We propose a Monte Carlo experiment to compare the performance of the *SSCM* procedure with that of the *PSME* in terms of bias reduction under both the conditions of heterogeneous and homogeneous covariates between regimes.

For this purpose, we generate 500 data sets of 2,000 units from the Two-Regime model above specified in Eqs. (1, 2a and 2b). The exogenous covariates $\mathbf{Z}$ are generated in order to reproduce the very frequent condition of heterogeneity in observed covariates between treatment and comparison group, and the condition of homogeneity in the observed covariates between regimes. More in detail, we consider a 5 $\times$ 1 vector $\mathbf{z}_i$ in Eq. (1), where the five variables are generated from the following distributions: z_{1i} from a Uniform ranging in [0,10]; z_{2i} from a Normal with mean 10 and variance 16, call it N(10, 16); $z_{3i} = z_{1i}N(-5; 16)$; $z_{4i} = z_{1i}z_{2i} + N(0; 4)$; $z_{5i} = N(-10; 16) + z_{2i}^2$. The coefficients included in the vector β are: $\beta_1 = 10$; $\beta_2 = -10$; $\beta_3 = 10$; $\beta_4 = 10$; $\beta_5 = -10$. . The choice of the distributions and their parameters has the purpose of including in the explanatory variables both independent and dependent (also quasi-collinear) variables.

In order to simulate the effect of endogeneity on the estimates, we consider two different DGP, with and without endogeneity, so as to fix two distinct sets of population parameters under the condition of endogeneity and exogeneity, respectively.

In order to embed endogeneity in the selection equation (Eq. 1) and in the outcome equations (Eqs. 2a and 2b), we generate the random variable $\xi_i = N(0; 1) + (u_{1i} - u_{0i})$, where u_{1i} and u_{0i} are the error terms of the outcome equations (Eqs. 2a and 2b), generated, respectively, as follows:

Table 1 Population ATT parameters derived from DGP

DGP	$\sigma_{1v}(\rho_{1v})^*$	$\sigma_{0v}(\rho_{0v})^*$	Population *ATT*	Population *ATT* with observed heterogeneity
Endogeneity (1)	5.4(0.9)	2.4(0.6)	6.12	6.15
Endogeneity (2)	5.4(0.9)	−2.4(−0.6)	9.80	9.80
Endogeneity (3)	5.4(0.9)	0.8(0.2)	8.47	8.47
Endogeneity (4)	5.4(0.9)	−0.8(−0.2)	8.47	8.47
No Endogeneity	0.0 (0.0)	0.0 (0.0)	5.00	5.00

Note *Taking into account the variances of u_1 and u_0, the corresponding correlation coefficients are approximately given by the values shown in brackets

$$u_{1i} = \sigma_{1v}\vartheta_i + \varepsilon_{1i} \text{ and } u_{0i} = \sigma_{0v}\vartheta_i + \varepsilon_{0i} \tag{16}$$

where ϑ_i is a $N(0;1)$ random variable, and ε_{1i} and ε_{0i} are independently generated by a $N(0;36)$ and a $N(0;16)$, respectively. Finally, we standardize ξ_i and obtain the disturbance term of the selection equation, v_i following a $N(0;1)$ distribution, as stated above specifying the *selection equation* (Eq. 1).

Note that we decide to fix different values of σ_{1v} and σ_{0v} in each experiment (reported, below, in Tables 1 and 2), in order to reproduce different conditions of endogeneity. In particular, if σ_{1v} and σ_{0v} are both positive, we obtain an unobserved heterogeneity that positively influences both the propensity to undergo the treatment and the ability of the subject to gain the outcome. The opposite occurs if one of these covariances has a negative value.[2]

To reproduce the probability to undergo treatment, we generate a cumulative Normal Standard distribution, given by $\Phi(\beta_0 + \beta_1 z_{1i} + ... + \beta_k z_{ki} + v_i)$.

Regarding to the outcome equations, we include in the right-hand sides of the outcome equations (Eqs. 2a and 2b) two components, μ_{1i} and μ_{0i}, exogenously generated by a $N(15;25)$ and a $N(10;16)$ random variables, respectively. Endogeneity in outcome equations (Eqs. 2a and 2b) is given by the relationship between the disturbance terms u_{1i} and u_{0i} and the error term of the selection equation, v_i, as above specified.

The "Population" treatment parameter considered in our analysis is given by the Average Treatment on Treated (*ATT*): $E(y_{1i} - y_{0i}|T_i = 1)$. Applying Eq. (5), Population *ATT* is equal to $E[\mu_{1i} - \mu_{0i} + (\sigma_{1v} - \sigma_{0v})v_i + (\varepsilon_{1i} - \varepsilon_{0i})|T_i = 1]$ and converges on different limits depending on the pre-determined values of the covariances σ_{1v} and σ_{0v}. As a consequence, setting both σ_{1v} and σ_{0v} equal to zero, we obtain the Population *ATT* in absence of endogeneity. Table 1 shows the Population *ATT* parameters generated under different values determined for σ_{1v} and σ_{0v} in order to reproduce endogeneity.

[2] For example, considering a two-regime model of wage for unionized and non-unionized workers, latent cultural factors may induce a worker who gains a higher wage not to join the union, even though an unionized worker should have greater economic protection.

Table 2 Estimated ATT parameters without heterogeneity in observed covariates. Population ATT value = 5

DGP	*SSCM*			*PSME*		
Endogeneity	***ATT***	*95% CI*		***ATT***	*95% CI*	
σ_{1v} 5.4; σ_{0v} 2.4	4.974	*4.930*	*5.018*	7.996	*7.970*	*8.022*
σ_{1v} 5.4; σ_{0v} −2.4	4.320	*4.279*	*4.362*	6.814	*6.782*	*6.846*
σ_{1v} 5.4; σ_{0v} 0.8	4.983	*4.939*	*5.027*	7.571	*7.534*	*7.607*
σ_{1v} 5.4; σ_{0v} −0.8	4.729	*4.688*	*4.77*	7.572	*7.536*	*7.608*
	% BIAS*	*St.Dev*	*t***	***% BIAS****	*St.Dev*	*t***
σ_{1v} 5.4; σ_{0v0v} 2.4	−0.51%	*0.022*	222.16	59.92%	*0.013*	*607.170*
σ_{1v} 5.4; σ_{0v} −2.4	−13.59%	*0.021*	205.72	36.28%	*0.016*	*414.280*
σ_{1v} 5.4; σ_{0v} 0.8	−0.35%	*0.022*	222.48	51.41%	*0.018*	*410.660*
σ_{1v} 5.4; σ_{0v} −0.8	−5.42%	*0.021*	225.71	51.45%	*0.018*	*412.000*

Note * % of *Bias* [(Est. *ATT*-5)/5]%; ** t-ratio: (*ATT*/*St.Dev*.)

Moreover, in order to consider heterogeneity in covariates, we replicate the experiment with some different specifications in the covariates **Z** with respect to the previous scheme. For $T_i = 1$, the variable z_{1i} is generated from a $U(0;13)$ and the variable z_{2i} from a $N(13,4)$, while the variables z_{3i}, z_{4i} and z_{5i} change according to the previous scheme. As a consequence of heterogeneity in covariates, negligible changes in the values of the population *ATT* occur, as reported in the last column of Table 1.

The differences between the estimated *ATT* obtained under endogeneity and the "unbiased" population *ATT* value (equal to 5) quantify the effect of endogeneity simulated by the *DGP*. Hence, we can evaluate the bias of the estimated *ATT* parameters, obtained by applying matching methods in different conditions of endogeneity.

3.1 Comparison of Matching Estimators

The proposed *SSMC* estimator is compared to the *PSME* on the simulated data.

For the sake of clarification, we provide a brief description of the steps needed to apply the estimation methods.

(i) ***SSCM***:

Step 1: let us start to run the *BO* decomposition of the outcome, using the treatment dummy T_i in order to indicate the choice of regime. The outcomes, y_{1i} and y_{0i}, are pooled to form the dependent variable, as well as the covariates of the selection, in order to generate the five explanatory variables. The decomposition is replicated for the dummy $1 - T_i$ (equal to 0 for treated, and equal to 1 for untreated).

Step 2: two new variables, named "*Split*" and "*Endowment*", are generated. The variable *Split* measures the "shift" effect on the outcome as a consequence of the

choice of the regime given by $\mathbf{z}_{0i}^{'}(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$, while the variable *Endowment* measures the effect of difference in covariates between regimes, given by $(\mathbf{z}_{1i} - \mathbf{z}_{0i})^{'}\boldsymbol{\alpha}_1$ (see, above, Sect. 3.1).

Step 3: the propensity score matching is performed using the estimated propensity scores in order to obtain a preliminary estimation of the causal effects $\widetilde{\Delta}_i = y_{1i} - y_{0c}$ (where y_{0c} is the counterfactual of y_{1i} belonging to the comparison group and characterized by the same propensity score of the *i*-th unit). The propensity to undergo the treatment is assumed to be conditional to the five above generated exogenous covariates and the variable *Split*.

The estimated propensity scores are then sorted in ascending order, and the estimated causal effects are indexed and ordered accordingly to the estimated propensity scores.

Step 4: A Maximum Likelihood estimation of the State-Space model is performed adopting $\widetilde{\Delta}_i = y_{1i} - y_{0c}$ as a dependent variable and the variables *Split* and *Endowment* as covariates in the measurement equation; while the transition equation is specified as in Eq. (7).[3] The predicted values, $\widehat{\Delta}_i$, are the estimates of endogenous components of the causal effects, $\sigma\nu_i$ (cf., above, Sect. 2).

Step 5: subtracting $\widehat{\Delta}_i$, from y_{1i}, we obtain the corrected outcomes $\hat{y}_i$. A matching procedure is then replicated to link the corrected outcomes $\hat{y}_i$ with the respective counterfactuals in order to obtain the causal effects and the treatment parameters.

(ii) ***PSME***:

A matching procedure is performed, using the propensity scores estimated by *Probit*.[4] The estimation of the causal effects is given by $\widehat{\Delta}_i = y_{1i} - y_{0c}$. The estimated propensity to undergo the treatment is assumed to be conditional to the five above generated exogenous covariates. The estimated causal effects allow us to compute the treatment parameters.

3.2 Comparing Simulation Results of Matching Procedures

We summarize in Tables 2 and 3 the estimated *ATT* values obtained by embedding different endogeneity conditions into the *DGP*. Note that, computing the bias with respect to the population *ATT* value (set to 5), the *SSCM* estimator performs better than the *PSME* procedure. The bias resulting from the application of *SSCM* is markedly smaller than of the one resulting from *PSME*. We can observe, in particular, that, if we reproduce the "more common" endogeneity conditions (characterized

[3] The different step of the estimator are performed using STATA 14 packages.

[4] The STATA 14 package used to perform matching is PSMATCH2 [11]. Performing PSMATCH2, a "one to one" linkage without replacement with a caliper equal to 0.05 is imposed. In addition, the "Common Support" condition is ensured.

Table 3 Estimated *ATT* parameters with heterogeneity in observed covariates. Population *ATT* value = 5

DGP	*SSCM*			*PSME*		
Endogeneity	***ATT***	*95% CI*		***ATT***	*95% CI*	
σ_{1v} 5.4; σ_{0v} 2.4	5.059	*3.482*	*6.527*	7.941	6.551	9.599
σ_{1v} 5.4; σ_{0v} −2.4	4.246	*2.879*	*5.681*	6.888	5.548	8.339
σ_{1v} 5.4; σ_{0v} 0.8	5.128	*3.356*	*6.722*	7.982	6.566	9.500
σ_{1v} 5.4; σ_{0v} −0.8	4.757	*3.433*	*6.083*	7.599	5.993	9.045
	% BIAS*	*St.Dev*	*t***	***% BIAS****	*St.Dev*	*t***
σ_{1v} 5.4; σ_{0v} 2.4	1.19%	*0.529*	9.56	58.82%	*0.498*	15.94
σ_{1v} 5.4; σ_{0v} −2.4	−15.07%	*0.512*	8.29	37.76%	*0.491*	14.03
σ_{1v}5.4; σ_{0v} 0.8	2.53%	*0.517*	9.93	59.63%	*0.479*	16.66
σ_{1v} 5.4; σ_{0v} −0.8	−4.86%	*0.523*	9.09	51.99%	*0.484*	15.71

Note * % of *Bias* [(Est. *ATT*-5)/5]%; ** t-ratio: (*ATT*/*St.Dev*.)

by covariances, σ_{1v} and σ_{0v}, with the same sign) in the *DGP*, the confidence intervals obtained by the *SSCM* estimates include the population *ATT* value. In the less frequent case, in which the propensity to undergo treatment is endogenously affected in the two regimes with opposite sign, confidence intervals of the *SSCM* estimates do not include the population parameter. However, the percentage of bias of *SSCM* estimation does not exceed 15% in absolute value.

While the mean of the estimated *ATT* using the *SSCM* procedure is not influenced by the presence of heterogeneity in covariates between the regimes, standard errors and confidence intervals are found to be markedly increased with respect to the case of no heterogeneity in covariates.

Tables 2 and 3 report the statistics on *ATT* and *Bias* only. In the Appendix simulation results are reported in detail, including the statistics measuring balancing in covariates between treated and untreated cases after matching. These statistics allow us to evaluate the extent to which the adopted matching procedure reduces differences in covariates between treated and untreated units (cf. [5, 15]). In addition, we also report the mean of estimated coefficients measuring dependence between causal effects and propensity score values, and the mean of the coefficients measuring the autoregressive component of the causal effects, resulting by the estimation of the Transition Equation of the *state space* model (see, above, Eq. 6). In the following table (Table 4), we summarize the description of the coefficients and indicators provided by the Monte Carlo experiments.

Analyzing the statistics, reported below in Appendix, obtained by the Monte Carlo experiments, we observe that the *RHO*(T; (Δ) coefficient, measuring, in the *SSCM* procedure, the correlation between the estimated propensity score and the causal effects, is generally higher if heterogeneity in covariates has been embedded in *DGP*. We found that, with heterogeneous covariates, *RHO*(T; Δ) ranges between 14 and 15%. Instead *RHO*(T; Δ) ranges between 2 and 5% if covariates are not

Table 4 Description of the coefficients and indicators provided by the Monte Carlo experiments

$RHO(T;\Delta)$	Correlation coefficient between estimated propensity score and causal effects
Estimated ATT_{SSCM}	Estimated ATT parameter obtained performing the SSCM matching Estimator
(ATT_{PSME})	Estimated ATT parameter obtained performing the PSME matching Estimator
Unbiased Pop. *ATT*	ATT generated by DGP without endogeneity
Biased Pop. *ATT*	ATT generated by DGP under endogeneity
Estim. Transition coeff	Coefficient of the Transition equation measuring the autoregressive effect in ordered causal effects
Shift coeff	Coefficient estimated in the "measurement equation", corresponding to the explanatory variable "Split" provided by the Blinder-Oaxaca decomposition
Endowment coeff	Coefficient estimated in the "measurement equation", with the introduction as regressor of the difference in covariates provided by the Blinder-Oaxaca decomposition
Mean bias after matching	Standardized mean difference between treatment and control units after matching
Median bias after matching	Median difference between treatment and control units after matching
BAFT	"Rubins' B" indicator: The absolute standardized difference of the means of the linear index of the propensity score in the treated and non-treated (matched)
RAFT	"Rubin's R" index: The ratio of variances of the propensity score index between treated and non-treated (matched)
No. of units on the common support	No. of observations belonging to the treatment group or to the comparison group who have an estimated propensity score equal to that of one or more observations belonging to the opposite group

affected by heterogeneity. In addition, we generally observe significant estimated values of the coefficient of the transition equation, ρ, measuring the autoregressive effect in ordered causal effects. Estimated values are close to -0.44 in all the Monte Carlo experiments.

The impact of the endogenous change of regime and of the difference in covariates across regimes in the *measurement equation* (Eq. 7), are measured by the "Shift" and the "Endowment" coefficient, respectively. Note that the Endowment coefficient is generally higher (ranging between 0.73 and 0.76) if covariates are not affected by heterogeneity. It ranges between 0.52 and 0.56 in the case of heterogeneous covariates across regimes. On the contrary, the "Shift" coefficient is higher if heterogeneity in covariates occurs (between 1.39 and 1.59) than the opposite case of homogeneous covariates (between -0.06 and 0.15).

The *BAFT* and *RAFT* indicators provide, in all the experiments, values compatible with a satisfactory balancing using both *SSCM* and *PSME* estimators.[5] In general, a better balancing is obtained setting the experiment with homogeneous covariates between regimes.

4 Concluding Remarks

The aim of this study is to improve the propensity-score matching approach so that estimation results do not overly suffer from the influence of the endogeneity of treatment. In doing this, the main innovation introduced here is given by the specification of the endogenous relationship between the individual propensity score (individual probability to undergo the treatment) and the individual causal effect of treatment. In practice, we start by assuming that the probability that a subject undergoes treatment endogenously depends, at least in part, on the potential effect of the treatment. This implies that two subjects with the same (or similar) propensity score should expect similar results in terms of causal effects. This allows us to consider the causal effect of treatment on each subject such as correlated with the causal effect of another subject with similar propensity score.

A consequence of this assumption is that the causal effects, ordered by their correspondent propensity scores, are autocorrelated via their endogenous component. As an empirical verification of this assumption, we apply a state-space model to estimate the autocorrelated endogenous component of the causal effects, so as to use the result of this estimate as a correction term. In particular, the results of the Monte Carlo experiments here reported confirm that, simulating endogeneity of the selection into treatment in a Two-Regime model, the predicted components of causal effects can be successfully used, at a second stage of the estimation procedure, to correct the matches outcomes. In addition, we verify that, simulating difference in covariates between treated and untreated groups, the bias in *ATT* estimation reduces more if we apply the *SSCM* procedure.

As the results of our empirical analysis show, this method allows us to reduce the selectivity bias in matching without imposing, to the data or the model, any restriction usually adopted to reproduce a condition comparable to randomization.

Both bias components, due, respectively, to differences in covariates between treated and untreated and to the influence of latent factors, is reduced using *SSCM,* as the above reported simulation results show. Note that, applying the usual matching procedures, the bias due to the influence of unobservable factors can be reduced to the extent that these latent factors are considered as "ignorable" (see [12]). Applying the *SSCM* approach, a reduction in bias is also achieved in "non-ignorability" conditions. This result is obtained by decomposing the measurement equation of *SSCM* model,

[5] Rubin [15] recommends that *BAFT* be less than 25 and that *RAFT* be ranged between 0.5 and 2 for the samples to be considered sufficiently balanced.

as reported above in Eq. (15). The correction introduced by estimating measurement equation allows us to reduce both bias components under the imposition of endogeneity, that is if ignorability conditions do not hold.

At this stage of our research, we have deepened the characteristics of the *SSCM* estimator only by means of Monte Carlo experiments. The next aim of this research is to verify the robustness of the *SSCM* estimator by performing Monte Carlo experiments on DGP generated under different distributive assumptions.

Considering our study at an early stage, the field of application currently investigated is limited to the comparison of propensity-score based methods between treated and untreated units in a cross-sectional context. Possible extensions could deal with a broader evaluation of the application of the *SSCM* method compared to other cross-sectional methods, such as the Inverse Probability Matching, Nearest Neighbor and Bias Corrected matching (cf. [1]).

Appendix

Results of the Monte Carlo Experiments.

(1) *Simulated data without heterogeneity in observed covariates* (Tables 5, 6, 7 and 8)

Table 5 No heterogeneity in covariates. No. of Reps. = 500. Simulated endogeneity: $\sigma_{1v} = 5.4$; $\sigma_{0v} = 2.4$

	SSCM (No. of units on common support: 1882)				*PSME* (No. of units on common support: 1995)			
	Mean	*Std. Dev*	C. I. 95%		Mean	*Std. Dev*	C. I. 95%	
$RHO(T;\Delta)$	0.0266	*0.0012*	0.0243	0.0289	0.0602	*0.0007*	0.0588	0.0615
Estimated ATT_{SSCM}	4.9744	*0.0224*	4.9304	5.0184	7.9958	*0.0132*	7.97	8.0217
Unbiased Pop. ATT	*4.9903*	*0.0091*	*4.9725*	*5.0081*	*4.9988*	*0.0064*	4.9862	5.0113
Biased Pop. *ATT*	6.123	*0.0137*	6.096	6.1499	6.146	*0.0098*	6.1267	6.1653
Estim. Transition coeff	−0.4426	*0.0009*	−0.4443	−0.4409				
Shift coeff	−0.0061	*0.0185*	−0.0425	0.0304				
Endowment coeff	0.7529	*0.0018*	0.7492	0.7565				
Mean bias after matching	2.6819	*0.0583*	2.5673	2.7965	2.0514	*0.0347*	1.9833	2.1194
Median bias after matching	2.5889	*0.0681*	2.455	2.7228				
BAFT	8.7095	*0.1517*	8.4114	9.0076	7.4477	*0.0873*	7.2763	7.6191
RAFT	1.0047	*0.0073*	0.9904	1.0191	1.0288	*0.0048*	1.0193	1.0383

Table 6 No heterogeneity in covariates. No. of Reps. = 500. Simulated endogeneity: $\sigma_{1v} = 5.4$; $\sigma_{0v} = 0.8$

	SSCM (No. of units on common support: 1891)				*PSME* (No. of units on common support: 1995)			
	Mean	*Std. Dev*	C. I. 95%		Mean	*Std. Dev*	C. I. 95%	
$RHO(T;\Delta)$	0.0245	*0.0011*	0.0223	0.0267	0.0502	*0.0009*	0.0483	0.052
Estimated ATT_{SSCM}	4.9825	*0.0224*	4.9385	5.0265	7.5705	*0.0184*	7.5342	7.6067
Unbiased Pop. ATT	*5.0073*	*0.0084*	4.9908	5.0238	*5.007*	*0.0088*	*4.9898*	*5.0242*
Biased Pop. *ATT*	7.2388	*0.014*	7.2114	7.2663	8.4698	*0.0154*	8.4395	8.5002
Estim. Transition coeff	− 0.4419	*0.0008*	− 0.4436	− 0.4403				
Shift coeff	0.0287	*0.017*	− 0.0047	0.062				
Endowment coeff	0.7594	*0.0017*	0.756	0.7628				
Mean bias after matching	2.495	*0.0565*	2.3841	2.606	2.1077	*0.0488*	2.0118	2.2036
Median bias after matching	2.3444	*0.0642*	2.2182	2.4705	1.9453	*0.055*	1.8373	2.0533
BAFT	8.4844	*0.1457*	8.1983	8.7705	7.6179	*0.1249*	7.3725	7.8633
RAFT	1.0012	*0.0068*	0.9879	1.0146	1.0128	*0.0062*	1.0007	1.0249

(2) *Simulated data introducing heterogeneity in observed covariates* (Tables 9, 10, 11 and 12)

Table 7 No heterogeneity in covariates. No. of Reps. = 500. Simulated endogeneity: $\sigma_{1v} = 5.4$; $\sigma_{0v} = -2.4$

	SSCM (No. of units on common support: 1901)				*PSME* (No. of units on common support: 1995)			
	Mean	*Std. Dev*	C. I. 95%		Mean	*Std. Dev*	C. I. 95%	
$RHO(T;\Delta)$	0.0208	*0.0011*	0.0187	0.0229	0.0441	*0.0008*	0.0425	0.0457
Estimated ATT_{SSCM}	4.3204	*0.0210*	4.2791	4.3616	6.8141	*0.0164*	6.7818	6.8464
Unbiased Pop. ATT	*4.9968*	*0.0083*	*4.9805*	*5.0132*	*4.9971*	*0.0079*	4.9816	5.0126
Biased Pop. *ATT*	9.7883	*0.0179*	9.7532	9.8234	9.7787	*0.0145*	9.7502	9.8072
Estim. Transition coeff	-0.4427	*0.0008*	− 0.4443	− 0.4411				
Shift coeff	0.1542	*0.0176*	0.1196	0.1888				
Endowment coeff	0.7332	*0.0018*	0.7296	0.7367				
Mean bias after matching	2.7085	*0.0691*	2.5728	2.8442	2.0963	*0.0454*	2.0071	2.1856
Median bias after matching	2.532	*0.0701*	2.3942	2.6698	1.9439	*0.0504*	1.8449	2.0429
BAFT	9.0678	*0.1807*	8.7127	9.4229	7.3877	*0.1087*	7.1742	7.6011
RAFT	1.0058	*0.0072*	0.9916	1.02	1.0223	*0.0057*	1.011	1.0336

Table 8 No heterogeneity in covariates. No. of Reps. = 500. Simulated endogeneity: $\sigma_{1v} = 5.4$; $\sigma_{0v} = -0.8$

	SSCM (No. of units on common support: 1895)				*PSME* (No. of units on common support: 1995)			
	Mean	*Std. Dev*	C. I. 95%		Mean	*Std. Dev*	C. I. 95%	
$RHO(T;\Delta)$	0.0225	*0.0012*	0.0202	0.0248	0.0502	*0.0009*	0.0484	0.052
Estimated ATT_{SSCM}	4.7288	*0.0210*	4.6876	4.77	7.5723	*0.0184*	7.5362	7.6084
Unbiased Pop. ATT	*5.0020*	*0.0093*	*4.9838*	*5.0203*	*5.0074*	*0.0088*	*4.9902*	*5.0247*
Biased Pop. *ATT*	8.4783	*0.0152*	8.4485	8.5081	8.4712	*0.0154*	8.4409	8.5015
Estim. Transition coeff	− 0.4419	*0.0008*	− 0.4435	−0.4402				
Shift coeff	0.1168	*0.0164*	0.0846	0.1491				
Endowment coeff	0.7482	*0.0017*	0.7449	0.7516				
Mean bias after matching	2.6499	*0.0601*	2.5317	2.768	2.1056	*0.0489*	2.0095	2.2016
Median bias after matching	2.5703	*0.0677*	2.4374	2.7033	1.9435	*0.055*	1.8353	2.0516
BAFT	8.7016	*0.1570*	8.3931	9.01	7.6145	*0.1251*	7.3687	7.8603
RAFT	0.9952	*0.0066*	0.9822	1.0081	1.012	*0.0061*	1.0000	1.0241

Table 9 Heterogeneity in covariates. No. of Reps. = 500. Simulated endogeneity: $\sigma_{1v} = 5.4$; $\sigma_{0v} = 2.4$

	SSCM (No. of units on common support: 1934)				*PSME* (No. of units on common support: 1968)			
	Mean	*Std. Dev*	C. I. 95%		Mean	*Std. Dev*	C. I. 95%	
$RHO(T;\Delta)$	0.1567	*0.0370*	0.0608	0.2684	0.2113	*0.0285*	0.1208	0.299
Estimated ATT_{SSCM}	5.0594	*0.5291*	3.4824	6.5270	7.9409	*0.4981*	6.5505	9.5988
Unbiased Pop. ATT	*5.0105*	*0.2105*	4.3976	5.5818	*5.0153*	*0.1974*	4.4988	5.6266
Biased Pop. *ATT*	6.1694	*0.3100*	5.3431	7.0190	6.1529	*0.3106*	5.1768	7.0055
Estim. Transition coeff	−0.4614	*0.0236*	− 0.5444	− 0.3922				
Shift coeff	1.4269	*0.322*	0.6882	2.8712				
Endowment coeff	0.5249	*0.0387*	0.4261	0.6429				
Mean bias after matching	4.9963	*1.7333*	0.8394	12.3551	5.7745	*1.9577*	1.1594	13.21
Median bias after matching	4.8836	*2.3745*	0.5446	14.7685	6.1974	*2.4247*	0.2901	14.9395
BAFT	21.3282	*4.4243*	8.8634	35.3954	22.1471	*4.1054*	5.7511	37.5376
RAFT	1.1155	*0.2168*	0.4998	1.7896	1.168	*0.1905*	0.6644	1.7616

Table 10 Heterogeneity in covariates. No. of Reps. = 500. Simulated endogeneity: $\sigma_{1v} = 5.4$; $\sigma_{0v} = 0.8$

	SSCM (No. of units on common support: 1927)				*PSME* (No. of units on common support: 1969)			
	Mean	*Std. Dev*	C. I. 95%		Mean	*Std. Dev*	C. I. 95%	
$RHO(T;\Delta)$	0.1677	*0.0386*	0.0459	0.2816	0.2162	*0.0268*	0.1464	0.2988
Estimated ATT_{SSCM}	5.1267	*0.5165*	3.3563	6.7218	7.9816	*0.479*	6.5655	9.4996
Unbiased Pop. ATT	*5.0107*	*0.2081*	4.3993	5.594	*5.0035*	*0.2056*	4.3844	5.5943
Biased Pop. *ATT*	7.2626	*0.3269*	6.2737	8.073	7.2141	*0.3304*	6.2855	8.0881
Estim. Transition coeff	−0.4589	*0.0226*	− 0.5227	− 0.3942				
Shift coeff	1.398	*0.3116*	0.7341	2.7256				
Endowment coeff	0.5202	*0.0382*	0.4229	0.6226				
Mean bias after matching	4.8514	*1.6374*	1.3317	10.4524	5.8488	*2.017*	0.7382	13.0318
Median bias after matching	4.7744	*2.4592*	0.5058	14.6848	6.1902	*2.4461*	0.5725	14.2794
BAFT	21.0182	*4.3332*	10.0042	36.3935	21.6758	*4.1721*	10.495	36.443
RAFT	1.1165	*0.2172*	0.4042	1.6645	1.1946	*0.2042*	0.6493	1.7908

Table 11 Heterogeneity in covariates. No. of Reps. = 500. Simulated endogeneity: $\sigma_{1v} = 5.4$; $\sigma_{0v} = -2.4$

	SSCM (No. of units on common support: 1943)				*PSME* (No. of units on common support: 1971)			
	Mean	*Std. Dev*	C. I. 95%		Mean	*Std. Dev*	C. I. 95%	
$RHO(T;\Delta)$	0.1390	*0.0375*	−0.0020	0.2672	0.2001	0.0281	0.1293	0.2913
Estimated ATT_{SSCM}	4.2464	0.5121	2.8788	5.6813	6.8882	0.4908	5.5479	8.3394
Unbiased Pop. ATT	5.0051	*0.2042*	4.3985	5.6521	***5.0095***	0.2065	4.4344	5.6268
Biased Pop. *ATT*	9.8091	*0.3589*	8.7284	11.0152	9.7874	0.3633	8.7151	10.7534
Estim. Transition coeff	−0.4572	*0.0240*	−0.5364	−0.3768				
Shift coeff	1.5859	*0.3853*	0.7016	3.3063				
Endowment coeff	0.5619	*0.0379*	0.4526	0.7021				
Mean bias after matching	5.0343	*1.8335*	0.778	12.2944	5.8729	2.0717	1.0809	13.0339
Median bias after matching	4.9884	*2.4593*	0.3312	13.2262	6.2542	2.4963	0.7634	13.8850
BAFT	20.518	*4.4625*	7.2853	36.2346	21.231	4.4589	10.5499	34.9559
RAFT	1.1321	*0.2230*	0.4343	1.9061	1.1906	0.1927	0.6864	1.8085

Table 12 Heterogeneity in covariates. No. of Reps. = 500. Simulated endogeneity: $\sigma_{1v} = 5.4$; $\sigma_{0v} = -0.8$

	SSCM (No. of units on common support: 1931)				*PSME* (No. of units on common support: 1969)			
	Mean	*Std. Dev*	C. I. 95%		Mean	*Std. Dev*	C. I. 95%	
$RHO(T;\Delta)$	0.1602	*0.0385*	0.0562	0.2763	0.2122	0.0273	0.1462	0.323
Estimated ATT—SSCM	4.7568	*0.5233*	3.4332	6.0831	7.5994	0.4838	5.9927	9.0451
Unbiased Pop. ATT	*5.0122*	*0.205*	4.4301	5.609	*5.0102*	0.2049	4.4301	5.6047
Biased Pop. *ATT*	8.4813	*0.3355*	7.366	9.5244	8.4679	0.3483	7.366	9.3284
Estim. Transition coeff	-0.4573	*0.0234*	− 0.532	−0.391				
Shift coeff	1.4295	*0.3139*	0.6726	2.4584				
Endowment coeff	0.535	*0.0408*	0.4236	0.6628				
Mean bias after matching	4.9532	*1.7404*	1.3584	10.1984	5.9958	2.17	1.1903	15.7593
Median bias after matching	4.865	*2.3882*	0.379	11.9008	6.4759	2.6393	0.787	18.581
BAFT	20.7891	*4.6601*	8.6572	37.0917	21.4835	4.6189	8.5545	38.2082
RAFT	1.1283	*0.2153*	0.5115	1.7224	1.1834	0.1941	0.5498	1.8284

References

1. Abadie, A., Imbens, G.W.: Bias-corrected matching estimators for average treatment effects. J. Bus. Econ. Stat. **29**(1), 1–11 (2011)
2. Blinder, A.S.: Wage discrimination: reduced form and structural estimates. J. Hum. Res. **8**(4), 436–455 (1973)
3. Carneiro, P., Hansen, K.T., Heckman, J.J.: Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. Int. Econ. Rev. **44**(2), 361–422 (2003)
4. Harvey, A.C.: Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press, Cambridge (1990)
5. Haviland, A., Nagin, D.S., Rosenbaum, P.R.: Combining propensity score matching and group-based trajectory analysis in an observational study. Psychol. Methods **12**(3), 247–267 (2007)
6. Heckman, J.J.: The principles underlying evaluation estimators with an application to matching. Ann. Econ. Stat. **91**(92), 9–73 (2008)
7. Heckman, J.J., Honoré, B.E.: The empirical content of the Roy model. Econometrica **58**, 1121–1149 (1990)
8. Heckman, J.J., Ichimura, H., Todd, P.E.: Matching as an econometric evaluation estimator: evidence from evaluating a job training programme. Rev. Econ. Stud. **64**(4), 605–654 (1997)
9. Heckman, J.J., Navarro-Lozano, S.: Using matching, instrumen-tal variables, and control functions to estimate economic choice models. Rev. Econ. Stat. **86**(1), 30–57 (2004)
10. Imbens, G.W., Angrist, J.D.: Identification and estimation of local average treatment effects. Econometrica **62**(2), 467–475 (1994)

11. Leuven, E., Sianesi, B.: PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Statistical Software Components S432001, Boston College Department of Economics, revised 01 Feb 2018 (2003). https://econpapers.repec.org/software/bocbocode/s432001.htm. Accessed 29 Sept. 2018
12. Morgan, S.L., Harding, D.J.: Matching estimators of causal effects: prospects and pitfalls in theory and practice. Sociol Method Res. **35**, 3–60 (2006)
13. Oaxaca, R.: Male–female wage differentials in urban labor markets. Int. Econ. Rev. **14**, 693–709 (1973)
14. Rosenbaum, P.R., Rubin, D.B.: the central role of the propensity score in observational studies for causal effects. Biometrika **70**(1), 41–55 (1983)
15. Rubin, D.B.: Using propensity scores to help design observational studies: application to the tobacco litigation. Health Serv. Outcomes Res. Methodol **2**(3–4), 169–188 (2001)

Gender Gap Assessment and Inequality Decomposition

Michele Costa

Abstract We propose to measure and to evaluate gender gaps and gender inequalities by means of the decomposition of an inequality measure. A three-terms decomposition of the Gini index is applied, thus allowing to take into account also the role of overlapping between female and male subpopulations. We develop an unified framework for the evaluation of gender gap, linking traditional measures, based on subgroups income means, to the approach related to inequality decomposition, and showing how overlapping component represents a key issue in gender gap analysis. An analysis of the income distribution of the Italian households shows how gender gaps represent a major source of inequality, without particular improvements during the last 20 years.

Keywords Gender gap · Gender inequality · Inequality decomposition · Gini index · Overlapping component

1 Introduction

Gender inequalities and gender gaps are a worldwide concern and represent the core of uncountable actions and policies developed by either governments and institutions. Gender inequalities are firstly a primary and fundamental issue of justice and represent the current expression of long standing questions about the sources of inequality advanced by philosphers, political scientists and economists during the centuries. Consequences of gender inequalities are frequently overlooked or underestimated, while it exists a thoughtful literature which analyzes the relation between gender inequality and welfare, pointing out gender gaps as a constraint for economic growth. Furthermore also overall inequality is positively related to gender gap.

A growing literature aims at the evaluation of gender gap and of its effects, with an impressive escalation of contributions and proposals during the last years (see e.g. [2, 4, 6, 14, 16–18]).

M. Costa (✉)
Department of Economics, University of Bologna, Piazza Scaravilli, 2, 40126 Bologna, Italy
e-mail: michele.costa@unibo.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_8

We assess the role of gender in income inequality by decomposing the Gini inequality ratio following the approach introduced by Dagum in 1997 [9]. The first advantage of our approach is to develop a measurement system of income gender gap based on differences not only in average incomes, but in female and male income distributions. The necessity of evaluating the gender gap as a comparison between female and male income distributions and not only as a difference between average incomes is addressed, among the others, by [14, 19].

An additional merit of the paper refers to the overlap which occurs between female and male income distributions and which is fundamental in the evaluation of gender gap. Overlap represents a key point in inequality analysis and it requires a careful assessment. In the absence of overlap, the female and male subgroups would be completely separated, and we would obtain a simplified situation, corresponding to a perfectly stratified society, where the gender gap is maximum. On the other side, increasing levels of overlap indicate a weaker gender gap, and, if female and male subgroups overlapped completely, the influence of gender on inequality would reache its minimum. Regarding this aspect, the Dagum's decomposition is particularly suitable, as it is among the few decompositions that explicitly consider overlap.

Our main contribution is to evaluate the overlap between male and female distributions and to include also this element in the gender gap analysis.

We also contribute to literature on gender gap and on inequality decomposition by extending traditional approaches, thus achieving a more complete and effective assessment of the gender gap. We propose two generalizations of the usual decomposition framework.

First, we focus on the differences at the lower and at the upper end of the male and female distributions, by comparing the related decompositions and by analysing how gender gap affects different parts of the distributions.

Second, by means of inequality decomposition we are able to analyse the inequality structure and take into account other variables (such as area of residence or educational level), able to influence the gender gap. In this way it is possible to develop a broader approach to the analysis of gender gap, where other factors can be added to the income dimension. Since gender gap is a multidimensional issue (see e.g. [1, 11]), which cannot be fully explained by only the income dimension, the inclusion of further inequality factors is of great importance.

The next Section briefly outlines the main aspects of the Dagum's Gini index decomposition, while Sect. 3 illustrates the methods proposed for gender gap evaluation. Section 4 presents a case study for Italian households and Sect. 5 concludes.

2 The Dagum's Gini Index Decomposition

The Gini index is one of the most important measure of inequality and, during its over 100 years of life, has experienced many different interpretations, expressions and formulas. For the case of a population disaggregated into k subgroups of size n_j, with $\sum_{j=1}^{k} n_j = n$, the Gini index can be expressed as

$$G = \frac{1}{2n^2\bar{y}} \sum_{j=1}^{k} \sum_{h=1}^{k} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} |y_{ji} - y_{hr}| \tag{1}$$

where $\bar{y}$ is the arithmetic mean of Y in the overall population, y_{ji} is the value of Y in the i-th unit of the j-th subgroup and, accordingly, y_{hr} is the value of Y in the r-th unit of the h-th subgroup.

A further expression of the Gini index which is extremely useful for the index decomposition is

$$G = \sum_{j=1}^{k} \sum_{h=1}^{k} G_{jh} p_j s_h \tag{2}$$

where $p_j = n_j/n$ and $s_j = (n_j \bar{y}_j)/(n\bar{y})$ are the population share and the income share of the j-th subgroup, respectively, while G_{jh} is the Gini index between subgroup j and subgroup h, with

$$G_{jh} = \frac{1}{n_j n_h (\bar{y}_j + \bar{y}_h)} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} |y_{ji} - y_{hr}| = \frac{\Delta_{jh}}{\bar{y}_j + \bar{y}_h}. \tag{3}$$

For $j = h$, G_{jh} is the Gini index of the j-th subgroup.

The literature on the Gini index is quite extensive and cannot be easily summarized, however for a detailed discussion of the Gini index see, e.g., [7, 12].

The information provided by the Gini index, related to the overall inequality level, can be successfully exploited by means of the Gini index decomposition. Our interest is specifically dedicated to this point, since it represents the key for our proposal to address gender gap analysis.

Among the many methods which allow to decompose the Gini index (see, e.g., [8, 13, 20]), we use the decomposition proposed by Dagum [9].

Besides the two traditional components that characterize all the index decompositions, i.e. the inequality within subgroups G_w and the inequality between subgroups G_b, Dagum explicitly considers a third component, G_t, related to the overlapping, or transvariation (using the Gini's terminology), between subgroups. Since overlapping represents a major feature of many real situations, such as the analysis of female and male income distributions, we think that it is necessary to include also this element into the decomposition. In this way we attribute to overlapping and its interpretation the same status as inequality within and inequality between, without considering this third component as an unfortunate and undesirable residual term.

The appeal of the Dagum's proposal is its great simplicity: the n^2 differences $|y_{ji} - y_{hr}|$ in (1) are directly assigned to the three components of the decomposition.

First, inequality within G_w is derived from the differences $|y_{ji} - y_{hr}|$ belonging to the same subgroup, that is for $j = h$.

Second, when $j \neq h$, that is when the quantities y_{ji} and y_{hr} involve two different subgroups, the difference $|y_{ji} - y_{hr}|$ is assigned to the inequality between G_b if

$\bar{y}_j \geq \bar{y}_h$ and $y_{ji} \geq y_{hr}$, that is when the sign of the difference is the same as the sign of the difference between the subgroups means.

Third, if $j \neq h$, $\bar{y}_j \geq \bar{y}_h$ and $y_{ji} < y_{hr}$ that is in case of an overlapping unit, the difference $|y_{ji} - y_{hr}|$ belongs to the transvariation component G_t.

The analysis of the gender gap implies the presence of $k = 2$ subgroups, as the total population is disaggregated into the female (f) and male (m) subgroups. The existence and the relevance of gender gap originate from $\bar{y}_f$, the average income of the female subpopulation, being strongly and systematically lower than $\bar{y}_m$, the average income of the male subpopulation. In the following, assuming $\bar{y}_f < \bar{y}_m$ as reference, it is illustrated the Dagum's decomposition for the analysis of gender gap.

3 Gender Gap Analysis by Means of Inequality Decomposition

When a population is divided into only 2 subgroups, the female (f) and male (m) subpopulations, the Gini index, defined as in (2), can be expressed as

$$G = G_f p_f s_f + G_m p_m s_m + G_{fm} p_f s_m + G_{mf} p_m s_f$$

The case of two subgroups also allows some important simplifications since

$$p_m = 1 - p_f, \qquad s_m = 1 - s_f, \qquad G_{fm} = G_{mf}.$$

By referring to the female and male subpopulations, we can write the Gini index between f and m as

$$G_{mf} = \frac{\Delta_{mf}}{\bar{y}_m + \bar{y}_f} = \frac{1}{n_m n_f (\bar{y}_m + \bar{y}_f)} \sum_{i=1}^{n_m} \sum_{r=1}^{n_f} |y_{mi} - y_{fr}|.$$

In the following the three components of the Dagum's Gini index decomposition are derived and interpreted within the framework of gender gap evaluation.

3.1 The Inequality Within Subgroups

The component of inequality within G_w can be obtained quite easily as a weighted sum of the Gini indexes of each subgroup, where the weights are given by the population share and the income share of the two subgroups:

$$G_w = G_f p_f s_f + G_m p_m s_m \tag{4}$$

Even if the measurement of G_w is the subject of some criticism (see, e.g. [10]), the Dagum's proposal is in line with the existing literature, where the majority of the decompositions obtains G_w as a function of G_f and G_m.

G_w allows to evaluate the contribution to total inequality related to the variability within the subgroups. Low values of G_w indicate homogenous subgroups and a reduced impact of variability within female and male subgroups on total inequality, while an high G_w provides the opposite indication.

3.2 *The Inequality Between Subgroups and the Overlapping Component*

Given G and G_w, we can obtain the inequality between subgroups G_b and the overlapping component G_t as

$$G - G_w = G_b + G_t = G_{fm} p_f s_m + G_{mf} p_m s_f$$

For G_b and G_t, which in the original version of the Dagum's decomposition require some substantial computational effort, are also available [5] simplified expressions. In order to derive the inequality between subgroups G_b and the overlapping component G_t, it is useful to start from the simplest case, that is only two non overlapping subgroups, and then to introduce the presence of overlapping.

3.2.1 The Case of no Overlapping

A population divided into two non overlapping subgroups, despite being a relevant departure from many real situations, offers a simple and straightforward solution for the measurement of the inequality between. Since $G_t = 0$ we have

$$G_b = G_{fm} p_f s_m + G_{mf} p_m s_f = \frac{p_f s_m + p_m s_f}{n_f n_m (\bar{y}_f + \bar{y}_m)} \sum_{i=1}^{n_f} \sum_{r=1}^{n_m} |y_{fi} - y_{mr}|$$

Given $\bar{y}_f < \bar{y}_m$, the absence of overlapping implies $|y_{fi} - y_{mr}| = (y_{mr} - y_{fi})$ and allows to semplify the expression of G_b:

$$G_b = \frac{p_f s_m + p_m s_f}{n_f n_m (\bar{y}_f + \bar{y}_m)} \sum_{i=1}^{n_f} \sum_{r=1}^{n_m} (y_{mr} - y_{fi}) = \frac{p_f s_m + p_m s_f}{n_f n_m (\bar{y}_f + \bar{y}_m)} n_f n_m (\bar{y}_m - \bar{y}_f) =$$

$$= p_f (s_m + p_m \bar{y}_f / \bar{y}) \frac{(\bar{y}_m - \bar{y}_f)}{(\bar{y}_f + \bar{y}_m)} = p_f p_m \frac{(\bar{y}_m + \bar{y}_f)}{\bar{y}} \frac{(\bar{y}_m - \bar{y}_f)}{(\bar{y}_f + \bar{y}_m)} =$$

$$= p_f p_m \bar{y}_m / \bar{y} - p_f p_m \bar{y}_f / \bar{y} = p_f (1 - s_f) - (1 - p_f) s_f = p_f - s_f.$$

Therefore our final result for G_b allow to evaluate the inequality between as the difference between the population share and the income share of the female subgroup.

The Dagum's Gini index decomposition results

$$G = G_w + G_b = (G_f p_f s_f + G_m p_m s_m) + (p_f - s_f)$$

where the gender gap can be evaluated by the difference $(p_f - s_f)$.

3.2.2 The Case of Overlapping

In the case of overlapping subgroups, some differences $(y_{mr} - y_{fi})$ result negative and therefore the quantity $(p_f - s_f)$, which is the sum of all differences $(y_{mr} - y_{fi})$, both positive and negative, will underestimate the inequality between G_b for an amount equal to the sum of the negative differences.

Since the negative differences correspond to the overlapping component, we have

$$G_b = \frac{p_f s_m + p_m s_f}{n_f n_m (\bar{y}_f + \bar{y}_m)} \left(\sum_{i=1}^{n_f} \sum_{r=1}^{n_m} (y_{mr} - y_{fi}) + \sum_{\substack{i=1 \\ y_f > y_m}}^{n_m} \sum_{r=1}^{n_f} (y_{fr} - y_{mi}) \right).$$

In a more compact form we can write $G_b = p_f - s_f + G_t$ and the Dagum's Gini index decomposition results

$$G = G_w + G_b + G_t = (G_f p_f s_f + G_m p_m s_m) + (p_f - s_f + G_t) + G_t$$

from which it is possible to derive G_t as $G_t = (G - G_w - p_f - s_f)/2$.

Inequality between G_b and overlapping component G_t allow to evaluate the contribution to total inequality attributable to the differences between the subgroups, that is, in our analysis, the gender gap.

The role of the two components is quite different. From one side, an high (low) G_b indicates a relevant (slight) gender gap, as total inequality is (is not) strongly influenced by inequality between. From the other side, an high (low) G_t points out to a slight (relevant) gender gap, since complete overlapping corresponds to the absence of gender gap, while $G_t = 0$ (female and male subgroups are perfectly separated) indicates a total stratification.

The gender gap can still be evaluated by means of G_b, but in the general case of possible overlapping, it is given by $(p_f - s_f + G_t)$.

It is worth to note how the use of G_b to evaluate the gender gap is direct and straightforward, and how it does not imply the inclusion of any additional assumption, or constraint, as frequently happens in the case of parametric methods, even simple ones, such as the linear ones.

3.3 The Informative Content of Subgroups Means

The use of the Dagum's Gini index decomposition for the evaluation of the gender gap also allows to shed some light on the debate about the informative content of the subgroups means $\bar{y}_m$ and $\bar{y}_f$.

The first gross evaluation of the gender gap based on $\bar{y}_m$ and $\bar{y}_f$ is given by the absolute difference

$$gap_1 = (\bar{y}_m - \bar{y}_f)$$

which is usually complemented by some relative difference such as

$$gap_2 = (\bar{y}_m - \bar{y}_f)/\bar{y}_f$$

$$gap_3 = (\bar{y}_m - \bar{y}_f)/\bar{y}.$$

The extreme case $\bar{y}_m = \bar{y}_f = \bar{y}$ of null gender gap implies $p_f = s_f$ and $p_m = s_m$, that is the equidistribution. It follows that, when $p_i = s_i$, the gender gap is equal to 0, while increasing differences $(p_i - s_i)$ report an increasing inequality. The case $p_f > s_f$, that is when the population share of the female subpopulation is greater than the income share of the female subpopulation, implies $\bar{y}_m > \bar{y}_f$ and indicates the existence of a gender gap. In the following we include also this indicator:

$$gap_4 = (p_f - s_f).$$

It is crucial to assess to what extent the subgroups means $\bar{y}_m$ and $\bar{y}_f$ are informative and if they allow to correctly evaluate the gender gap.

We are able to propose a solution by resorting to G_{bBM}, that is the pioneering measure of the inequality between proposed by [3], who evaluate the differences between the subgroups on the basis of the subgroups means only. For the case of two subgroups, m and f, we have

$$G_{bBM} = p_f p_m (\bar{y}_m - \bar{y}_f)/\bar{y}$$

and it is possible to show that $(\bar{y}_m - \bar{y}_f)$ is linked to $(p_f - s_f)$:

$$p_f p_m (\bar{y}_m - \bar{y}_f)/\bar{y} = p_f p_m \bar{y}_m/\bar{y} - p_f p_m \bar{y}_f/\bar{y} = p_f s_m - p_m s_f =$$

$$= p_f(1 - s_f) - (1 - p_f)s_f = p_f - p_f s_f - s_f + p_f s_f = p_f - s_f.$$

Since the quantity $(p_f - s_f)$ corresponds (see Sect. 3.2.1) to the inequality between in the Dagum's Gini index decomposition for the case of two non overlapping subgroups, we are able to affirm that the subgroups means are fully informative in absence of overlapping. However, their informative content declines for increasing levels of overlapping.

From Sect. 3.2.2 we also derive that, when the female and male income distributions overlap, the use of the means leads to underestimate the gender gap, with the underestimation being proportional to the degree of overlapping.

3.4 The Inequality Structure

Inequality decomposition is a powerful tool to get a deep understanding of the inequality structure. We exploit this property to provide insight on gender gap.

Usually the Gini index decomposition is applied to all n available observations on a variable y (such as income, expenditure, etc.) disaggregated on the basis of some inequality factor of interest, in our case the gender of the head of the household.

From the decomposed Gini index

$$G_y = G_{wy} + G_{by} + G_{ty}$$

we analyze the effect of the inequality factor by means of the three ratios G_{wy}/G_y, G_{by}/G_y, G_{ty}/G_y. In particular, by referring to the measurement of the gender gap, we add to our set of measures the second ratio,

$$gap_5 = G_b/G$$

which is based on the inequality between, that is the component of the inequality able to evaluate the relevance of the underlying inequality factor.

3.4.1 The Effect of Low and High Income Values

A first possible extension with respect to the traditional framework is to compare the decomposition obtained by using all n observations to the decompositions obtained by referring only to subsamples of observations.

In particular, it is useful to analyze the decompositions for the lower values of y,

$$G_{y|ymin} = G_{wy|ymin} + G_{by|ymin} + G_{ty|ymin}$$

as well as for the higher values of y,

$$G_{y|ymax} = G_{wy|ymax} + G_{by|ymax} + G_{ty|ymax}.$$

Then we evaluate the inequality structure by analysing the relations

$$\frac{G_{wy}}{G_y} = \frac{G_{wy|ymin}}{G_{y|ymin}} = \frac{G_{wy|ymax}}{G_{y|ymax}},$$

$$\frac{G_{by}}{G_y} = \frac{G_{by|ymin}}{G_{y|ymin}} = \frac{G_{by|ymax}}{G_{y|ymax}},$$

$$\frac{G_{ty}}{G_y} = \frac{G_{ty|ymin}}{G_{y|ymin}} = \frac{G_{ty|ymax}}{G_{y|ymax}}.$$

When the structure of the decomposed indices $G_{y|ymin}$ and $G_{y|ymax}$ is similar, that is, when the equivalences hold, we get that the underlying inequality factor operates uniformly on y. On the contrary, different structures indicate that particular regions of y are more affected by the inequality factor.

For the analysis of the gender gap we are specifically interested to the equivalences related to G_b, which allow to evaluate more thoroughly the relation between income and gender gap, understanding in particular if gender gap is more related to low or high incomes or if it maintains the same size for any value of y.

3.4.2 The Effect of Further Inequality Factors

The previous approach, where the decomposition obtained by using all n observations is compared to decompositions related to subsamples of observations, can be generalized to include additional inequality factors in the gender gap analysis.

In order to evaluate the influence of an inequality factor x, we can rank y on the values of x, select two subsamples of observations related to particular values of x (usually the lowest and highest) and analyze the decomposition for female and male subgroups.

In this case we analyse the inequality structure by means of the relations

$$\frac{G_{wy}}{G_y} = \frac{G_{wy|xmin}}{G_{y|xmin}} = \frac{G_{wy|xmax}}{G_{y|xmax}},$$

$$\frac{G_{by}}{G_y} = \frac{G_{by|xmin}}{G_{y|xmin}} = \frac{G_{by|xmax}}{G_{y|xmax}},$$

$$\frac{G_{ty}}{G_y} = \frac{G_{ty|xmin}}{G_{y|xmin}} = \frac{G_{ty|xmax}}{G_{y|xmax}}$$

Similar decompositions suggest that the inequality factor x does not provide any additional information on gender gap, while different decompositions indicate a relation between x and the gender gap. In the latter case we are able to exploit the additional information on gender gap provided by x, thus moving gender gap evaluation to a multidimensional framework. Since gender gap is not exclusively related to income, but it has a multidimensional nature, this extension is of particular interest in order to assess the relevance of gender gap.

4 Gender Income Inequality Among Italian Households

The Dagum's decomposition of the Gini index presented in Sect. 2 is extremely useful to analyze the relevance of gender in income inequality. The component G_w allows to evaluate how the income variability existing within the female and male subpopulations influence total inequality, while the contribution attributable to the differences between the female and male subpopulations is given by G_b and G_t. The meaning of G_b is straightforward, but as far as G_t it is useful to point out that high levels of overlapping indicate a small contribution of gender to income inequality, while low levels of overlapping suggest a stronger contribution.

4.1 The Data

The data used in this study are from the Survey on Households Income and Wealth, a multidimensional survey on Italian households performed every two years by the Bank of Italy. The analyses and the results illustrated in this section refer to three waves of the survey: 1993, 2004 and 2014.

Table 1 reports the basic information set for the analysis of income gender gap: the population share, the income share, the average income and the Gini index for the Italian households by gender of the reference person. In the case of the Bank of Italy survey, the reference person is the responsible for the family economy or the person more informed about it. We observe how, over the past 20 years, the proportion of families with a female reference person has increased considerably and today it represents more than a third of families.

Table 1 also includes further information about the income distribution by means of the median $\bar{y}_{me}$ and the coefficient of variation cv, given by the ratio of the standard deviation by the mean. The distance between $\bar{y}$ and $\bar{y}_{me}$ confirms the relevance of the skewness; we can also note how the women income distribution is greater affected by this aspect. Overall the difference $\bar{y} - \bar{y}_{me}$ and the cv point out to relevant differences

Table 1 Population share, income share, mean, median, coefficient of variation and Gini index for Italian household income by gender of the reference person

	1993			2004			2014		
	Female	Male	Tot	Female	Male	Tot	Female	Male	Tot
p	0.275	0.725	1.000	0.302	0.698	1.000	0.352	0.648	1.00
s	0.193	0.807	1.000	0.238	0.762	1.000	0.293	0.707	1.00
$\bar{y}$	14372	22730	20432	23204	32200	29483	25393	33319	30525
$\bar{y}_{me}$	11155	18901	16578	18450	26200	23833	19958	28251	25104
cv	0.794	0.700	0.748	0.734	0.873	0.869	0.871	0.661	0.733
G	0.382	0.342	0.366	0.362	0.339	0.353	0.375	0.328	0.350

between men and women distributions, thus suggesting that the use of an indicator based on pairwise differences between of the observed units, such as the Gini index, should be a preferred choice for income distribution analysis.

4.2 The Gender Gap by Subpopulations Means

The first, basic, traditional evaluation of gender gap is based on the subgroups means $\bar{y}_m$ and $\bar{y}_f$. In Table 2 and Fig. 1 it is reported the absolute difference $(\bar{y}_m - \bar{y}_f)$, as well as some relative differences between $\bar{y}_m$ and $\bar{y}_f$, showing a declining income gender gap for the Italian households from 1993 to 2014. We can observe a decrease both in absolute and in relative values, with gap_2 and gap_3 which fall by 46% and 36% respectively.

The same indication is provided by $gap_4 = (p_f - s_f)$, which is null when the gender gap is equal to 0, while it increases for increasing level of gender gap: from 1993 to 2014 this indicator goes from 0.082 to 0.059, with a decrease of 28%.

Overall the aggregate data of Table 2 suggest the presence of a gender gap, but also its reduction over time.

Table 2 Gender income gap for the Italian households by income means

	gap_1	gap_2	gap_3	gap_4
	$(\bar{y}_m - \bar{y}_f)$	$(\bar{y}_m - \bar{y}_f)/\bar{y}_f$	$(\bar{y}_m - \bar{y}_f)/\bar{y}$	$(p_f - s_f)$
1993	8358	0.582	0.409	0.082
2004	8996	0.388	0.305	0.064
2014	7927	0.312	0.260	0.059

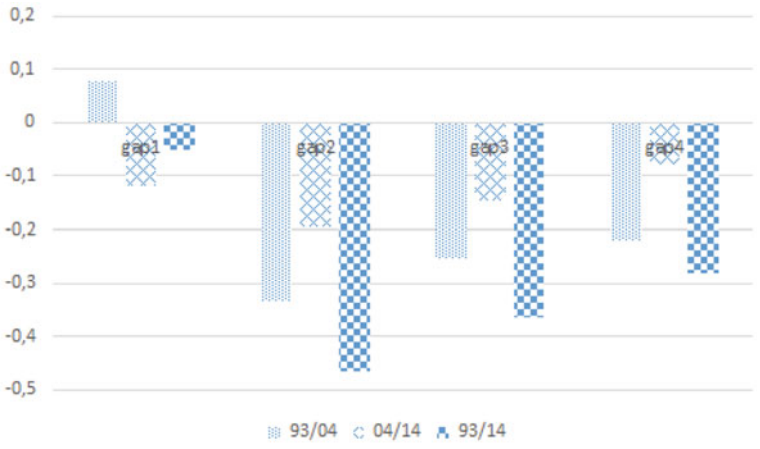


Fig. 1 Gender gap by subpopulations means: variations 1993/2004, 2004/2014, 1993/2014

4.3 The Gender Gap by Inequality Decomposition

Moving from the aggregate and mean-based evaluation provided in the previous Section to the more detailed and accurate information contained on the decomposed Gini index (Table 3 and Fig. 2), we obtain a different picture on gender income inequality.

First, the importance of inequality within on total inequality strongly decreases (from 60% in 1993, to 58% in 2004 and to 54% in 2014), thus indicating a weaker variability within the female and male subpopulations and a stronger effect of the underlying inequality factor, in our case the gender.

Second, the overlapping between the female and male subpopulations increases: the importance of the G_t component rises from 8.7% in 1993, to 11.7% in 2004 and to 14.6% in 2014. A greater overlapping represents a positive signal for the reduction of the gender gap, since it suggests that the distributions of the subpopulations share a larger area.

Third, the inequality between is stable, as the importance of the G_b component goes from 31% in 1993, to 29.9% in 2004 and to 31.5% in 2014.

Table 3 Income inequality decomposition by gender of the reference person

				gap_5		
	Gw	Gb	Gt	Gw/G	Gb/G	Gt/G
1993	0.221	0.113	0.032	0.603	0.310	0.087
2004	0.206	0.105	0.041	0.585	0.299	0.117
2014	0.189	0.110	0.051	0.539	0.315	0.146

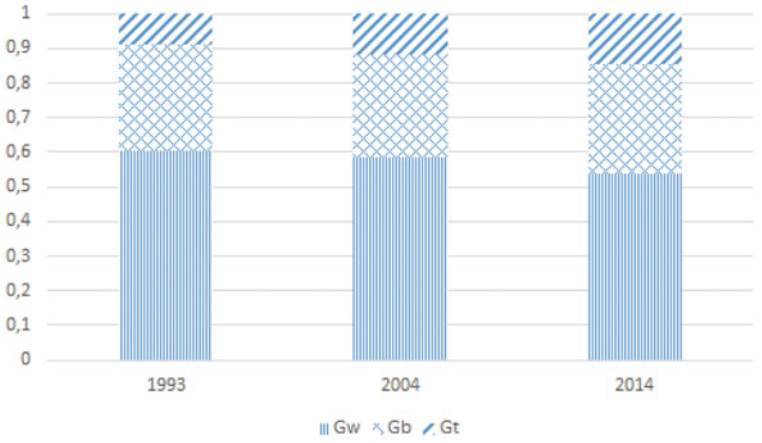


Fig. 2 Decomposed Gini index (male and female subgroups)

Table 4 Population share, income share and Gini index for the Italian households by gender of the reference person

	1993			2004			2014		
	Female	Male	Tot	Female	Male	Tot	Female	Male	Tot
20% bottom income									
p	0.558	0.442	1.000	0.530	0.470	1.000	0.431	0.569	1.000
s	0.548	0.452	1.000	0.510	0.490	1.000	0.426	0.574	1.000
G	0.186	0.211	0.199	0.171	0.157	0.167	0.244	0.214	0.227
20% top income									
p	0.118	0.882	1.000	0.195	0.805	1.000	0.246	0.754	1.000
s	0.112	0.888	1.000	0.179	0.821	1.000	0.244	0.756	1.000
G	0.171	0.180	0.179	0.155	0.218	0.207	0.177	0.168	0.170

Overall, the decrease of inequality within is balanced by the increase of the overlapping component. While a greater G_t alleviates the role of gender as inequality factor, this increase is not sufficient to reduce G_b, which shows a stable gender income inequality during the period 1993–2014. Our empirical findings are in line with existing literature on gender gap, where [15] inequality within accounts for 50% of total inequality and the overlapping component plays a relevant role.

In order to better understand the results of Table 3, we focus on the left and the right tail of the income distribution, taking into account the bottom and the top 20% of the income. Table 4 reports the population share, the income share and the Gini index for the female and male subpopulations for the two cases and it is possible to observe some relevant differences.

Overall, the female subpopulation share p_f increases steadily from 1993 to 2014, but the changes are not homogenous in the two subgroups: p_f decreases in the bottom 20%, while it assumes a far greater relevance for the top 20% incomes. In 1993 the two subgroups are quite different, with the bottom 20% characterized by an unicum $p_f > p_m$ and the top 20% where the female subpopulation represents only a small fraction of the total. 20 years later, in 2014, the two subgroups are more similar, for the bottom 20% is as usually $p_f < p_m$, while for the top 20% the female subpopulation share is doubled since 1993.

We analyze the bottom and the top of the income distribution, with the aim to compare the structure of the decomposed Gini indices.

By comparing the decomposed Gini indexes for the bottom and the top incomes (Table 5), we note that the two decompositions, initially quite different, are more or less similar in 2014. We can attribute the reduction of the importance of G_w, already observed on Table 3, to the 20% top income. The overlapping component, initially stronger for the bottom incomes, represents around the 20% of total inequality for both subgroups, and the increase of its weight is quite evenly distributed between the bottom and top incomes. On the contrary, the stability of the importance of the

Table 5 Income inequality decomposition by gender of the reference person

	Gw/G	Gb/G	Gt/G	Gw/G	Gb/G	Gt/G
	20% bottom income			20% top income		
1993	0.497	0.275	0.227	0.790	0.118	0.083
2004	0.494	0.313	0.193	0.722	0.177	0.101
2014	0.505	0.258	0.237	0.623	0.195	0.182

Table 6 Population share, income share and Gini index for the Italian households by gender of the reference person

	1993			2004			2014		
	Female	Male	Tot	Female	Male	Tot	Female	Male	Tot
Up to elementary school									
p	0.388	0.612	1.000	0.435	0.565	1.000	0.488	0.512	1.000
s	0.289	0.711	1.000	0.348	0.652	1.000	0.402	0.598	1.000
G	0.355	0.323	0.353	0.296	0.279	0.303	0.279	0.276	0.294
With university degree									
p	0.172	0.828	1.000	0.304	0.696	1.000	0.400	0.600	1.000
s	0.143	0.857	1.000	0.236	0.764	1.000	0.331	0.669	1.000
G	0.281	0.301	0.302	0.306	0.338	0.339	0.332	0.322	0.335

inequality between observed on Table 3 is the results of an increase related to the top incomes and a slight decrease observed for the 20% bottom income.

A further analysis of the gender income inequality refers to the study of particular population characteristics: in the following we evaluate the effects related to the educational level. An exhaustive analysis would require the inclusion of many other factors, but would go beyond the objectives of the work, which aims to show the advantages of inequality decomposition in gender gap evaluation.

The Gini index decomposition is applied not to all n observations of the overall population, but only to the subsample of households with the particular characteristic which we are analyzing. More specifically, we compare the female/male decompositions obtained on two subgroups related to two different values of the character under examination. When the two decompositions are substantially similar, the underlying factor is not relevant for the interpretation of the gender inequality, while, on the contrary, different decompositions indicate an influence on gender inequality.

Table 6 illustrates the population share, the income share and the Gini index for two subgroups: for the educational level we compare the up-to-elementary school group to the group with a university degree.

The related decompositions of the Gini index for the analysis of the gender gap are shown in Table 7. The comparison between the decompositions suggests that the educational level strongly influences the gender income inequality. We also confirm

Table 7 Income inequality decomposition by gender of the reference person, Italy 1993–2014

	Gw/G	Gb/G	Gt/G	Gw/G	Gb/G	Gt/G
	Up to elementary school			With university degree		
1993	0.512	0.384	0.104	0.730	0.182	0.088
2004	0.488	0.400	0.113	0.594	0.305	0.101
2014	0.473	0.410	0.117	0.517	0.345	0.138

the decrease of the importance of the inequality within, together with an increase of the relevance of the overlapping component and of the inequality between, especially for the more affluent subgroup.

5 Conclusions

The decomposition of an inequality index allows powerful insights on the inequality structure and can be extremely useful into the study of the gender income inequality, where the decomposition refers to the female and male subpopulations.

We develop an unified framework for the gender gap analysis, where the approach based on the subgroups means is linked to the methods related to the inequality decomposition. We show how the two approaches are equivalent when female and male income distributions are not overlapping, while, in case of overlapping, income means lead to underestimate the gender gap.

On the basis of the information provided by the average incomes, and without taking into account the overlapping between female and male income distributions, we obtain for the Italian case a declining gender gap from 1993 to 2014.

Within a framework based on the Gini index decomposition, which considers all distributional characteristics, gender gap explains 31% of total inequality in 1993, a level which remains stable until 2014.

Inequality decomposition also allows a better understanding on the sources of the gender gap: we analyze the role of educational level, which, for Italian households, remains a main driver of gender income inequality, and we are also able to link our results on gender gap to the right part of the income distribution.

Our proposal allows a more complete and effective assessment of gender gap, able to provide useful indications both for the advancement of the theoretical debate and the implementation of successful socio-economic policies.

References

1. Amici, M., Stefani, M.L.: A gender equality index for the Italian regions. Questioni di Economia e Finanza, Banca d'Italia **190** (2013)
2. Atkinson, A.B., Casarico, A., Voitchovsky, S.: Top incomes and the gender divide. J. Econ. Inequal. **16**, 225–256 (2018)
3. Bhattacharya, B., Mahalanobis, B.: Regional disparities in household consumption in India. J. Am. Stat. Assoc. **62**, 143–161 (1967)
4. Bonnet, C., Meurs, D., Rapport, B.: Gender inequalities in pensions: different components similar levels of dispersion. J. Econ. Inequal. **16** (2018)
5. Costa, M.: Gini index decomposition for the case of two subgroups. Commun. Stat. **37**, 631–644 (2008)
6. Cupak, A., Fessler, P., Schneebaum, A., Silgoner, M.: Decomposing gender gaps in financial literacy: new international evidence. Econ. Lett. **168** (2018)
7. Dagum, C.: Gini ratio. In: The New Palgrave Dictionary of Economics. Mac Millian Press, London (1987)
8. Dagum, C., Zenga, M.: Income and Wealth Distribution Inequality and Poverty. Springer, Berlin (1990)
9. Dagum, C.: A new decomposition of the Gini income inequality ratio. Empir. Econ. **22**, 515–531 (1997)
10. Frosini, B.V.: Approximation and decomposition of Gini, Pietra-Ricci and Theil inequality measures. Empir. Econ. **43**, 175–197 (2012)
11. Gender Equality Index Report, European Institute for Gender Equality (2013)
12. Giorgi, G.M.: Gini's scientific work: an evergreen. Metron **63**, 299–315 (2005)
13. Giorgi, G.M.: The Gini inequality index decomposition. An evolutionary study. In: Deutsch, J., Silber, J. (eds.), The Measurement of Individual Well-Being and Group Inequalities, Routledge, London (2011)
14. Goraus, K., Tyrowicz, J., van der Velde, L.: Which gender gap estimatest to trust? A comparative analysis. Rev. Income Wealth **63**, 118–146 (2017)
15. Larraz, S.: Decomposing the Gini inequality index: an expanded solution with survey data applied to analyze income inequality. Sociol. Methods Res. **44**, 508–531 (2015)
16. Mussida, C., Picchio, M.: The gender wage gap by education in Italy. J. Econ. Inequal. **12**, 117–147 (2014)
17. Pittau, M.G., Yitzhaki, S., Zelli, R.: The "make-up" of a regression coefficient: gender gaps in the European labor market. Rev. Income Wealth **61**, 401–421 (2015)
18. Schneebaum, A., Rehm, M., Mader, K., Hollan, K.: The gender wealth gap across European countries. Rev. Income Wealth **64**, 295–331 (2018)
19. Selezneva, E., Van Kerm, P.: A distribution-sensitive examination of the gender wage gap in Germany. J. Econ. Inequal. **14**, 21–40 (2016)
20. Yitzhaki, S., Lerman, R.: Income stratification and income inequality. Review Income Wealth **37**, 313–329 (1991)

Functional Linear Models for the Analysis of Similarity of Waveforms

Francesca Di Salvo, Renata Rotondi, and Giovanni Lanzano

Abstract In seismology methods based on waveform similarity analysis are adopted to identify sequences of events characterized by similar fault mechanism and propagation pattern. Seismic waves can be considered as spatially interdependent, three dimensional curves depending on time and the waveform similarity analysis can be configured as a functional clustering approach, on the basis of which the membership is assessed by the shape of the temporal patterns. For providing qualitative extraction of the most important information from the recorded signals, we propose the use of metadata, related to the waves, as covariates of a functional response regression model. The temporal patterns of this effects, as well as of the residual component, obtained after having taken into account the most relevant predictors, are investigated in order to detect a cluster structure. The implemented clustering techniques are based on functional data depth.

Keywords Functional response regression · Structured functional principal component · Waveforms clustering · Functional data depth

1 Introduction

When an earthquake occurs, shockwaves of energy, called seismic waves, are released from the earthquake source and travel to a receiver via some path through the subsurface. Seismometers, hydrophones (in water) or accelerometers, measuring motion of

F. Di Salvo (✉)
Department SAAF, University of Palermo, Palermo, Italy
e-mail: francesca.disalvo@unipa.it

R. Rotondi
CNR IMATI, Milan, Italy
e-mail: reni@mi.imati.cnr.it

G. Lanzano
INGV, Milan, Italy
e-mail: giovanni.lanzano@ingv.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_9

the ground and its acceleration (the rate of the change of the velocity) are arranged in networks recording a large number of signals.

From the statistical point of view, the problem of investigating the seismotectonic structures of an area involves several methods based on the analysis of the similarity of the waves recorded. In particular, seismic networks often record signals characterized by similar shapes and methods studying their similarity are adopted to identify sequences of foreshock, main shock and aftershock; in this field the goal is the definition of group of events characterized by similar fault mechanism and propagation pattern, under the hypothesis that a group of dependent events (multiplets) represents a chain led by seismogenetics background of a common earthquake.

The detection of earthquake families or multiplets is finalized to the identification of sources related to the same fault [9] or to obtain instrumental catalogues similarity of independent earthquakes cleaned of dependent ones [3].

Statistical approaches are powerfull tools for detecting dependent events in a seismic data set; waveform similarity analysis is considered to join seismic episodes into a single multiplet [3]; clustering has been demonstrated as a useful method for identifying members of the same group that possess similar waveform. Different techniques for assessing the cluster membership of a earthquake are also reported in literature [1, 14].

A seismic events is characterized by a number of attributes, that make up the set of integrating metadata concerning the source, the localization of the recording, the dynamics of the registration. In this paper, the methdology integrates information from waveforms with information from metadata. For what concerns seismic waves, a common opinion is that this class of data are also noisy and most techniques, including clustering, can be optimized by using appropriate data preprocessing. Signal filtering, as singular value decomposition as well as short-time Fourier transforms (STFT), are recognized as proper techniques for extracting the key features [7, 15]. The filtered signal are analyzed through a functional regression model in order to estimate the effects of some covariates. A clustering procedure is applied to the functional residuals obtained by subtracting the predicted from the signals, to ascertain whether similar structures persist on waves, after having removed the effects of the main covariates. In Sect. 2, estimate functional structure of the waves is described as well as the set of metadata. In Sect. 3, the methodology is presented: in Sect. 3.1 functional regression models on scalar covariates are described. Within this framework, the Karhunen-Loéve expansion of the effects is applied to detect valuable information from the waveforms. In Sect. 3.2 an algorithm based on the modified band depth [10] is proposed for clustering functional residuals waves from linear models. In Sect. 4, the approach is applied to a set of recordings of the main four seismic events of the sequence Amatrice - Norcia - Visso, from August 2016 to January 2017. Discussion and concluding remarks are in Sect. 5.

2 The Data

Tha data analyzed in this paper come from ESM, the Engineering Strong-Motion Database, a source for ground-motion data and associated metadata, relatively to seismic events with magnitude greater than 4.0 [11].

When an earthquake occurs, shockwaves of energy, called seismic waves, are released from the earthquake focus; a network of sismographs records a large number of seismic waves: records are taken in three directions: up-down, east-west and north-south directions. As result, each wave is characterized by three interdependent curves, recorded onto a common time scale (200 Hz), respectively named vertical (Up-Down), longitudinal (East-West) and transversal (North-South).

In our analysis the fundamental observational unit is a seismic wave, a three-dimensional signal recorded in few seconds, whose structure is functional as it depends on time, as in Fig. 1.

Seismic waveform data are stored, on the basis of international seismological standards, with the inclusion of comprehensive meta information, such as event (the

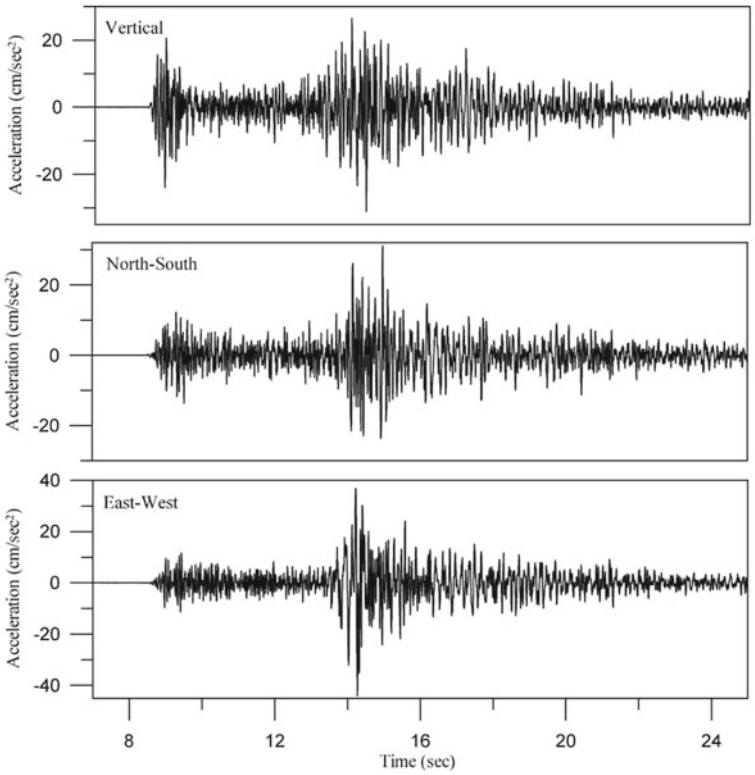


Fig. 1 Seismic wave recorded in three directions

Table 1 Types of data (source, receiver and waveforms information)

Event	Seismograph	Waveform
Latitude/Longitude degree, azimuth degree, time, event depth, magnitude	Latitude/Longitude degree, site class, morphologic class, epicentral distance	PGA, time PGA, direction, frequency, duration, acceleration

source) or seismograph (the receiver) information: metadata can be arranged in data concerning the generating seismic event, as magnitude, and data concerning each single wave, as peak ground acceleration. Another set of attributes is referred to the seismographs; the complete list of attributes is in Table 1. These characteristics are generally taken into consideration by scientists for studying ground motions in specific area.

From a statistical point of view, the volume and the complexity of this data is an enormous challenge for the analysis.

3 The Methodology

3.1 The Model

Seismic waves, that are three dimensional spatially interdependent curves, can be considered as realizations of a multivariate functional random field:

$$Y^p(t) = f^p(t) + \varepsilon^p(t)$$

The couple $(t, Y^p(t))$ denotes the time and the function $Y^p(t)$ at time t and $t \in [0, T]$. Standardizing the time interval in [0, 1]:

$$Y^p = \{Y^p : [0, 1] \to R\}, \quad p = 1, 2, 3$$

is the set of real-valued functions on the closed interval [0, 1].

The curves are indicized by a couple of indexes (i, s) indicating the recordings $Y^p_{is}(t)$ of the ith seismic event (or source) at sismograph (or receiver) $s, i = 1, \ldots, I$, $i = 1, \ldots, S$:

$$Y^p_{11}(t), \ldots Y^p_{is}(t) \ldots, \quad Y^p_{is} \in Y^p, \ t \in [0, 1]$$

The aim is to look for differences in temporal dynamics, taking into account some covariates; for simplicity we consider one covariate referred to the source and one to the receiver.

Effects of these covariates on the dynamics of the waves are modeled in a linear regression model with scalar covariates and functional response [13]:

$$Y_{is}^{p}(t) = X_i \beta_1^p(t) + Z_s \beta_2^p(t) + \varepsilon_{is}^p(t) \tag{1}$$

Each wave is decomposed into the linear combination of functional regression parameters $\beta_1^p(t)$ and $\beta_2^p(t)$ measuring the partial effect of events and receivers covariates on the response at position t. X_i and Z_s are the scalar covariates and $\varepsilon_{is}^p(t)$ the residual component. There is a large body of literature dealing with this class of models and further variants [6, 12, 16, 17]: non-parametric regression handling non-linear and non-additive association structures [5]; functional regression models of pre-smoothed curves [13]; projections into a coefficient space for a given set of basis functions and subsequent modelling on the transformed space [12]; structured Functional Principal Component Analysis with decomposition of covariance structure of the latent processes [16].

Regularization of the functional coefficients $\beta_j^p(t)$ $(j = 1, 2)$ potentially increases estimation and precision accuracy [12] and in the structured additive regression model of the Eq. (1) it is typically achieved using either truncation and roughness penalties in some chosen basis space:

$$\hat{Y}^p(t) = X \sum_{l=1}^{L_1} \phi_l^1(t)\theta_{1l}^p + Z \sum_{l=1}^{L_2} \phi_l^2(t)\theta_{2l}^p \tag{2}$$

Each effect is specified by a linear expansion in L_j basis and coefficients, for $j = 1, 2$, where $\Phi^1 = \left[\phi_1^1, \phi_2^1, \ldots \phi_{L_1}^1\right]$ is the finite basis system spanning the time interval for the event-covariate and $\Phi^2 = \left[\phi_1^2, \phi_2^2, \ldots \phi_{L_2}^2\right]$ is the basis system spanning the time interval for the receiver-covariate. The coefficients $\Theta_j^p = \left[\theta_{j_1}^p, \theta_{j_2}^p, \ldots, \theta_{j_{Lj}}^p\right]$, for $j = 1, 2$, represent a finite set of parameters measuring the partial effects of predictor j on the response p.

A generalization of this model admits functional predictors, as discussed in [12] and [13], using truncated basis function expansions for the predictors; in this case the effects of covariates are represented by mean of a row-wise tensor product of two marginal bases. Let $\Phi^z(z)$ and $\Phi^x(x)$ be the basis systems spanning the domain of the covariates X and Z respectively; defining $\Phi^\eta(t) = \left[\phi_1^\eta, \phi_2^\eta, \ldots \phi_{c_1}^\eta\right]$ and $\Phi^\gamma(t) = \left[\phi_1^\gamma, \phi_2^\gamma, \ldots \phi_{c_2}^\gamma\right]$ as the row-wise tensor products of two marginal bases matrix [8]:

$$\Phi^\eta(t) = \Phi^x(x) \square \Phi^1(t)$$

$$\Phi^\gamma(t) = \Phi^z(z) \square \Phi^2(t)$$

the resulting linear representation of the functional response is:

$$\hat{Y}^p(t) = \sum_{l=1}^{c_1} \phi_l^{\eta}(t)\theta^{*p}_{1l} + \sum_{l=1}^{c_2} \phi_l^{\gamma}(t)\theta^{*p}_{2l}$$

with a different set of parameters $\Theta^{*p}_j = \left[\theta^{*p}_{j_1}, \theta^{*p}_{j_2}, \ldots, \theta^{*p}_{j_{cj}}\right]$, for $j = 1, 2$.

In this paper considering scalar predictors, we refer to Eq. (2).

Identifiability and computational efforts are the challenging points of the model presented in Eq. (2) [6]; a solution is based on functional principal components, using the Karhunen-Loéve expansion for the effects $X\beta_1^p(t)$ and $Z\beta_2^p(t)$ the model (2) becomes:

$$\hat{Y}^p_{is}(t) = \hat{\mu}^p(t) + \sum_{k=1}^{\infty} \xi_k^{X^p}(t)\delta_{ik}^{X^p} + \sum_{k=1}^{\infty} \xi_k^{Z^p}(t)\delta_{jk}^{Z^p} \tag{3}$$

where $\xi_k^{X^p}(t)$ and $\xi_k^{Z^p}(t)$ are respectively the kth eigenfunctions of the two expansions, and $\delta_k^{X^p}$, $\delta_k^{Z^p}$ are the kth vectors of principal scores, i.e. the projections of the effects on the space of the principal components.

A basis truncation, taking into account only the first principal components, is common to explain a high percentage of the total variability; considering the case when most variability of each latent process is captured by the first m_1 and m_2 principal components, the structured covariance matrices for the effects in Eq. (3) are derived in [16]. The two functional processes in Eq. (3) are characterized by the observed variability; their covariance operator can be estimated using method of moments from the estimated curves, [16]. Principal component decomposition of the respective covariance operators allows the description of the effects of the two attributes on the pattern of the waves: other structures on the data could be detected on residuals.

3.2 Analysis of Residuals Through Functional Data Depth

The residual functions are obtained by subtracting point to point the predicted in Eq. (3) from the observed response functions in Eq. (1).

The proposal consists in clustering the residual functions with similar structures on the basis of the concept of data depth. Depth measures the centrality of an observation within a sample and allows the definition of a natural ordering from center outwards; several depth notions generalize unidimensional concept. Here we focus on Modified Band Depth [10].

Given the set of n continuous residuals functions, $f(t)$ in [0, 1] (the double indicization i, s is not required here) and λ, a Lebesgue measure in [0, 1], for any f of the sample, the *Modified Band Depth* is the portion of time that $f(t)$ is inside the regions, made up of $2, 3, \ldots, k, \ldots, K$ of the n curves:

$$MBD_n(f) = \sum_{k=2}^{K} n_{(k)}^{-1} \sum_{1<\ldots r_1, r_2 \ldots<n} \frac{\lambda(A(f;\, f_h r1,\, f_{r2}))}{\lambda(T)} \tag{4}$$

where $A(f;\, f_{r1},\, f_{r2})$ the region delimited by f_{r1} and f_{r2}, is defined as:

$$A(f;\, f_{r1},\, f_{r2}) = \left\{ f(t) : min_{r=r_1,r_2} f_r\,(t) \leq f(t) \leq max_{r=r_1,r_2} f_r\,(t) \right\}$$

The curves are ordered on the basis of the depth MBD in Eq. (4), deepest curve having the maximum value. The underlying idea is to determine clusters of ordered curves having the maximum cohesiveness; two previous paper [2, 4] describe the algorithm based on the notion of Modified Band Depth. In this paper the algorithm is adopted for clustering the curves and the principal steps of the algorithm are resumed:

1. First step:
   a. The $center \rightarrow outward$ ordering provided by MBD_n is exploited for the residual components $f(t)$
   b. Given α, $(0 \leq \alpha \leq 1)$, the $n_1 = (1-\alpha)n$ of deepest curves are grouped in a first cluster C_1, $n_1 = |C_1|$.
   c. The αn most external curves make up the cluster C_2, $|C_2| = \alpha n = n2$.
2. Iterative step:
   a. The members of the existing clusters are separated in two groups: the kernels, i.e. the $(1-\alpha)100\%$ of deepest functions, and the group of the outer functions;
   b. each of the $\alpha\%$ of outer curves, can be allocated into one of the previous clusters, on the basis of the maximized MBD.
   c. Repeat the iterative step until the allocation of the curves improves in terms of increasing MBD
3. Stopping rule:
   Stop when each curve has the highest MBD within its current cluster and no other allocation improves the the MBD of the curves; in this case the optimal configuration is reached.
   Alternatively, stop when a maximum number of iterations is reached.

The goodness of the configuration obtained is measured by the reciprocal of the areas of the kernels, that can be interpreted as a measure of the clusters cohesiveness.

4 The Application

The data analyzed in this paper come from ESM, the Engineering Strong-Motion Database and consists in the set of recordings of four seismic events of the sequence "Amatrice-Norcia-Visso", from August 2016 to January 2017, with epicentres Accumuli, Ussita, Norcia, Pizzoli. The seismic events and a summary of the principal

Table 2 Seismic sequence in Amatrice-Norcia-Visso (2016–17)

Epicentre and date of the events	Latitude	Longitude	$Magnitude_L$	Number of recordings
Accumoli, 2016-08-24	42.6983	13.2335	6.0	14
Ussita, 2016-10-26	42.9087	13.1288	5.9	11
Norcia, 2016-10-30	42.8322	13.1107	6.5	8
Pizzoli, 2017-01-18	42.5293	13.2823	5.5	12

Table 3 Summary of Metadata

	Range
Event depth km	7.50–9.20
Epicentral distance km	51.30–98.50
Backazimuth degree	0.80–357.90
PGA cm.sec^2	−57.11–102.52
Time PGA s	15.27–59.82
Duration s	43.18–202.86

metadata are presented in Tables 2 and 3. In Fig. 2 the map of the seismic events and the position of seismographs is reported.

A pre-processing step is generally recommended before implementing the analysis; here data are pre-processed by a filtering the original signals, through discrete Short Time Fourier Transform (STFT), separately on the three spatial directions, U-D, E-W, N-S. The algorithm splits the time interval $[0, T]$ into frames, then computes the Fourier transform of the observed functions to determine the sinusoidal frequency and phase content of local sections of the signal as it changes over time. A representation of the spectrum on frequencies and time domain is in Fig. 3.

The partition with the minimum number of frames, retaining at least 75% of the whole variability, is selected; this criterion allows to cut the signals at the same length, sampling it at 10 Hz. Filtered waves are represented in Fig. 4a. In order to estimate specific effects in the dynamics of the filtered waves, the event and the epicentral distance, are considered as covariates of the functional regression model.

Following Eq. (3), the first eigenfunctions $\xi_k^{X^p}(t)$, $k = 1, ..., 4$ for the first covariate, and the first eigenfunctions $\xi_k^{Z^p}(t)$, $l = 1, ..., 3$ for the second covariate, are represented Fig. 4, respectively in (b) and (c).

The estimated waves reveal four different dynamics of the signals for the four main events in Fig. 5.

The last part of the application focuses on the functional residuals, obtained subtracting estimated waves from filtered waves. The modified band depth is computed

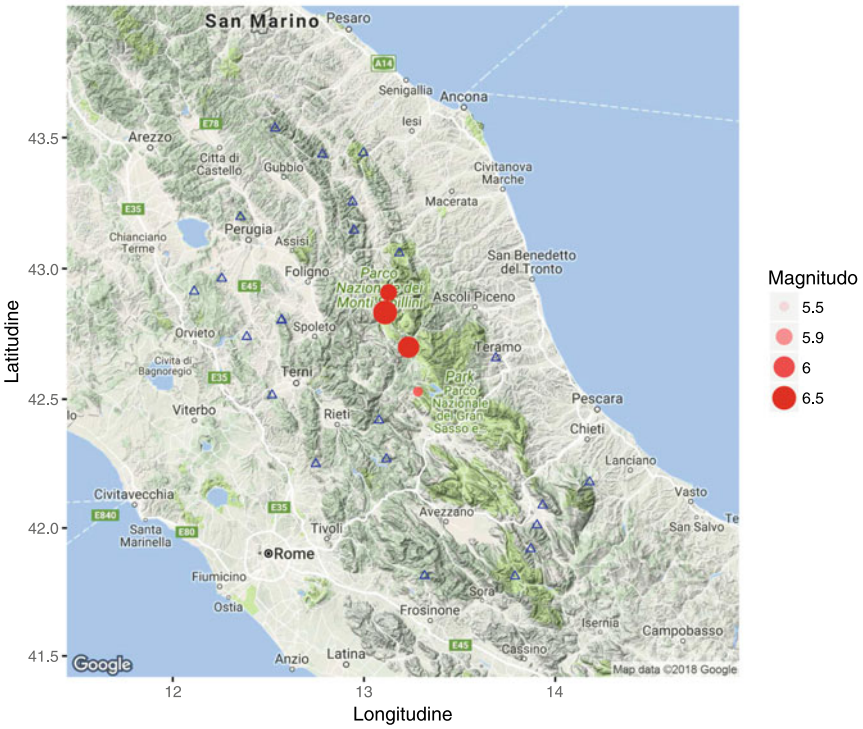


Fig. 2 Geographic coordinates of the seismic events (red points) and seismographs (blue triangle points)

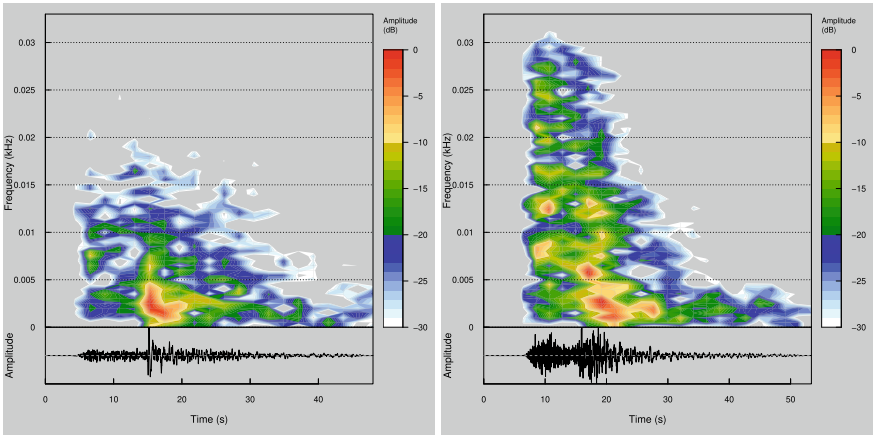


Fig. 3 Discrete STFT representation of two waves

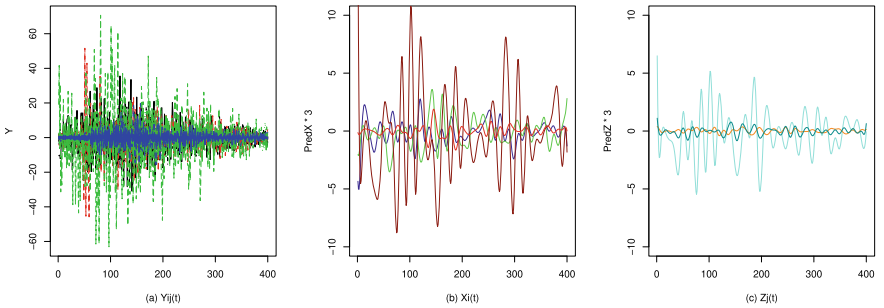


Fig. 4 Filtered waves (**a**) and estimated eigenfunctions (**b**) and (**c**)

as depth measure for ordering the residuals. In Fig. 6 (top) the kernel, consisting in 50% of inner residuals curves (the deepest curve is the red line) is separated from the 50% of the most external ones, Fig. 6 (bottom); it is evident that the two groups differ more for the amplitude variability than for the shape. The residuals curves for the four main events are in Fig. 7.

This initial configuration with two clusters is the input of the Modified Band depth algorithm decribed in Sect. 3.2. At each step, the algorithm finds an intrinsic order inside the clusters of the current configuration and it decides whether one of the most external functions should be reallocated in a different cluster on the basis of its maximum value of depth.

The final configuration has five clusters and it is optimal in the sense that the each residual curves is allocated in the cluster in which it achieves the highest depth, Fig. 8. The deepest curves of the clusters are also compared in the bottom-right plot.

For each cluster, on the basis of the order defined by the depth MBD, it is possible to measure, for each p [0, 1], the area containing the proportion p of most central samples.

The area is a function of p and it is a measure of the dispersion of the clusters (the smaller is the value, the higher is the cohesiveness of the clusters); in Fig. 9 plotting the values of the areas for each p, the scale curve graph obtained shows the dispersion inside the clusters as p increases.

5 Results and Discussion

The present study investigates functional response regression on scalar predictors in the field of the waveforms similarity. In our proposal, using metadata as predictors, some of the most relevant characteristics of the waves can be described in terms of the effects of some attributes concerning the seismic events, the seismographs or the wave itself. The proposed approach is an adaptive data-driven method, implementing functional principal components techniques that can be easily extended to qualitative predictors, in two-way or m-way models. The complementary clustering algorithm,

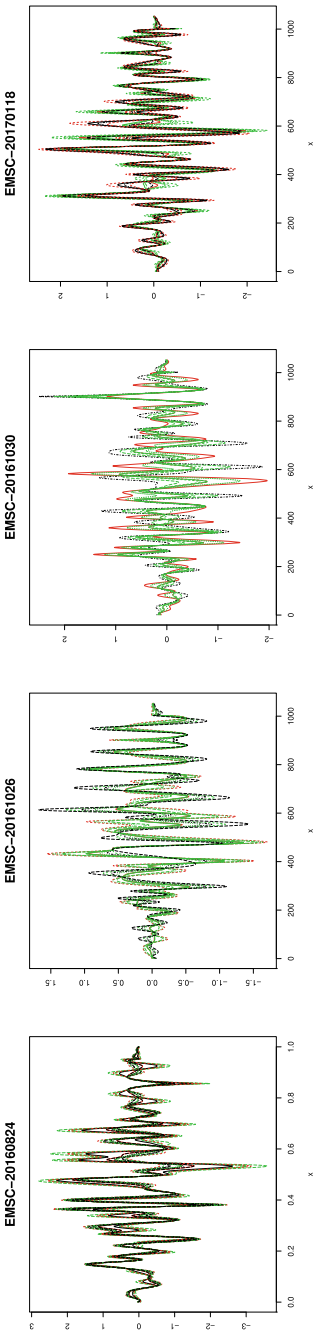


Fig. 5 Estimated waves (FPCA approach)

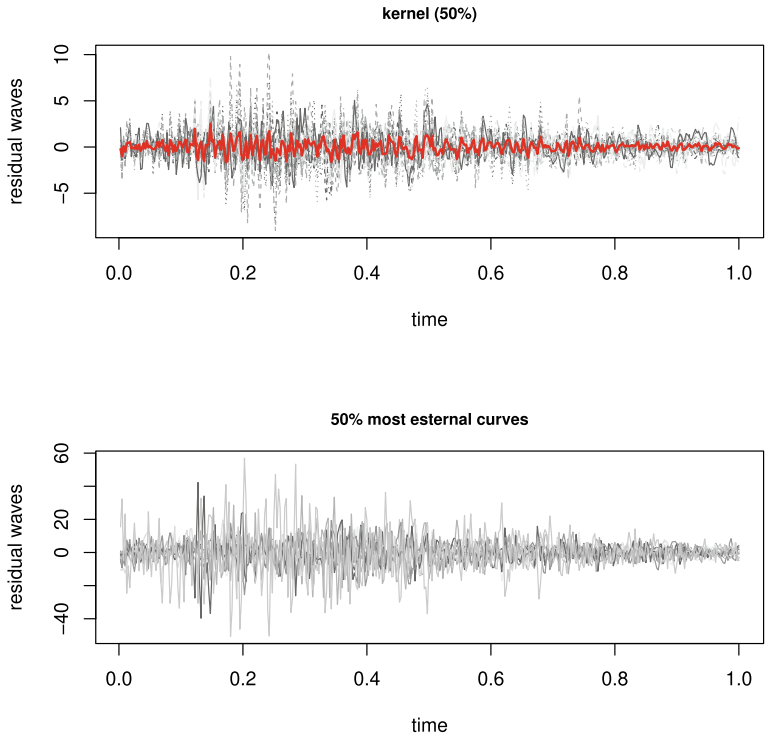


Fig. 6 Ordered residual curves

on the residual part of the models, exploits functional depth measures and adapts them to identify clusters of waves with similar temporal pattern. The identification of the presence of these clusters is a results of interest in seismology, where the waveform similarity analysis is adopted to identify multiplets, i.e. groups of dependent events, and to isolate main events. In the application, the procedure is applied to explore similarity of seismic waves recorded by seismographs of central Italy for four events, after having taken into account the most relevant predictors. The results obtained for the data analyzed are encouraging, but further research is necessary to test the reliability of the proposed technique taking into account several sequences and/or wider magnitude range.

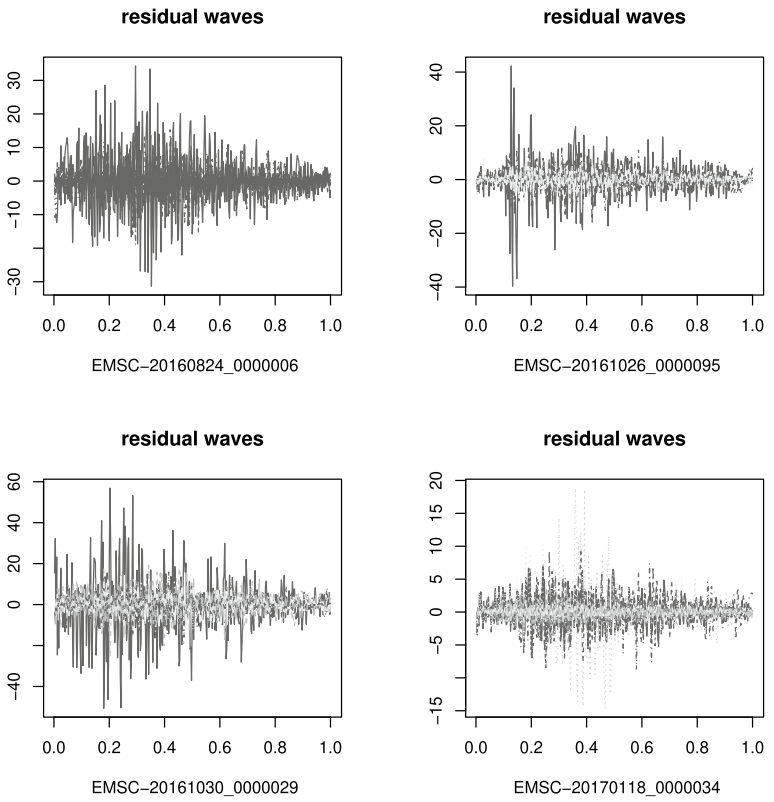


Fig. 7 Residual curves of the four main events

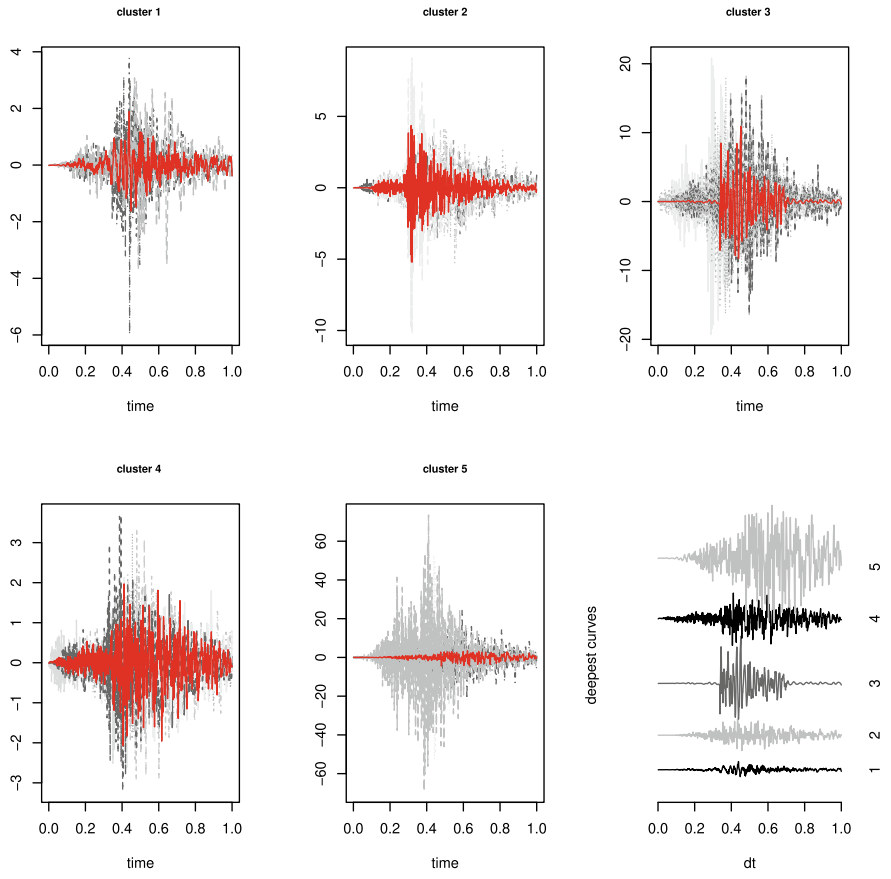


Fig. 8 Optimal configuration

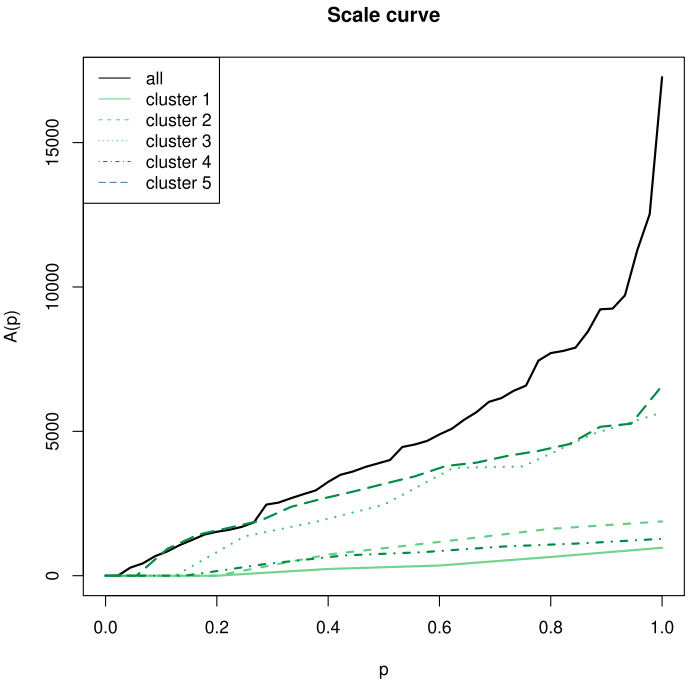


Fig. 9 Scale curves of the configuration: Area of the band containing a percentage p of inner functions

References

1. Adelfio, G., Chiodi, M., D'Alessandro, A., Luzio, D., D'Anna, G., Mangano, G.: Simultaneous seismic wave clustering and registration. Comput. Geosci. **8**(44), 60–69 (2012)
2. Adelfio G., Di Salvo F., Sottile G.: Depth-based methods for clustering of functional data TIES 2017 Conference, Bergamo, Italy, July 24th - 26th (2017)
3. Barani, S., Ferretti, G., Massa, M., Spallarossa, D.: The waveform similarity approach to identify dependent events in instrumental seismic catalogues. Geophys. J. Int. (2006). https://doi.org/10.1111/j.1365-246X.2006.03207.x
4. Di Salvo, F., Rotondi, R., Lanzano, G.: Detecting clusters in spatially correlated waveforms. In: GNGTS Conference, Trieste, November 13th–16th (2017)
5. Ferraty, F., Vieu, P.: Nonparametric Functional Data Analysis: Methods, Theory, Applications and Implementations. Springer, London (2006)
6. Greven, S., Scheipl, F.: A general framework for functional regression modelling. Stat. Model. **17**(1–2), 1–35 (2017). https://doi.org/10.1177/1471082X16681317
7. Hao-kun, D., Jun-xing, C., Ya-juan, X., Xing-jian, W.: Seismic facies analysis based on self-organizing map and empirical mode decomposition. J. Appl. Geophys. **112**, 52–61 (2015)
8. Hastie,T., J., Tibshirani, R. J. Generalized Additive Models, Chapman and Hall/CRC Monographs on Statistics and Applied Probability, (1990)
9. Jagla, E.A., Kolton, A.B.: A mechanism for spatial and temporal earthquake clustering. J. Geophys. Res. (2010). https://doi.org/10.1029/2009JB006974
10. Lopez-Pintado, S., Romo, J.: Depth-based inference for functional data. Comput. Stat. Data Anal. **51**(10), 4957–4968 (2007)

11. Luzi L., Puglia R., Russo E., D'Amico M., Felicetta C., Pacor F., Lanzano G., Çieken U., Clinton J., Costa G., Duni L., Farzanegan E., Gueguen P., Ionescu C., Kalogeras I., Özener H., Pesaresi D., Sleeman R., Strollo A., Zare M.: The Engineering strong-motion database: a platform to access pan-European accelerometric data. Seismol. Res. Lett. **87**(4), 987–997 (2016). https://doi.org/10.1785/0220150278
12. Morris , J.S.: Functional regression. Ann. Rev. Stat. Appl. **2**(1) (2014). https://doi.org/10.1146/annurev-statistics-010814-020413
13. Ramsay J., Silverman, B.: Functional Data Analysis, 2nd edn. Springer, New York (2005)
14. Reasenberg, P.: Second-order moment of Central California seismicity, 1969–1982. J. Geophys. Res. **90**, 5478–5495 (1985)
15. Silvestrov, I., Tcheverda V.: SVD analysis in application to full waveform inversion of multicomponent seismic data. J. Phys.: Conf. Ser. **290** (2011). https://doi.org/10.1088/1742-6596/290/1/012014
16. Shou, H., Zipunnikov, V., Crainiceanu, C.M., Greven, S.: Structured functional principal component analysis. Biometrics **71**(1), 247–257 (2015). https://doi.org/10.1111/biom.12236
17. Suk, H.W., Hwang, H.: Functional generalized structured component analysis. Psychometrika **81**(4), 940–968. https://doi.org/10.1007/s11336-016-9521-1

Capturing Measurement Error Bias in Volatility Forecasting by Realized GARCH Models

Richard Gerlach, Antonio Naimoli, and Giuseppe Storti

Abstract This paper proposes generalisations of the Realized GARCH model, in three different directions. First, heteroskedasticity of the noise term in the measurement equation is modelled letting the variance of the measurement error to vary over time as a function of an estimator of the Integrated Quarticity obtained from intra-daily returns. Second, to account for attenuation bias effects, volatility dynamics are allowed to depend on the accuracy of the realized measure letting the response coefficient of the lagged realized measure be a function of the time-varying variance of the volatility measurement error. Therefore, the model tends to assign more weight to lagged volatilities when they are measured more accurately. Finally, a further extension is proposed by introducing an additional explanatory variable into the measurement equation, aiming to quantify the bias due to the effect of jumps.

Keywords Realized GARCH · Realized volatility · Realized quarticity · Attenuation bias · Heteroskedasticity · Jumps

1 Introduction

The use of realized volatility measures has led to new approaches in fitting and forecasting daily volatility. In a GARCH-type framework, the main idea is to replace a noisy volatility proxy, such as the squared daily returns, with a more efficient realized measure or to jointly use both low (daily returns) and high (realized measures) frequency information in modelling the dynamics of the conditional variance of returns. The HEAVY model of [16] and the Realized GARCH model of [9] provide

R. Gerlach
The University of Sydney, Business School, Darlington, Australia
e-mail: richard.gerlach@sydney.edu.au

A. Naimoli · G. Storti (✉)
Dipartimento di Scienze Economiche e Statistiche (DISES), Università di Salerno, Fisciano, Italy
e-mail: storti@unisa.it

A. Naimoli
e-mail: anaimoli@unisa.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_10

a convenient structure for the joint modelling of daily returns and realized volatility measures.

However, realized measures are noisy estimates of the underlying integrated variance, generating a classical errors-in-variables problem. This typically leads to the rise of what is usually known as attenuation bias with the realized measure being less persistent than the latent integrated variance. Some recent works have highlighted how correcting for this attenuation bias can potentially lead to better volatility forecasts. For example, [6] found that, letting the volatility persistence depend on the estimated degree of measurement error, through an interaction term between realized volatility and the square root of realized quarticity, leads to significant improvements in the predictive performance of HAR models. Similarly, [17] found evidence that, in a GARCH-X model, the magnitude of the response coefficients associated with different realized volatility measures is related to the quality of the measure itself. Finally, [8] observe that the response of the current conditional variance to past unexpected volatility shocks is negatively correlated with the accuracy of the associated realized volatility measure.

In this paper, we propose a novel modelling approach that accounts for the attenuation bias effect in a natural and fully data-driven way. To this purpose, we first extend the standard Realized GARCH model of [9] by letting the variability of the measurement error vary over time as a function of an estimator of the integrated quarticity of intra-daily returns. In this way, we obtain a model-based time-varying estimate of the accuracy of the realized measure used. Consequently, we adjust the volatility dynamics for attenuation bias effects by allowing the response coefficient of the lagged realized volatility to depend on this quantity. In particular, the model is designed so that more weight is given to lagged volatilities when these are more accurately measured. Finally, the proposed modelling approach is further extended to explicitly model the impact of jumps on the predicted conditional variance of returns. This is achieved by introducing into the measurement equation an additional component that controls for the amount of bias generated by the occurrence of jumps. A notable feature of the proposed solution is that the jump correction only occurs on days in which jumps are effectively observed, while resorting to the use of more efficient standard measures, such as realized variances and kernels, in jump-free periods.

The paper is organised as follows. In Sect. 2 the basic theoretical framework behind the computation of realized measures is reviewed, while Sect. 3 discusses the Realized GARCH model of [9]. Section 4 illustrates the proposed time-varying parameter heteroskedastic Realized GARCH model. A jumps-free setting is considered first, then a modification of the proposed model, explicitly taking into account the impact of jumps, is introduced. Quasi Maximum Likelihood (QML) estimation of the proposed models is discussed at the end of the same section. Sections 5–7 are dedicated to the empirical application. Section 5 presents the main features of the data used for the analysis; Sect. 6 focuses on the in-sample performance of the proposed models, taking the standard Realized GARCH model as a benchmark, whereas the out-of-sample forecasting performance is analysed in Sect. 7. Finally, Sect. 8 concludes.

2 A Brief Review of Realized Measures

Let the logarithmic price p_t of a financial asset follow the stochastic differential process

$$dp_t = \mu_t dt + \sigma_t dW_t + dJ_t \qquad 0 \leq t \leq T \ , \tag{1}$$

with μ_t and σ_t the drift and instantaneous volatility processes, respectively, while W_t is a standard Brownian motion; σ_t is assumed to be independent of W_t and J_t is a finite activity jump process. Under assumption of jump absence ($dJ_t = 0$) and a frictionless market, the logarithmic price p_t follows a semi-martingale process. In this case, the Quadratic Variation (QV) of log-returns $r_t = p_t - p_{t-1}$ coincides with the Integrated Variance (IV): $IV_t = \int_{t-1}^{t} \sigma_s^2 ds$. The IV can be consistently estimated by summing intra-daily squared returns [4] by defining the Realized Volatility (RV)

$$RV_t = \sum_{i=1}^{M} r_{t,i}^2 \ , \tag{2}$$

where $r_{t,i} = p_{t-1+i\Delta} - p_{t-1+(i-1)\Delta}$ is the ith Δ-period intraday return, $M = 1/\Delta$. Thus, for $\Delta \longrightarrow 0$, RV consistently estimates the true latent volatility, but in practice, due to data limitations, it follows that

$$RV_t = IV_t + \varepsilon_t \tag{3}$$

and

$$\varepsilon_t \sim N(0, 2\Delta IQ_t), \tag{4}$$

where $IQ_t = \int_{t-1}^{t} \sigma_s^4 ds$ is the Integrated Quarticity (IQ). In turn, IQ can be consistently estimated as

$$RQ_t = \frac{M}{3} \sum_{i=1}^{M} r_{t,i}^4 \ . \tag{5}$$

However, in the presence of jumps, QV will differ from IV, with the difference given by $dJ_t = k_t dq_t$, where $k_t = p_t - p_{t-}$ is the size of the corresponding jump in the logarithmic price process[1] and q_t is a counting process, with possibly time-varying intensity λ_t, such that $P(dq_t = 1) = \lambda_t dt$. Under the assumptions in [1]

$$RV_t \underset{p}{\rightarrow} QV_t = IV_t + \sum_{t-1 \leq s \leq t} k^2(s) \ ,$$

[1] Note that Eq. (1) implies that the jump-diffusion process is right-continuous. Therefore, we use p_{t-} to denote the left-continuous version of the process.

hence, RV is a consistent estimator of QV, but not of IV. An alternative here is to use jump-robust estimators, such as the *Bipower* and *Tripower Variation* [5], *minRV* or *medRV* [2], that are consistent for IV even in the presence of jumps. In addition, in the presence of jumps, IQ will be not consistently estimated by RQ; thus, also in this case, some alternative jump-robust estimator will be needed.

A further issue is how to consistently estimate QV in the presence of market microstructure frictions. In this direction, several estimators are proposed in the literature to mitigate the influence of market microstructure noise, such as the Two Time Scales approach of [18], the Realized Kernel of [3] and the pre-averaged RV of [11], among others.

3 The Realized GARCH Model

The Realized GARCH (RGARCH), introduced by [9], provides a simultaneous structure for the joint modelling of stock returns and realized measures of volatility by completing the GARCH-X through a measurement equation that explicitly accounts for the contemporaneous relationship between the realized measure and the latent conditional variance. Formally, let $\{r_t\}$ be a time series of stock returns and $\{x_t\}$ be a time series of realized measures of volatility. The log-linear RGARCH model is specified by the following equations

$$r_t = \mu_t + \sqrt{h_t}\, z_t \,, \tag{6}$$

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma\, log(x_{t-1}) \,, \tag{7}$$

$$log(x_t) = \xi + \varphi\, log(h_t) + \tau(z_t) + u_t \,, \tag{8}$$

where $h_t = var(r_t|\mathcal{F}_{t-1})$ is the conditional variance and $\mathcal{F}_{t-1}$ the historical information set at time $t-1$. The innovations z_t and u_t are assumed to be mutually independent, with $z_t \overset{iid}{\sim} (0,1)$ and $u_t \overset{iid}{\sim} (0, \sigma_u^2)$. Also, it is assumed that the conditional mean $\mu_t = E(r_t|\mathcal{F}_{t-1}) = 0$. The function $\tau(z_t)$ allows for leverage effects, capturing dependences between returns and future volatility. Following [9], we set $\tau(z_t) = \tau_1\, z_t + \tau_2(z_t^2 - 1)$.

The RGARCH also provides an AR(1) representation for $log(h_t)$ since substituting the measurement equation into the GARCH equation we have

$$log(h_t) = (\omega + \xi\gamma) + (\beta + \varphi\gamma)\, log(h_{t-1}) + \gamma\, w_{t-1} \,, \tag{9}$$

where $w_t = \tau(z_t) + u_t$ and $E(w_t) = 0$. The coefficient $(\beta + \varphi\gamma)$ reflects the volatility persistence, while γ represents the impact of both the lagged return and realized measure on future volatility. To ensure the volatility process h_t is stationary the required restriction is $\beta + \varphi\gamma < 1$.

Our analyses focus on the log-linear specification of the RGARCH model since it offers two main advantages over the linear RGARCH. First, the positivity of the conditional variance automatically holds by construction. Second, the use of log-measures substantially reduces, but does not eliminate, the heteroskedasticity of the measurement error term.

4 Modelling Heteroskedastic Measurement Errors and Dynamic Attenuation Bias Effects

In this section, a generalisation of the basic RGARCH is proposed to model the natural heteroskedasticity of the measurement error u_t, as well as the dynamic attenuation bias. As discussed in Sect. 2, any consistent estimator of the IV can be written as the sum of the conditional variance plus a random innovation. Since the variance of this innovation term is function of the IQ, it seems natural to model the variance of the noise u_t in Eq. (8) as function of the RQ. Thus, it is assumed that the measurement noise variance is time-varying, i.e. $u_t \overset{iid}{\sim} (0, \sigma^2_{u,t})$.

In order to model the time-varying variance of the measurement noise, the specification

$$\sigma^2_{u,t} = exp\left\{\delta_0 + \delta_1 log\left(\sqrt{RQ_t}\right)\right\} \tag{10}$$

is considered, where the exponential formulation guarantees the positivity of the estimated variance, without imposing constraints on the parameters δ_0 and δ_1. The resulting model is denoted the *Heteroskedastic* Realized GARCH (HRGARCH).

In order to account for dynamic attenuation effects in the volatility persistence, the basic HRGARCH specification is further extended allowing for time-varying persistence in the volatility equation. This is achieved by letting γ, the impact coefficient of the lagged realized measure, depend on the time-varying variance of the measurement noise u_t. In line with [6], the impact of past realized measures on current volatility is expected to be down-weighted in periods in which the efficiency of the realized measure is low. The resulting model is called the *Time Varying Heteroskedastic* Realized GARCH (TV-HRGARCH). The volatility updating equation of the TV-HRGARCH is given by

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma_t\, log(x_{t-1})\,, \tag{11}$$

where

$$\gamma_t = \gamma_0 + \gamma_1\, \sigma^2_{u,t-1} \tag{12}$$

and $\sigma^2_{u,t}$ follows the specification in (10).

The setting which has been considered so far does not allow for the occurrence of jumps. Consideration is now given to a variant of the proposed modelling approach that, in order to capture this additional source of bias, features a jumps component as an additional variable in the measurement equation. This is achieved by adding the log-ratio between a non jump-robust realized measure x_t, such as the standard RV estimator or the Realized Kernel, and a jump-robust realized measure x_t^J, as an explanatory variable.[2]

Generally, let $C_t = x_t / x_t^J$. In the limit, this ratio will converge in probability to the ratio between QV and IV. Values of $C_t > 1$ are interpreted as providing evidence of jumps occurring at time t, while the discrepancy between the two measures is expected to disappear in absence of jumps, leading to values of $C_t \approx 1$. Naturally, sampling variability will play a role here and values $C_t < 1$ will be possible, in a small proportion of cases. This is compatible with the fact that the observed C_t is given by the combination of a latent signal $\bar{C}_t \geq 1$ and a measurement error, thus explaining observed values of C_t below the threshold 1.

Considering the bias correction variable C_t, the measurement equation is specified as

$$log(x_t) = \xi + \varphi\, log(h_t) + \eta\, log(C_t) + \tau(z_t) + u_t^* \tag{13}$$

or equivalently

$$log(x_t^*) = \xi + \varphi\, log(h_t) + \tau(z_t) + u_t^* \ , \tag{14}$$

where $log(x_t^*) = log(x_t / C_t^\eta) = (1 - \eta)\, log(x_t) + \eta\, log(x_t^J)$.

If $0 < \eta < 1$, as it results from our empirical analysis, $log(x_t^*)$ is a weighted average of $log(x_t)$ and $log(x_t^J)$ and the parameter η can be seen as the weight to be assigned to the jump-robust log-transformed realized measure $log(x_t^J)$, whereas $(1 - \eta)$ is the weight corresponding to the non-jump robust realized measure $log(x_t)$. If $\eta = 0$, $log(x_t^*)$ coincides with the non-robust realized measure $log(x_t)$, while for $\eta = 1$, $log(x_t^*)$ reduces to the jump robust log-transformed realized measure $log(x_t^J)$.

Considering the chosen RGARCH and the AR(1) representation for the log-conditional variance, it follows that

$$log(h_t) = (\omega + \xi\gamma) + (\beta + \varphi\gamma) log(h_{t-1}) + \gamma\, w_{t-1}^* \ , \tag{15}$$

where

$$w_t^* = \tau(z_t) + u_t^*$$

and

$$u_t^* = log(x_t^*) - \xi - \varphi log(h_t) - \tau(z_t). \tag{16}$$

[2] In the empirical application, among the different proposals arising in the literature, we focus on the *medRV* estimator.

By substituting Eq. (16) in (15), the log-conditional variance can be alternatively written as

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma\, log(x_{t-1}) - \gamma\eta\, log(C_{t-1}) \tag{17}$$

or equivalently

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma\, log(x^*_{t-1}). \tag{18}$$

In this modified framework, looking at Eq. (17), it then turns out that the log-conditional variance $log(h_t)$ is driven not only by past values of the realized measure but also, with opposite sign, by past values of the associated bias. The additional parameter η allows to adjust the contribution of C_{t-1}. From a different point of view, Eqs. (14) and (18) suggest that the volatility updating equation can be rewritten in a form similar to that of the standard RGARCH model, with the substantial difference that volatility changes are driven instead by the bias-corrected measure $log\left(x^*_{t-1}\right)$; the amount of correction is determined by the estimated scaling parameter η. This specification, of course, extends to the HRGARCH and TV-HRGARCH models.

In the remainder, models incorporating the bias correction variable C_t in the measurement equation will be denoted by using the superscript " * ", namely: RGARCH*, HRGARCH* and TV-HRGARCH*.

The model parameters can be estimated by standard QML techniques. Let Y_t indicate any additional explanatory variable eventually included in the measurement equation. Following [9], the quasi log-likelihood function, conditionally on past information $\mathcal{F}_{t-1}$ and Y_t, is given by

$$\mathcal{L}(\mathbf{r}, \mathbf{x}; \boldsymbol{\theta}) = \sum_{t=1}^{T} log\, f(r_t, x_t | \mathcal{F}_{t-1}, Y_t)\,,$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}'_h, \boldsymbol{\theta}'_x, \boldsymbol{\theta}'_\sigma)'$ with $\boldsymbol{\theta}_h, \boldsymbol{\theta}_x$ and $\boldsymbol{\theta}_\sigma$ respectively being the vectors of parameters appearing in the volatility equation ($\boldsymbol{\theta}_h$), in the level of the measurement equation ($\boldsymbol{\theta}_x$) and in the noise variance specification ($\boldsymbol{\theta}_\sigma$).

An attractive feature of the RGARCH structure is that the conditional density $f(r_t, x_t | \mathcal{F}_{t-1}, Y_t)$ can be easily decomposed as

$$f(r_t, x_t | \mathcal{F}_{t-1}, Y_t) = f(r_t | \mathcal{F}_{t-1})\, f(x_t | r_t; \mathcal{F}_{t-1}, Y_t).$$

Assuming a Gaussian specification for z_t and u_t, such as $z_t \overset{iid}{\sim} N(0,1)$ and $u_t \overset{iid}{\sim} N(0, \sigma_u^2)$, the quasi log-likelihood function is

$$\mathcal{L}(\mathbf{r}, \mathbf{x}; \boldsymbol{\theta}) = \underbrace{-\frac{1}{2}\sum_{t=1}^{T} log(2\pi) + log(h_t) + \frac{r_t^2}{h_t}}_{\ell(\mathbf{r})} + \underbrace{-\frac{1}{2}\sum_{t=1}^{T} log(2\pi) + log(\sigma_u^2) + \frac{u_t^2}{\sigma_u^2}}_{\ell(\mathbf{x}|\mathbf{r})}\,. \tag{19}$$

Since standard GARCH models do not include an equation for x_t, the overall maximised log-likelihood values given by RGARCH models are not comparable to those returned from the estimation of standard GARCH-type models; the former will tend to be larger. Nevertheless, the partial log-likelihood value of the returns component, $\ell(\mathbf{r}) = \sum_{t=1}^{T} log\ f(r_t|\mathcal{F}_{t-1})$, can still be meaningfully compared to the maximised log-likelihood value achieved for a standard GARCH-type model.

5 The Data

To assess the performance of the proposed models, an empirical application to four stocks traded on the Xetra Market in the German Stock Exchange has been performed. This section presents the salient features of the data analysed. In particular, the following assets are considered: Allianz (ALV), a financial services company dealing mainly with insurance and asset management; Bayerische Motoren Werke (BMW), a company engaged in vehicle and engine manufacturing; Metro Group (MEO), a cash and carry group and RWE (RWE), a company providing electric utilities. There are two main reasons for choosing these stocks. First, they are consistently ranked among the 30 largest capitalisation stocks on the German market (DAX 30). Second, belonging to different industrial sectors, they allow us to assess the robustness of our empirical results in variable market contexts.

The original dataset included tick-by-tick data on transactions (trades only) in the period 02/01/2002 to 27/12/2012. The raw data have been cleaned, using the procedure described in [7], then converted to an equally spaced series of five-minute log-returns, which are aggregated on a daily basis to compute a time series of 2791 daily log-returns; the realized measures RV and $medRV$ and two realized quarticity measures, RQ and $medRQ$. Only continuous trading transactions during the regular market hours 9:00 am–5:30 pm are considered.[3]

Figure 1 displays the daily returns for the four analysed stocks. These reveal three periods of high volatility common to all assets: the first relates to the dot com bubble in 2002; the second is the financial crisis starting in mid 2007 and peaking in 2008; the crisis in Europe then progressed from the banking system to a sovereign debt crisis, with the highest turmoil level in the late 2011, the 3rd period. These are clearly evident in Fig. 2, reporting the time plots of the daily 5-minute RV series. Finally, Fig. 3 shows the evolution of the bias correction variables C_t over time. This fluctuates approximately around a base level $\approx$1, with an evident positive skewness due to the the upward peaks (jumps), while downward variations due to measurement noise appear to be much less pronounced and negligible.

[3] See [13] for more details on the cleaning procedure.

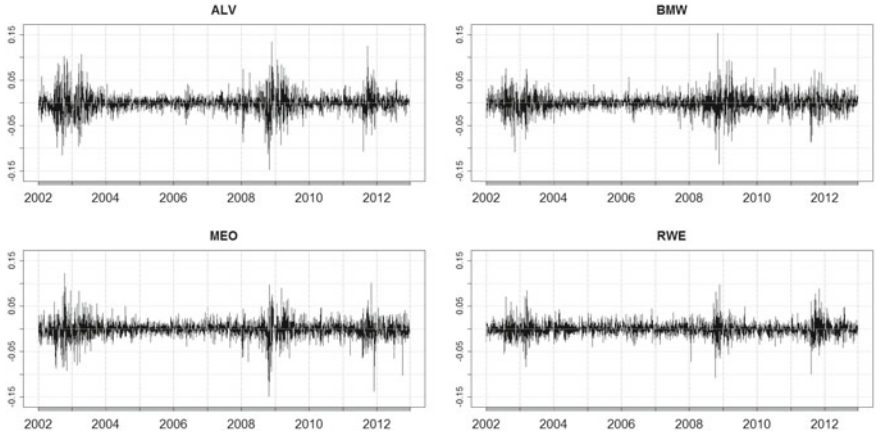


Fig. 1 Time series of daily log-returns. Daily log-returns for the stocks ALV (top-left), BMW (top-right), MEO (bottom-left) and RWE (bottom-right) for the full sample period 02/01/2002–27/12/2012

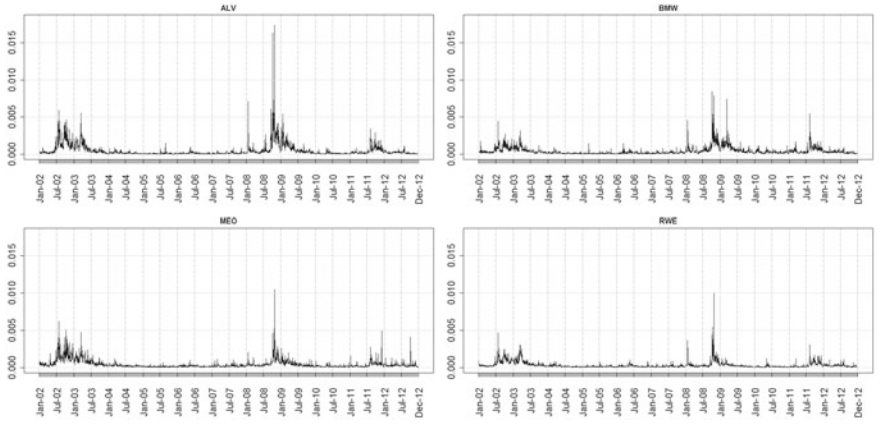


Fig. 2 Daily Realized Volatility. Daily Realized Volatility computed using a sampling frequency of 5 min. ALV (top-left), BMW (top-right), MEO (bottom-left) and RWE (bottom-right). Full sample period 02/01/2002–27/12/2012

6 In-sample Analysis

This section discusses the in-sample performance of the proposed models. It is worth noting that in presence of jumps, RQ_t could also be affected by bias. So, as a robustness check, in addition to the jump-free models (RGARCH, HRGARCH and TV-HRGARCH) and the models in which the impact of jumps is considered in the measurement equation (RGARCH*, HRGARCH* and TV-HRGARCH*), we also consider a variant of the latter class of models where, in the specification of $\sigma^2_{u,t}$,

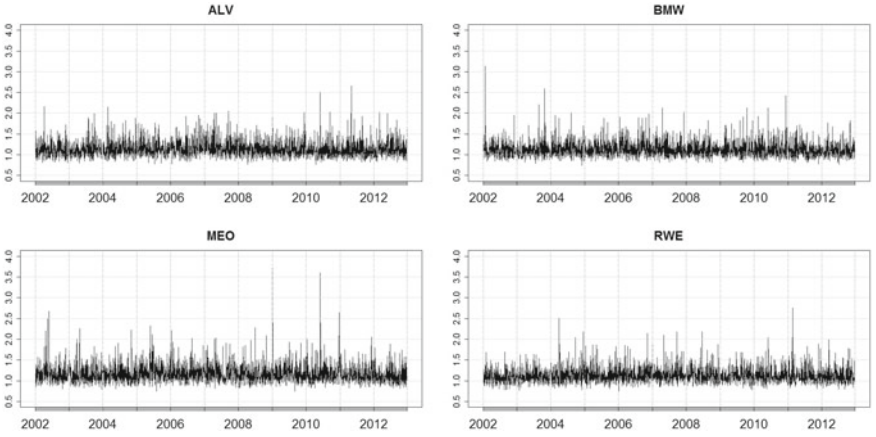


Fig. 3 Time series of daily bias correction variable C_t. Daily bias correction variable $C_t = RV_t/medRV_t$ for the stocks ALV (top-left), BMW (top-right), MEO (bottom-left) and RWE (bottom-right) for the full sample period 02/01/2002–27/12/2012

RQ_t is replaced by the jump-robust estimator $medRQ$. In order to distinguish this class of models the subscript "MRQ" has been used.

Estimation results using the 5-min RV as realized measure are reported in Table 1, showing parameter estimates, together with values of the log-likelihood $\mathcal{L}(\mathbf{r}, \mathbf{x})$, partial log-likelihood $\ell(\mathbf{r})$ and Bayesian Information Criterion (BIC), for the four analysed stocks.[4]

The parameter β is between 0.565 and 0.752, always taking the highest value for HRGARCH and the lowest for TV-HRGARCH, whereas φ takes values close to one with a limited variability across different models. This suggests that $log(x_t)$ is roughly proportional to the log-conditional variance. These results are in line with the findings in [9]. The parameters of the leverage function $\tau(z)$ are always significant (except for τ_1 using the RGARCH* for MEO stock), with τ_1 negative and τ_2 positive, as expected.

The parameter δ_1 is always positive and statistically significant at the 0.05 level, thus giving empirical confirmation to the intuition that the variance of the measurement error $\sigma^2_{u,t}$ is time-varying and, in accordance with the asymptotic theory, suggesting that this is positively related to the IQ. This also implies that $\sigma^2_{u,t}$ tends to take on higher values in periods of turmoil and lower values when volatility tends to stay low, as it can be easily seen in Fig. 4 which compares the constant variance σ^2_u estimated by RGARCH with the time-varying variance $\sigma^2_{u,t}$ given by the HRGARCH model. For the four analysed stocks, the trend of $\sigma^2_{u,t}$ follows the dynamics of the realized measure, being higher in turbulent periods and lower in calm periods, while the

[4] To save space, we do not report robust standard errors but results are available on request.

Table 1 In-sample estimation results using 5 min realized volatility

		ω	γ	γ_0	γ_1	β	ξ	φ	τ_1	τ_2	η	σ_u^2	δ_0	δ_1	$\ell(\mathbf{r})$	$\mathcal{L}(\mathbf{r}, \mathbf{x})$	BIC
ALV	RG	**0.008**	0.402	–	–	0.598	**−0.399**	0.953	−0.069	0.108	–	0.185	–	–	7609.434	6006.082	−4.281
	HRG	**0.009**	0.381	–	–	0.620	−0.473	0.948	−0.069	0.111	–	–	**−0.405**	0.162	7609.002	6032.686	−4.297
	TV-HRG	1.995	–	0.489	1.060	0.565	−0.462	0.949	−0.069	0.111	–	–	**−0.394**	0.166	7611.877	6063.920	−4.317
	RG*	**0.002**	0.408	–	–	0.590	**−0.382**	0.960	−0.068	0.107	0.398	0.180	–	–	7609.635	6042.332	−4.304
	HRG*	**0.010**	0.392	–	–	0.607	−0.471	0.953	−0.068	0.110	0.388	–	**−0.493**	0.154	7609.412	6066.969	−4.319
	TV-HRG*	1.265	–	0.446	0.747	0.575	−0.466	0.952	−0.069	0.110	0.285	–	**−0.530**	0.151	7610.934	6082.354	−4.327
	HRG^*_{MRQ}	**0.010**	0.394	–	–	0.605	−0.467	0.954	−0.069	0.110	0.403	–	**−0.539**	0.145	7609.429	6065.011	−4.318
	TV-HRG^*_{MRQ}	1.534	–	0.449	0.885	0.579	−0.469	0.954	−0.069	0.110	0.466	–	**−0.499**	0.151	7613.060	6083.501	−4.328
BMW	RG	**0.044**	0.304	–	–	0.705	**−0.520**	0.927	−0.029	0.082	–	0.171	–	–	7470.471	5976.587	−4.260
	HRG	**0.036**	0.280	–	–	0.727	**−0.591**	0.924	−0.032	0.086	–	–	**0.321**	0.270	7470.853	6026.783	−4.293
	TV-HRG	1.614	–	0.410	0.617	0.692	−0.577	0.926	−0.032	0.085	–	–	**0.147**	0.250	7469.434	6053.708	−4.310
	RG*	**0.047**	0.309	–	–	0.699	**−0.526**	0.930	−0.029	0.081	0.277	0.169	–	–	7471.600	5993.826	−4.270
	HRG*	**0.039**	0.285	–	–	0.722	−0.588	0.927	−0.031	0.085	0.223	–	**0.201**	0.256	7471.749	6037.920	−4.298
	TV-HRG*	1.358	–	0.389	0.524	0.696	−0.588	0.926	−0.032	0.085	0.138	–	**0.115**	0.247	7470.247	6058.335	−4.310
	HRG^*_{MRQ}	**0.041**	0.287	–	–	0.720	**−0.574**	0.928	−0.031	0.084	0.243	–	**−0.097**	0.212	7471.793	6024.967	−4.289
	TV-HRG^*_{MRQ}	1.067	–	0.349	0.469	0.707	−0.522	0.935	−0.031	0.083	0.309	–	**−0.117**	0.211	7469.812	6035.934	−4.294

(continued)

Table 1 (continued)

		ω	γ	γ_0	γ_1	β	ξ	φ	τ_1	τ_2	η	σ_u^2	δ_0	δ_1	$\ell(\mathbf{r})$	$\mathcal{L}(\mathbf{r}, \mathbf{x})$	BIC
MEO	RG	**−0.080**	0.252	–	–	0.740	**−0.082**	0.985	−0.018	0.082	–	0.189	–	–	7463.731	5828.595	−4.154
	HRG	**−0.077**	0.240	–	–	0.752	**−0.157**	0.982	−0.023	0.090	–	–	**−0.045**	0.209	7463.768	5860.955	−4.174
	TV-HRG	1.433	–	0.339	0.699	0.712	**−0.178**	0.978	−0.024	0.090	–	–	**−0.225**	0.188	7469.276	5890.380	−4.193
	RG*	**−0.072**	0.257	–	–	0.734	**−0.117**	0.987	**−0.016**	0.081	0.362	0.185	–	–	7464.330	5860.417	−4.174
	HRG*	**−0.069**	0.246	–	–	0.745	**−0.181**	0.984	−0.022	0.089	0.341	–	**−0.151**	0.198	7464.372	5888.727	−4.191
	TV-HRG*	**0.945**	–	0.307	0.487	0.722	**−0.191**	0.980	−0.023	0.089	0.253	–	**−0.278**	0.183	7467.683	5905.626	−4.201
	HRG^*_{MRQ}	**−0.066**	0.248	–	–	0.744	**−0.193**	0.982	−0.020	0.088	0.362	–	**−0.253**	0.180	7464.360	5883.987	−4.188
	TV-HRG^*_{MRQ}	1.222	–	0.314	0.633	0.721	**−0.211**	0.980	−0.020	0.088	0.430	–	**−0.308**	0.174	7470.239	5903.041	−4.199
RWE	RG	−0.406	0.315	–	–	0.644	0.736	1.067	−0.041	0.082	–	0.162	–	–	7991.189	6569.620	−4.685
	HRG	−0.366	0.292	–	–	0.671	**0.606**	1.058	−0.040	0.083	–	–	0.293	0.262	7989.793	6614.132	−4.714
	TV-HRG	0.933	–	0.389	0.613	0.634	**0.660**	1.063	−0.041	0.083	–	–	**−0.128**	0.212	7992.232	6631.893	−4.724
	RG*	−0.410	0.317	–	–	0.640	**0.743**	1.072	−0.041	0.080	0.284	0.160	–	–	7991.585	6588.595	−4.696
	HRG*	−0.377	0.294	–	–	0.666	**0.639**	1.064	−0.040	0.082	0.245	–	**0.179**	0.250	7990.196	6628.284	−4.721
	TV-HRG*	**0.458**	–	0.357	0.374	0.646	**0.675**	1.067	−0.040	0.082	0.180	–	**−0.046**	0.223	7991.355	6638.507	−4.726
	HRG^*_{MRQ}	−0.379	0.295	–	–	0.665	**0.649**	1.065	−0.041	0.081	0.271	–	**0.115**	0.236	7990.271	6622.699	−4.717
	TV-HRG^*_{MRQ}	**0.284**	–	0.338	**0.303**	0.654	**0.677**	1.069	−0.041	0.081	0.308	–	**−0.119**	0.209	7990.301	6627.033	−4.718

In-sample parameter estimates for the full sample period 02 January 2002–27 December 2012. $\ell(\mathbf{r})$: partial log-likelihood. $\mathcal{L}(\mathbf{r}, \mathbf{x})$: log-likelihood. BIC: Bayesian Information Criterion. Parameters that are not significant at the 5% level are reported in bold. RG: RGARCH

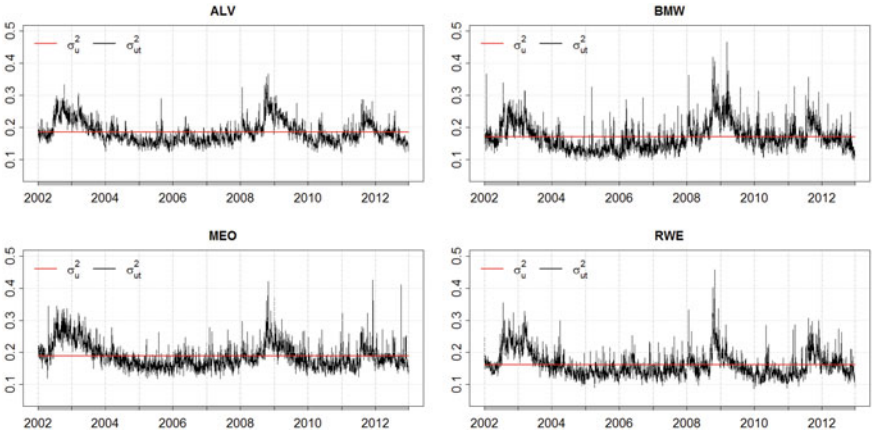


Fig. 4 Constant verses time-varying variance $\sigma^2_{u,t}$. Figure shows the constant variance σ^2_u (red-line) estimated with RGARCH together with the time-varying variance $\sigma^2_{u,t}$ (black-line) estimated with HRGARCH. Both models have been fitted taking the 5 min RV as volatility proxy. Full sample period 02 January 2002–27 December 2012

constant variance σ^2_u estimated within the RGARCH (red line in the plots) is approximately equal to the average level of the time-varying variance of the measurement noise.

Another key parameter is the coefficient γ, which summarises the impact of the realized measure on future volatility. It ranges from 0.240 to 0.402 for both RGARCH and HRGARCH and its value is generally increased by the introduction of the bias correction variable C_t in the measurement equation; this provides additional evidence supporting the idea that accounting for jumps further reduces the attenuation bias effect on γ. For the TV-HRGARCH this effect is adaptively explained by the time-varying coefficient γ_t, depending on the past noise variance $\sigma^2_{u,t-1}$ through the slope coefficient γ_1. This is always positive and statistically significant at the 0.05 level, except for the TV-HRGARCH$^*_{MRQ}$ model fitted to RWE, with an associated p-value of 0.107. Interestingly, in the TV-HRGARCH* and TV-HRGARCH$^*_{MRQ}$ models, the introduction of the bias correction variable in the measurement equation has the effect of reducing the value of the fitted γ_1, compared to what we find for the TV-HRGARCH.

The value of γ_1 determines the amount by which the response to past volatility is corrected to account for attenuation bias effects. Since γ_1 is always positive and $log(x_t)$ is negative, when the lagged variance of the error term of the realized measure $\sigma^2_{u,t-1}$ is high, the impact of the lagged log-transformed realized measure $log(x_{t-1})$ on $log(h_t)$ will be negative and lower than what would have been implied by the same value of $log(x_{t-1})$ in correspondence of a lower value of $\sigma^2_{u,t-1}$. Said differently, the impact of x_{t-1} on h_t will be down-scaled towards zero when $\sigma^2_{u,t-1}$ increases. Equivalently, variations in h_t ($\nabla h_t = h_t - h_{t-1}$) will be negatively correlated with the values of γ_t and $\sigma^2_{u,t-1}$. These results are in line with the recent findings of [6].

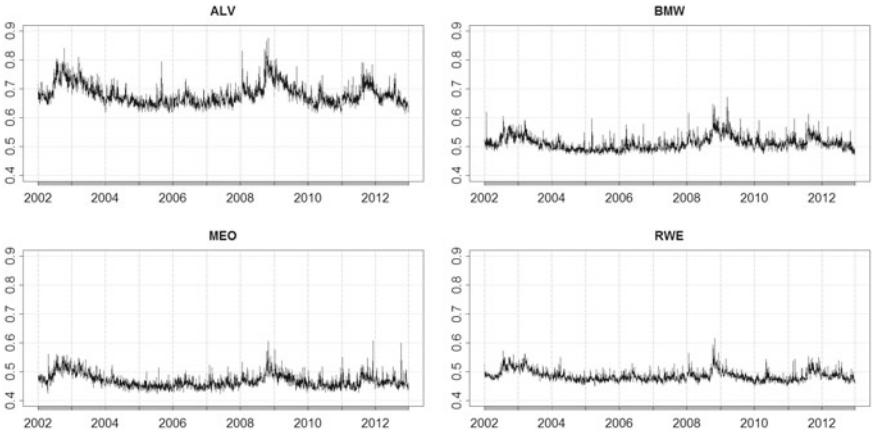


Fig. 5 Time-varying coefficient γ_t given by the TV-HRGARCH model. Figure shows the time-varying coefficient $\gamma_t = \gamma_0 + \gamma_1 \sigma^2_{u,t-1}$. Full sample period 02 January 2002–27 December 2012

Figure 5 displays the time plot of the γ_t coefficient for the four considered stocks. It is evident that when the variance of the measurement error is high, γ_t is also high, leading to a less substantial increase of h_t compared to days in which, ceteris paribus, $\sigma^2_{u,t}$ is low and the realized measure provides a stronger more reliable signal. Further, the value of γ_t tends to be higher than the value of the time invariant γ estimated within the RGARCH and HRGARCH models.

Finally, it is interesting to note that, for each series, the variance $\sigma^2_{u^*}$ of the measurement equation error u^*_t of the RGARCH* model is slightly lower than what observed for the RGARCH model, providing evidence of an improved goodness of fit in the modified measurement equation and efficiency of the bias corrected realized measure x^*_t. In this context, an important role is played by the smoothing parameter η, determining the amount of jump-implied bias correction associated to the ratio C_t. Overall, the variability of the estimated η coefficients across assets makes evident the flexibility of our modelling approach: the amount of smoothing in $log(x^*_t)$ is not arbitrarily chosen, but data driven through the estimated parameter η.

Reminding the interpretation of η discussed in Sect. 4, since the estimated η results lower than 0.5 for all the analysed assets, the impact of the realized measure $log(x_t)$ is, on average, greater than the impact of the jump-robust realized measure $log(x^J_t)$. Furthermore, as a general trend we observe that, for all stocks, the fitted value of η tends to be higher for heteroskedastic models relying on the jump robust quarticity estimator $medRQ$, (TV-) HRGARCH$^*_{MRQ}$, rather than for models based on RQ, (TV-)HRGARCH*.

To evaluate the performance of the estimated models we refer to the commonly used Bayesian Information Criterion (BIC). From the last column of Table 1 it clearly emerges that the specifications accounting for both heteroskedasticity and attenuation bias effects, within the class of models corrected by the C_t variable, tend to

minimise the BIC. In particular, the TV-HRGARCH* features the lowest value of the considered information criterion in three cases out of four, whereas for the stock ALV the TV-HRGARCH$^*_{MRQ}$ prevails, closely followed by the TV-HRGARCH* model. Interestingly, in the latter case, the parameter η given by the use of $medRQ_t$ instead of RQ_t as state variable of $\sigma^2_{u,t}$ into TV-HRGARCH* specification, takes on the highest estimated values within the class of models built using information on C_t, meaning that the jump-robust realized measure $medRV$ has an higher impact in determining volatility dynamics for ALV compared to the other stocks. The BIC analysis also suggests that the standard RGARCH provides the highest BIC values, but the simple inclusion of the C_t variable (RGARCH*) remarkably improves the fit highlighting the importance of modelling the bias due to the occurrence of jumps.

Summarising: the in-sample results show that the introduction of heteroskedasticity and time-varying persistence, as well as the bias correction for jumps and measurement errors, has positive effects on the accuracy of the estimated volatility. Consequently, in-sample, the proposed specifications show notable improvements over the standard RGARCH in terms of goodness of fit. Overall, the models incorporating heteroskedasticity, time-varying volatility response coefficient and jumps correction are those returning the best performances.

7 Out-of-Sample Analysis

In this section, the out-of-sample predictive ability of the models estimated in the previous section is assessed via a rolling window forecasting exercise, using an estimation window of 1500 days. The out-of-sample period starts on 26 November 2007 and includes 1270 daily observations, covering the credit crisis and the turbulent period from November 2011 to the beginning of 2012.

In order to assess the forecasting performance of the proposed models, the predictive (quasi) log-likelihood and the QLIKE loss function [14] are employed. Furthermore, the Model Confidence Set (MCS) of [10] is used to evaluate the comparative predictive ability of all the models, considering a confidence level of 75%. In particular, the Semi-Quadratic statistic,[5] based on a block-bootstrap procedure with 5000 re-samples, is employed to sequentially test the hypothesis of equal predictive ability, where the optimal block length has been chosen through the method described in [15]. It is worth noting that the QLIKE loss function specifically measures the ability in forecasting the conditional variance of returns, while the predictive log-likelihood assesses the ability of a given model to predict the conditional distribution of $(r_t, x_t | \mathcal{F}_{t-1}, Y_t)$, thus considering both components of the RGARCH quasi likelihood.

[5] We do not report p-values for the Range statistic as we obtain practically the same results as for the Semi-Quadratic statistic.

Table 2 Predictive log-likelihood using 5-min RV as volatility proxy (left): in **bold** the preferred model according to predictive log-likelihood. MCS p-values of predictive log-likelihood (right): in box models $\in$ 75% MCS. The p-values refer to the negative predictive log-likelihoods

	Average				MCS p-values			
	ALV	BMW	MEO	RWE	ALV	BMW	MEO	RWE
RGARCH	2617.961	2478.886	2498.497	2938.425	0.0006	0.0022	0.0000	0.0026
HRGARCH	2635.038	2500.582	2542.889	2974.401	0.0060	0.0322	0.0012	0.0578
TV-HRGARCH	2654.591	2515.229	2560.055	2982.102	0.3128	0.1182	0.0886	0.2200
RGARCH*	2634.241	2488.724	2523.193	2958.191	0.0830	0.0322	0.0008	0.0468
HRGARCH*	2651.794	2507.150	2560.524	2987.250	0.1086	0.1182	0.1858	0.5568
TV-HRGARCH*	**2660.343**	**2517.435**	2569.528	**2990.083**	1.0000	1.0000	0.8324	1.0000
HRGARCH$^*_{MRQ}$	2651.363	2502.983	2555.388	2983.233	0.0856	0.0730	0.0386	0.2692
TV-HRGARCH$^*_{MRQ}$	2659.316	2507.181	**2571.215**	2982.434	0.8762	0.1182	1.0000	0.2200

Since the RGARCH can be obtained as a special case of each of the models estimated in Sect. 6, the first criterion we use for assessing predictive accuracy is, as in [9], the predictive log-likelihood, given for time $t+1$ by:

$$\hat{\mathcal{L}}(\mathbf{r},\mathbf{x})_{t+1} = -\frac{1}{2}\left[log(2\pi) + log(\hat{h}_{t+1}) + \frac{r^2_{t+1}}{\hat{h}_{t+1}}\right] - \frac{1}{2}\left[log(2\pi) + log(\hat{\sigma}^2_u) + \frac{u^2_{t+1}}{\hat{\sigma}^2_u}\right]. \tag{20}$$

Subsequently, the aggregated predictive log-likelihood is computed by summing the density estimates for each day in the forecast period. Table 2 shows the values of the predictive log-likelihood corresponding to all models employed. The models including the bias correction variable C_t in the measurement equation maximise the predictive log-likelihood in all cases. In particular, for ALV, BMW and RWE, the specification that allows for heteroskedasticity and time-varying persistence, TV-HRGARCH*, returns the highest values of the predictive log-likelihood, whereas for MEO, the TV-HRGARCH$^*_{MRQ}$, replacing RQ_t by $medRQ_t$ in the specification of $\sigma^2_{u,t}$, prevails.

The right panel of Table 2 shows the MCS p-values associated to the predictive log-likelihood (multiplied by -1). Interestingly, the only model always coming into the MCS at the 75% confidence level is the TV-HRGARCH*, while the standard RGARCH never enters the set of superior models. For the asset ALV, all the time-varying specifications enter the MCS at the considered confidence level, while for BMW, the TV-HRGARCH* uniquely enters the set. On the other hand, for MEO, both the TV-HRGARCH* and TV-HRGARCH$^*_{MRQ}$ are included into the MCS. Finally, the HRGARCH* and the HRGARCH$^*_{MRQ}$ belong to the set of superior models only in the analysis of the RWE stock.

Table 3 Average values of QLIKE loss using 5-min RV as volatility proxy (left) and MCS p-values (right). For each stock in **bold** minimum loss and in box models $\in$ 75% MCS

	QLIKE				MCS p-values			
	ALV	BMW	MEO	RWE	ALV	BMW	MEO	RWE
RGARCH	-6.9723	-6.6765	-6.9460	-7.2618	0.0766	0.1616	0.0286	0.0294
HRGARCH	-6.9721	-6.6755	-6.9464	-7.2606	0.0772	0.0778	0.0714	0.0230
TV-HRGARCH	**-6.9755**	**-6.6788**	**-6.9529**	**-7.2652**	1.0000	1.0000	1.0000	1.0000
RGARCH*	-6.9728	-6.6775	-6.9480	-7.2625	0.0772	0.4294	0.1336	0.0310
HRGARCH*	-6.9728	-6.6763	-6.9482	-7.2612	0.0772	0.1534	0.2692	0.0280
TV-HRGARCH*	-6.9746	-6.6788	-6.9525	-7.2647	0.4234	0.7902	0.8038	0.0758
HRGARCH$^*_{MRQ}$	-6.9728	-6.6766	-6.9482	-7.2614	0.0772	0.2548	0.2174	0.0280
TV-HRGARCH$^*_{MRQ}$	-6.9754	-6.6776	-6.9525	-7.2620	0.9418	0.4294	0.8196	0.0294

As a further criterion for assessing and comparing the forecasting accuracy of the fitted models, the QLIKE loss function is employed. This choice is motivated by two considerations. First, the QLIKE is robust to noisy volatility proxies [14]. Second, compared to other robust alternatives, this loss function has been found to be more powerful in rejecting poorly performing predictors [12]. The QLIKE loss has been computed according to the formula

$$QLIKE = \frac{1}{T}\sum_{t=1}^{T} log(\hat{h}_t) + \frac{x_t}{\hat{h}_t}, \tag{21}$$

where the 5 min RV is chosen as volatility proxy.

Observing the average values of the QLIKE loss function (left panel of Table 3), we find that the lowest value is always obtained by using the TV-HRGARCH model. The MCS p-values, in the right panel of Table 3, show that the TV-HRGARCH is always included in the MCS and, for RWE, it is the only specification appearing in the MCS. Within the class of models which include a jump correction component in the measurement equation, the RGARCH* is included into the MCS for BMW, whereas the HRGARCH* falls into the MCS only for MEO. The TV-HRGARCH* is always included in the set of superior models, except for RWE. This result also applies to the TV-HRGARCH$^*_{MRQ}$, while the simplified heteroskedastic specification HRGARCH$^*_{MRQ}$ enters the MCS just for BMW. On the other hand, the standard RGARCH never comes into the set of superior models, nor does the HRGARCH.

8 Conclusion

We have proposed a generalisation of the class of Realized GARCH models that accounts for heteroskedasticity in measurement error and which explicitly reduces the magnitude of the attenuation bias through the temporal variation of the parameters driving the volatility persistence. Furthermore, in order to deal with the presence of jumps in stock prices, we further extend the proposed modelling approach introducing a bias correction variable that allows to control the impact of jumps on the predicted volatility in a fully data driven fashion.

Our empirical findings point out that the proposed modelling approach outperforms the standard Realized GARCH both in fitting and forecasting volatility. In particular, we find evidence in favour of the use of models accounting for heteroskedasticity, time-varying attenuation bias as well as the presence of jumps. More in detail, focusing on the out-of-sample predictive performance, we find that, when the predictive log-likelihood is used as a measure of accuracy, the TV-HRGARCH* model, incorporating all the mentioned features, is the only model always entering the 75% MCS, while the TV-HRGARCH model, not incorporating the jump correction term, only enters the MCS for the stock ALV. Moving to consider the QLIKE as a measure of predictive accuracy, the performance of the TV-HRGARCH* model is still remarkably good since this model enters the MCS in three cases out of four. However, it is slightly outperformed by the TV-HRGARCH model entering the MCS for all the stocks considered.

As a robustness check, the empirical analysis, both in-sample and out-of-sample, has also been repeated using RK and $medRQ$ as alternative volatility and quarticity measures, respectively, confirming the results obtained with the RV (results are available upon request). This exercise confirms the robustness of our findings to the choice of the volatility and quarticity measures used for modelling. Projects for future research include extending the analysis to consider implications for tail risk forecasting.

References

1. Andersen, T.G., Bollerslev, T., Diebold, F.X.: Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. Rev. Econ. Stat. **89**(4), 701–720 (2007)
2. Andersen, T.G., Dobrev, D., Schaumburg, E.: Jump-robust volatility estimation using nearest neighbor truncation. J. Econom. **169**(1), 75–93 (2012)
3. Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N.: Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. Econometrica **76**(6), 1481–1536 (2008)
4. Barndorff-Nielsen, O.E., Shephard, N.: Econometric analysis of realized volatility and its use in estimating stochastic volatility models. J. Roy. Stat. Soc.: Ser. B (Stat. Methodol.) **64**(2), 253–280 (2002)
5. Barndorff-Nielsen, O.E., Shephard, N.: Power and bipower variation with stochastic volatility and jumps. J. Financ. Economet. **2**(1), 1–37 (2004)

6. Bollerslev, T., Patton, A.J., Quaedvlieg, R.: Exploiting the errors: a simple approach for improved volatility forecasting. J. Econom. **192**(1), 1–18 (2016)
7. Brownlees, C.T., Gallo, G.M.: Financial econometric analysis at ultra-high frequency: Data handling concerns. Comput. Stat. Data Anal. **51**(4), 2232–2245 (2006)
8. Hansen, P.R., Huang, Z.: Exponential garch modeling with realized measures of volatility. J. Bus. Econ. Stat. **34**(2), 269–287 (2016)
9. Hansen, P.R., Huang, Z., Shek, H.H.: Realized garch: a joint model for returns and realized measures of volatility. J. Appl. Econom. **27**(6), 877–906 (2012)
10. Hansen, P.R., Lunde, A., Nason, J.M.: The model confidence set. Econometrica **79**(2), 453–497 (2011)
11. Jacod, J., Li, Y., Mykland, P.A., Podolskij, M., Vetter, M.: Microstructure noise in the continuous case: the pre-averaging approach. Stoch. Process. Their Appl. **119**(7), 2249–2276 (2009)
12. Liu, L.Y., Patton, A.J., Sheppard, K.: Does anything beat 5-minute rv? a comparison of realized measures across multiple asset classes. J. Econom. **187**(1), 293–311 (2015)
13. Naimoli, A.: Essays on the modelling and prediction of financial volatility and trading volumes (2017)
14. Patton, A.: Volatility forecast comparison using imperfect volatility proxies. J. Econom. **160**(1), 246–256 (2011)
15. Patton, A., Politis, D.N., White, H.: Correction to "automatic block-length selection for the dependent bootstrap" by d. politis and h. white. Econom. Rev. *28*(4), 372–375 (2009)
16. Shephard, N., Sheppard, K.: Realising the future: forecasting with high-frequency-based volatility (heavy) models. J. Appl. Econom. **25**(2), 197–231 (2010)
17. Shephard, N., Xiu, D.: Econometric analysis of multivariate realised qml: efficient positive semi-definite estimators of the covariation of equity prices (2016)
18. Zhang, L., Mykland, P.A., Aït-Sahalia, Y.: A tale of two time scales: Determining integrated volatility with noisy high-frequency data. J. Am. Stat. Assoc. **100**(472), 1394–1411 (2005)

Zero Inflated Bivariate Poisson Regression Models for a Sport (*in*)activity Data Analysis

Maria Iannario, Ioannis Ntzoufras, and Claudia Tarantola

Abstract In this paper, we analyze a sport (*in*)activity case study using a zero inflated bivariate Poisson model. We use the "(*in*)activity" term in order to embrace both active and passive sport participation (practicing or watching a sport, respectively). The paper investigates the determinants of sport (*in*)activity: the frequency and the probability of sports participation. It distinguishes between genuine "non-participants" and the ones who do not participate at a time of the survey but might under different circumstances.

Keywords Bivariate Poisson distribution · Paired count data · Sport · Zero inflated distributions

1 Introduction

Paired count data arise in several contexts including medicine (number of hospital admissions in several regions), epidemiology (number of positive to a specific illness after a full recovery by different treatments or geographical area), accident analysis (number of accidents in a site before and after traffic laws changes), econometrics (number of voluntary and involuntary job changes), risk analysis (the number of insurance claims with and without bodily injuries), marketing (number of selected products before and after an online marketing campaign), sports (the number of goals scored by each one of the two opponent teams in soccer), and so on.

M. Iannario
University of Naples Federico II, Via L. Rodinò, 22, Naples, Italy
e-mail: maria.iannario@unina.it

I. Ntzoufras
Athens University of Economics and Business, Patission 76, Athens, Greece
e-mail: ntzoufra@aueb.gr

C. Tarantola (✉)
University of Pavia, Via San Felice, 5, Pavia, Italy
e-mail: claudia.tarantola@unipv.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_11

Often the two count variables are correlated, hence Bivariate count models are used to jointly estimate their distribution. The most used model—introduced by [25] and further presented by [30]—is the Bivariate Poisson model. However, the definition of the Bivariate Poisson distribution is not univocal and alternative formulations have been presented (see [34], for details).

One disadvantage of Bivariate Poisson models is the lack of generality due to the use of Poisson for marginal distributions. The use of Poisson for marginal distributions does not accommodate for over/under dispersion or negative correlation. Thus, mixed Poisson models and other alternative models that allow for negative correlation have been introduced (see [2, 6, 8, 11, 19, 22, 23, 31, 40, 47], among others).

Another critical issue may arise in situations where the diagonal elements of the probability table present an inflation. For example, in pre and post treatment studies, the treatment may not have an effect on some specific patients for peculiar reasons due to the course of events [10]. Another example arises in sports such as soccer where the number of zeros in a game is larger than those predicted by a simple Bivariate Poisson model [32]. In such cases the diagonal observations present higher probabilities than the ones fitted under a Bivariate Poisson model. It may also be related to a redundancy of frequency in some specific categories as generally happens in the marginal *c*-inflated models due to some response style or *shelter* effect [28].

In this paper we focus on Zero Inflated Bivariate Poisson model (ZIBP model) which represents a special case of the Diagonal Inflated Bivariate Poisson model (DIBP model) by [33]. The DIBP model describes both correlation between two variables and over-dispersion (or alternatively under-dispersion) of the corresponding marginal distributions by taking into account the excess of zero. It has been applied by [48] for occupational injuries analyses and by [7] in the automobile insurance context.

In this contribution we employ a ZIBP model to explain the relation among *active* and *passive* sport activities. Active sport is related to physical activities, while watching a sports game, match, or event may be considered a passive domain; see e.g. [17, 38] for a detailed examination of consumer behaviour in sport. Here we apply the prefix "*in*" before the word "activity" in order to embrace both aspects of sport; thus we use the expression sport (*in*)activity (see also [29]).

The aim of the paper is to investigate the determinants of both active and passive activities; the frequency and the probability of active participation. In both cases there is an anchor that represents the absence of the activity, such as none, never. This level usually tagged zero may be scored by respondents who do not practice any kind of activity and respondents with a non-zero probability of contributing (i.e. respondents who do not practice at the time of the interview but are usually doing it). It distinguishes between genuine "non-participants" and the ones who do not participate at a time but might do under different circumstances.

Some survey data, particularly those that refer to an explicit time dimension, may include genuine non-participants whatever the circumstances are, as well as individuals who would decide to participate if the circumstances were different. It is, therefore, likely that these two types of zeros are driven by different behaviour.

With respect to the standard models for counting data the examined methodology exceeds some gaps related to the modeling of zeros by taking into account the potentially two-fold decision made with respect to participation.

The remainder of the paper is as follows. Section 2 introduces the motivating example. Section 3 reviews the methods used for the analysis starting with marginal models introduced for the selection of covariates and then presents the ZIBP model. Section 4 describes the main estimation results with a summary of the main findings and opportunities for further research. Some conclusions end the paper.

2 Motivating Example

The determinants of sport (*in*)activity will be discussed on the basis of a case study involving data collected in 2017 ($n = 549$ observations) through a survey developed for the BDsports project (http://bodai.unibs.it/bdsports/). The case study has been selected to allow the distinction between genuine inactive respondents and the ones who do not play sport at a time of the survey both in relation to sport activity and fun consumption behaviour (passive activity, that is subjects that spend time watching sport in TV).

In this study, the dependent variables of interest are count responses of the time dedicated to sport practice (in hours per week; denoted by X)—see Fig. 1, upper panel—and the time dedicated to watch sport events in TV (in number of sport events per week; denoted by Y)—see Fig. 1, center panel.

The dependence of the two variables is clearly evident in the mosaic plot which confirms an excess of zero counts (Fig. 1, bottom panel). In fact we can clearly see that the proportion of $(0, 0)$ is quite larger than the other frequencies. Furthermore, using the maximum likelihood G-test we rejected the null hypothesis of independence between active (`Sport hours`) and passive (`Week events`) sport activities ($G = 193.95$ p-value $= 0.226$).

In this study, we consider five possible covariates:

- `Gender`, gender of the respondent (1 = female, 0 = male),
- `Free time`, the respondents were asked to evaluate the amount of their free time on a scale from from (0 no free time) to 3 (having a lot of free time),
- `Life sed`, the respondents were asked to evaluate their sedentary lifestyle on a scale from 1 (no sedentary at all) to 100 (totally sedentary),
- `Stress`, the respondents were asked to evaluate the level of stress of their life on a scale from from 1 (not at all) to 100 (very stressful),
- `Age`, age in completed years of the respondents.

Table 1 reports the main characteristics of the individuals involved in the study. Evidence in the literature reveals that the probability of sports activity (`Sport hours`) decreases with age ([3, 14], among others). The reduction presents less difference between male and female for older adults [4]. `Gender`, in fact, is the

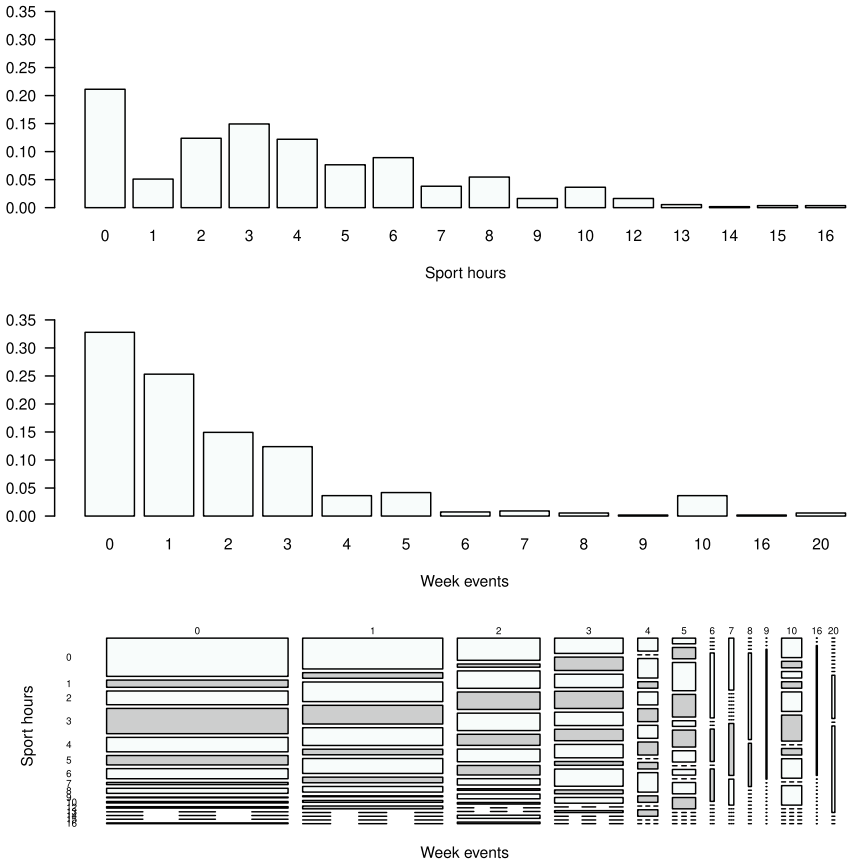


Fig. 1 Upper panel: Frequency distribution of the `Sport hours` per week. Center panel: Frequency distribution of the `Week events`. Bottom panel: Mosaic plot for previous counting variables

Table 1 Summary statistics concerning the characteristics of the 549 respondents of the survey

`Gender`			
Female (46.45%)	Male (53.55)%		
`Free time` (Self-reported on 4 point scale)			
0 (15.30%)	1 (38.80%)	2 (38.98%)	3 (6.92%)
`Life sed` (Self-reported on 100 point scale)			
Q_1 (30)	Q_2 (50)	Q_3 (70)	Mean (49.22)
`Stress` (Self-reported on 100 point scale)			
Q_1 (40)	Q_2 (70)	Q_3 (80)	Mean (60.52)
`Age`			
Min (14)	Max (75)	sd(8.27)	Mean (25.99)

other covariate that has a highly important influence on sports activity. There is evidence about the fact that men, in general, are not only more active than women (see [14, 15, 26, 36]) but they also show a higher frequency of sport participation when compared with active women (see [3, 15]). These differences may be attributed to biological factors, and cultural and social influences [27]. Women experience greater difficulties to access sport due to provision of facilities, commuting and time obligations [14]. One of the consequences for this is time constraint, with less `Free time` for women that often are the primary care providers for children and family members [21].

Another inverse relation with probability of sports activity has been highlighted in the context of sedentary life. Several definitions related to physical activity, inactivity and sedentary behaviors has been reported (see [44] and reference therein). Notice that sports is part of the physical activity spectrum corresponding to any institutionalized and organized practice, based on specific rules. Some very active individuals might not be sport athletes even though they regularly train and show a high level of physical activity. In comparison to the general population, older adults accumulate the highest level of sedentary time (see [18]). Sport activity is also related to free time with a focus on young people [20]. Studies on health related lifestyle choices—including leisure-time physical inactivity—are recently discussed (see [5]). These studies allow to connect sport (*in*)activity with time dedicated to watching sport on television with a gender determinant (see [24]). Watching sport, whether on TV or live at a venue, involves a prolonged sedentary period; it has been consistently associated with higher risk of adverse health outcomes. It has been also hypothesised that higher levels of sedentary behaviour would be linked to higher levels of stress (see [9, 41], among others).

3 Methods

The zero-inflated model is used when a count data set shows a large proportion of zeros. A bivariate zero-inflated model can be constructed by increasing the probability of the event $(Y_1 = 0, Y_2 = 0)$ and decreasing the other joint probabilities. In the following we firstly present its univariate version and then the extension to the bivariate case.

3.1 Zero Inflated Univariate Poisson Model

Zero-inflated poisson (ZIP) regression is used to model count data that have an excess of zero counts [35]. In the standard notation it represents a mixture which combines a Poisson count model and a degenerate distribution at zero. It has been proposed to deal with the case in which the number of zeros exceeds those expected for a regular Poisson distribution. The ZIP model can be considered as an adding structure

to the regular Poisson model. It allows separate consideration of those who are not exposed to an event of interest (no active respondents) and those who are exposed (respondents participating to some activity) and faced with the event several times during a specific time period [12].

Let U be a latent binary variable that indicates an individual's participation state: $U = 0$ if the subject does not participate to an event of interest (i.e., the subject is in the non-susceptible group); $U = 1$ if the subject participate to the event (i.e., the subject is in the susceptible group).

Let Y denote the event count, defined only when $U = 1$, which follows a Poisson distribution whose probability mass function is denoted by

$$f(y; \lambda|U = 1) = Pr(Y = y; \lambda|U = 1) = \exp(-\lambda)\lambda^y/y! \,.$$

When $U = 0$, we have the function $f(y; \lambda|U = 0) = \mathcal{I}(y = 0)$ which is a degenerate distribution among individuals who not participate to the event for $y = 0, 1, \ldots$. Let $\pi = Pr(U = 0)$ be the *inflation* weight. Here $\mathcal{I}(y = 0)$ is the indicator function, which is equal to one if $y = 0$; zero otherwise.

Let $f(y; \pi, \lambda) = Pr(Y = y; \pi, \lambda)$ be the unconditional or marginal distribution of Y, that is

$$\begin{aligned} f(y; \pi, \lambda) &= Pr(U = 0) f(y; \lambda|U = 0) + Pr(U = 1) f(y; \lambda|U = 1) \\ &= \pi\, \mathcal{I}(y = 0) + (1 - \pi)\, \frac{\exp(-\lambda)\lambda^y}{y!}, \qquad y = 0, 1, \ldots \end{aligned} \tag{1}$$

Both π and λ can be dependent on a set of covariates (not necessarily the same ones). The identifiability of a ZIP model is discussed in [37] whereas inference is in [16, 42, 43, 45], among others.

3.2 Zero Inflated Bivariate Poisson Model

The ZIBP regression [46] is used to model data that have an excess of counts in the $(0, 0)$ cell. It is obtained as a mixture of a Bivariate Poisson (BP) model and a degenerate distribution at $(0, 0)$. It is a special case of the Diagonal inflated model of [32]

$$f(x, y; \pi, \boldsymbol{\lambda}) = \pi \mathcal{I}(x = 0, y = 0) + (1 - \pi) f_{BP}(x, y|\lambda_1, \lambda_2, \lambda_3) \tag{2}$$

for $x = 0, 1, \ldots$ and $y = 0, 1, \ldots$; $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$ represents the parameter vector and f_{BP} denotes the BP distribution, that is

$$f_{BP}(x, y|\lambda_1, \lambda_2, \lambda_3) \exp^{(\lambda_1+\lambda_2+\lambda_3)} \frac{\lambda_1^x}{x!} \frac{\lambda_2^y}{y!} \sum_{i=0}^{min(x,y)} \binom{x}{i}\binom{y}{i} i! \left(\frac{\lambda_3}{\lambda_1\lambda_2}\right). \quad (3)$$

The marginal distribution of X and Y obtained from Eq. (3) are Poisson with $E(X) = \lambda_1 + \lambda_3$ and $E(Y) = \lambda_2 + \lambda_3$. The common parameter $\lambda_3 = cov(X, Y)$ is a measure of the dependence among the two variables.

Following a standard convection, we model the influence of covariates on $\boldsymbol{\lambda}$ via a loglinear formulation. The model can be expressed as

$$\begin{aligned}(X, Y) &\sim BP(\lambda_1, \lambda_2, \lambda_3)\\ \log(\lambda_1) &= \boldsymbol{W}_1\boldsymbol{\beta_1}\\ \log(\lambda_2) &= \boldsymbol{W}_2\boldsymbol{\beta_2}\\ \log(\lambda_3) &= \boldsymbol{W}_3\boldsymbol{\beta_3}\end{aligned}$$

where W_ℓ denotes the vector of covariates for λ_ℓ and $\boldsymbol{\beta}_\ell$ the corresponding vector of coefficients ($\ell = 1, 2, 3$).

Notice that when $\pi = 0$ the joint distribution is the simple BP model. Furthermore, even if -before introducing the inflation- the two variables are independent (that is $\lambda_3 = 0$) the corresponding inflated distribution introduces a degree of dependence among the two variables under examination. See [48] for the covariance structure of the ZIBP model. Inference for ZIBP models (estimation and hypothesis testing) is based on the likelihood function, that can be maximised via an EM algorithm as described in the Appendix of [33]. Diagnostics can be based on residuals analysis.

4 Data Analysis

We apply a BP model and a ZIBP model to our data. For comparative reasons we also consider a Double Poisson (DP) model obtained setting $\lambda_3 = 0$ in BP model. In this case the two variables are independent and the Bivariate Poisson distribution reduces to the product of two independent Poisson distributions.

The covariates included in the model have been selected via a preliminary analysis on the marginal distributions (1) of the two response variables.

More precisely, as a first attempt to capture the relationship among the number of respondents who express inactivity in the two examined dependent variables and all regressors included in the data, we fitted on each marginal distribution a basic Poisson model. Subsequently, we used a backward selection criteria to identify the relevant covariates. By a visual inspection of the marginal distributions we noticed a higher number of zero observations in comparison to the basic count data distributions. Hence we fitted a ZIP model (Sect. 3) for the marginal distributions using the same set of covariates selected for the marginal Poisson models. All the predictors in both count and inflation portions of the models were statistically significant. The previous

models fit the data significantly better than the null model, i.e., the intercept-only model (Deviance 264.337 with p-value $< 0.001 (df = 7)$ and 226.972 with p-value $< 0.001 (df = 8)$, respectively). Furthermore, performing a Vuong test [13] between the zero-inflated models and the standard Poisson regressions for the two marginal distributions we observed that the former are an improvement over the latter (test-statistic is -3.030 with a p-value < 0.001 and -6.259 with a p-value < 0.001, respectively).

In the examined bivariate model, we do not take into account any covariate on λ_3, hence we assume a constant covariance term. With respect to this point we tested several covariates for λ_3 but none of them resulted relevant. Furthermore, as reported in [32] models with constant covariance terms are easier to interpret. `Gender` and `life sed` are considered as covariates both on λ_1 and on λ_2. `Free time` and `Age` affect `Sport hours` whereas `Stress` weighs on `Week events`. To estimate the models we use the R package `bivpois` [32].

The results are reported in Table 2. Both AIC and BIC indicate that the ZIBP model is the most suitable for our data. Similar results can be obtained via a subsequent pairwise comparison of the examined models. More precisely the likelihood ratio test statics (LRTS) for comparing model DP versus BP is 10.890 (p-value $< 0.01 df = 1$) and the LRT for comparing model ZIBP versus BP is 77.612 (p-value $< 0.01 df = 1$). Finally, the Vuong test between BP and ZIBP models confirms the result obtained with AIC and BIC criteria in favour of ZIBP model (test-statistic is -3.215 with p-value < 0.01).

Table 2 Results for the BP, ZIBP and DP models

	Covariate	BP model		ZIBP model		DP model	
		Coef.	St.Er.	Coef.	St.Er	Coef.	St.Er
λ_1	`Constant`	1.811	0.110	1.843	0.101	1.839	0.109
	`Gender`	−0.179	0.043	−0.101	0.045	−0.175	0.044
	`Life sed`	−0.014	0.001	−0.013	0.001	−0.014	0.001
	`Age`	−0.007	0.003	−0.007	0.003	−0.006	0.003
	`Free time`	0.231	0.028	0.210	0.030	0.228	0.030
λ_2	`Constant`	1.264	0.089	1.253	0.084	1.326	0.080
	`Gender`	−1.391	0.110	−1.236	0.110	−1.246	0.074
	`Life sed`	−0.007	0.001	−0.006	0.001	−0.007	0.001
	`Stress`	0.002	0.001	0.003	0.001	0.002	0.001
λ_3	`Constant`	−1.718	0.766	−2.237	1.692		
π				0.069	0.001		
Number of parameters			10		11		9
Log-likelihood			−2426.032		−2387.226		−2431.477
AIC			4872.065		4796.452		4880.954
BIC			4922.077		4851.466		4925.965

Furthermore, we used 200 bootstrap replications to obtain the standard errors of the coefficients, asymptotic t-tests have been performed to evaluate the significance of the coefficients. The effects of all covariates are statistically significant.

Furthermore, the effect of inflation term is significant (p-value < 0.01).

From fitted ZIBP model we obtain

$$\log(\lambda_1) = 1.843 - 0.101(\texttt{Gender}) - 0.013(\texttt{Life sed}) - 0.007(\texttt{Age}) + 0.210(\texttt{Free time})$$
$$\log(\lambda_2) = 1.253 - 1.236(\texttt{Gender}) - 0.006(\texttt{Life sed}) + 0.003(\texttt{Stress})$$
$$\log(\lambda_3) = -2.237.$$

In accordance with the literature in our study active sport is more frequent among young people than older one, men are more involved than women both in active and passive sport activities and a high level of stress is positively related to the practice of passive sport activity. What is quite surprising is the behaviour of variable `Life sed` that is negatively related both with sport activity (as expected) and inactivity. One possible explanation for the latter is that sedentary people are not necessary interested in watching sports and they may spend their free time with other activities, such as watching movies or reading (see [38], among others).

The value of mixing proportion, related to the effect of $(0, 0)$ for both the marginal distributions, is not very high ($\simeq$6.9%) even if it has a relevant effect on most of the estimated parameters. The negative value of the constant term in $\widehat{\lambda}_3$ corresponds to an estimated covariance term equal to $\widehat{\lambda}_3 = e^{-2.237} = 0.107$ which indicates rather small correlation between the two examined responses.

From a diagnostic point of view, the graphical residuals plot in Fig. 2 does not provide evidence of the presence of outliers or anomalous data. The residuals of X

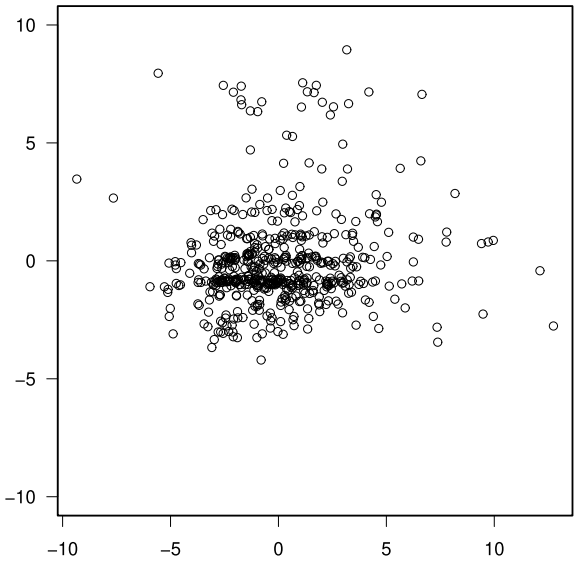


Fig. 2 Scatter-plot of the residuals related to the estimated ZIBP model

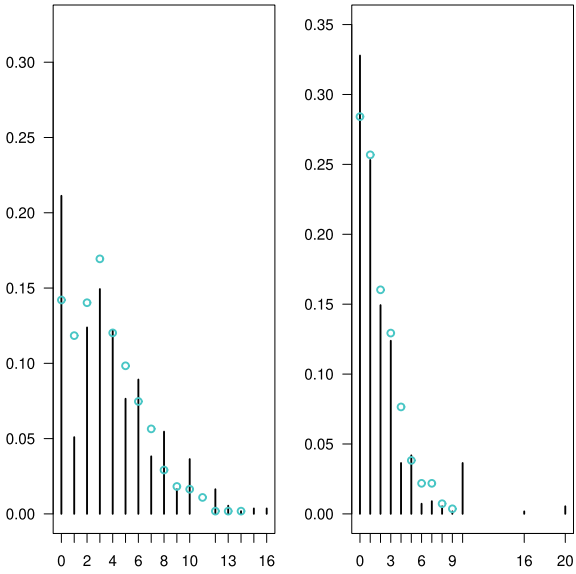


Fig. 3 Observed marginal frequencies and fitted probabilities (dots) concerning `Sport hours` (left panel) and `Week events` (right panel) for the ZIBP model

and Y represent deviations between observed and fitted values; they are given by $X - \hat{X}$ and $Y - \hat{Y}$; where $\hat{X}$ and $\hat{Y}$ are the fitted values for X and Y respectively. The latter are given by

$$\hat{X} = (1-\pi)(\lambda_1 + \lambda_3) \qquad \hat{Y} = (1-\pi)(\lambda_2 + \lambda_3) \ \textit{if}\ X \neq Y$$
$$\hat{X} = (1-\pi)(\lambda_1 + \lambda_3) + \pi \mathbb{E}_D(X) \quad \hat{Y} = (1-\pi)(\lambda_2 + \lambda_3) + \pi \mathbb{E}_D(X) \ \textit{if}\ X = Y$$

where $\mathbb{E}_D(X)$ is the mean of the distribution used to inflate the diagonal. In our case $\mathbb{E}_D(X) = 1$ if $X = Y = 0$, $\mathbb{E}_D(X) = 0$ otherwise.

Fitted marginal distributions versus the observed ones are displayed in Fig. 3. Some mismatch appears for the zero categories in both the marginal distributions opening to possible further studies to improve the results. Figure 4 displays the fitted joint distribution of the two responses represented by the shades of the bins.

5 Conclusion

The aim of this paper was to show, starting from a sport (*in*)activity case study, how a class of bivariate models for counting data, called ZIBP model, can be used in the analyses where zero inflated marginal distributions are present.

The advantages of the ZIBP model in the analysis of the bivariate counting variables include: an appropriate treatment of the counting nature of the data; the possibility to measure the connection between the two dependent variables related to different

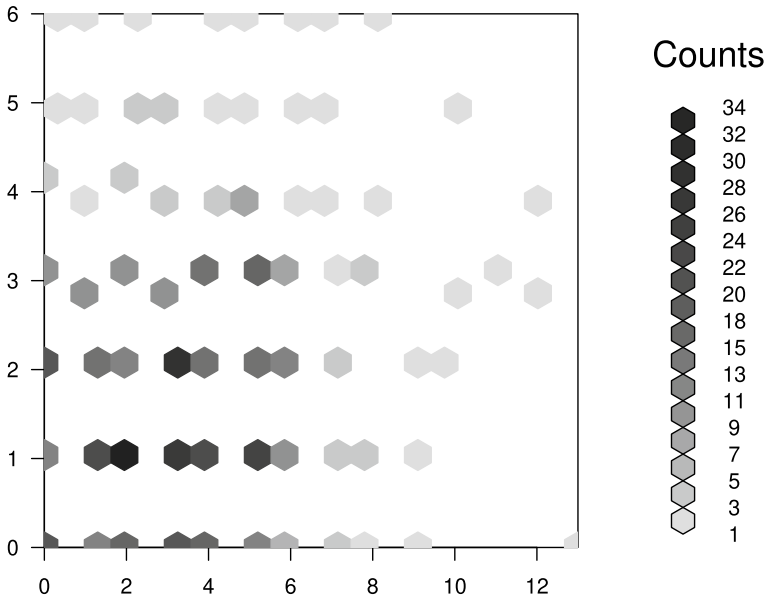


Fig. 4 Joint probability distribution concerning `Sport hours` and `Week events` for the ZIBP model

perspectives of sport activity (active and passive sport practice); the possibility to investigate if and how individual characteristics affect both categorical variables of interest; the feasibility to measure the mixing proportion which explains the increased number of zeros indicating people that are not interested in sports at all.

Results obtained in the analysis underline that young people practice more sport than older one, men are more involved than women both in active and passive sport activities and an high level of stress is positively related to the practice of passive sport activity. Sedentary life affects sport by reducing the practice but watching sport. A preliminary study based on the information content included in the survey did not give evidence of possible relevant covariates affecting the mixing proportion.

A study of observed marginal frequencies and fitted probabilities allows to point out the presence of several focal values around some specific real numbers, which are caused by rounding effects [39] (see for instance the *shelter* at 10 for both counting variables). This suggests the possibility to extend the mixture model taking into account also inflation caused by specific response styles; see [1].

Furthermore, more general studies based on the selection of covariates may be developed. Finally, tests on the zero inflated component by mimicking the approaches in [42, 43, 45], among others, may be also a further topic of research.

References

1. Allen, M.: The Sage Encyclopedia of Communication Research Methods, vol. 1–4. SAGE Publications, Thousand Oaks (2017)
2. AlMuhayfith, F.E., Alzaid, A.A., Omair, M.A.: On bivariate Poisson regression models. J. King Saud Univ. - Sci. **28**, 178–189 (2016)
3. Barber, N., Havitz, M.E.: Canadian participation rates in ten sport and fitness activities. J. Sport. Manage. **15**, 51–76 (2001)
4. Bauman, A., Sallis, J., Dzewaltowski, D., Owen, N.: Toward a better understanding of the influences on physical activity. Amer. J. Prev. Med. **23**(2S), 5–14 (2002)
5. Beenackers, M.A., Kamphuis, C.B., Giskes, K., Brug, J., Kunst, A.E., Burdorf, A., van Lenthe, F.J.: Socioeconomic inequalities in occupational, leisure-time, and transport related physical activity among European adults: a systematic review. Int. J. Behav. Nutr. Phys. Activ. **9**, 116 (2012)
6. Berkhout, P., Plug, E.: A bivariate Poisson count data model using conditional probabilities. Stat. Neerl. **58**, 349–364 (2004)
7. Bermùdez, L.: A priori ratemaking using bivariate Poisson regression models. Insur. Math. Econom. **44**, 135–141 (2009)
8. Bermùdez, L., Karlis, D.: A finite mixture of bivariate Poisson regression models with an application to insurance rate making. Comput. Stat. Data Anal. **56**, 3988–3999 (2012)
9. Charney, D.S., Manji, H.K.: Life stress, genes, and depression: multiple pathways lead to increased risk and new opportunities for intervention. Sci. STKE (2004)
10. Cheung, Y.B., Lam, K.F.: Bivariate poisson-poisson model of zero-inflated absenteeism data. Stat. Med. **25**, 3707–3717 (2006)
11. Chib, S., Winkelmann, R.: Markov chain Monte Carlo analysis of correlated count data. J. Bus. Econ. Stat. **19**, 428–435 (2001)
12. Dietz, K., Böhning, D.: The use of two-component mixture models with one completely or partly known component. Comput. Stat. **12**, 219–234 (1997)
13. Desmarais, B.A., Harden, J.J.: Testing for Zero Inflation in Count Models: Bias Correction for the Vuong Test. Stata J. **13**, 810–835 (2013)
14. Downward, P., Rasciute, S.: The relative demands for sports and leisure in England. Eur. Sport Manag. Q. **10**(2), 189–214 (2010)
15. Eberth, B., Smith, M.: Modelling the participation decision and duration of sporting activity in Scotland. Econom. Model. **27**, 822–834 (2010)
16. El-Shaarawi, A.H.: Some goodness-of-fit methods for the Poisson plus added zeros distribution. Appl. Environ. Micro. **49**, 1304–1306 (1985)
17. Funk, D.C.: Consumer Behaviour in Sport and Events. Business & Economics, Routledge (2008)
18. Gennuso, K.P., Gangnon, R.E., Matthews, C.E., Thraen-Borowski, K.M., Colbert, L.H.: Sedentary behavior, physical activity, and markers of health in older adults. Med. Sci. Sports Exerc. **45**, 1493–1500 (2013)
19. Ghitany, M.E., Karlis, D., Al-Mutairi, D.K., Al-Awadhi, F.A.: An EM algorithm for multivariate mixed Poisson regression models and its application. Appl. Math. Sci. **6**, 6843–6856 (2012)
20. Green, K.: Key Themes in Youth Sport. Routledge (2010)
21. Grima, S., Grima, A., Thalassinos, E., Seychell, S., Spiteri, J.V.: Theoretical models for sport participation: literature review. Int. J. Econ. Bus. Admin. **V**, 94–116 (2017)
22. Gurmu, S., Elder, J.: Generalized bivariate count data regression models. Econom. Lett. **68**, 31–36 (2000)
23. Gurmu, S., Elder, J.: A bivariate zero-inflated count data regression model with unrestricted correlation. Econom. Lett. **100**, 245–248 (2008)
24. Hamer, M., Weiler, R., Stamatakis, E.: Watching sport on television, physical activity, and risk of obesity in older adults. BMC Public Health **14**, 10 (2014)
25. Holgate, P.: Estimation for the bivariate Poisson distribution. Biometrika **51**, 241–245 (1964)

26. Hovemann, G., Wicker, P.: Determinants of sport participation in the European Union. European J. Sport Soc. **6** (1), 51–59 (2009)
27. Humphreys, B., Ruseski, J.E.: The economic choice of participation and time spent in physical activity and sport in Canada. Working Paper No 2010-14. Department of Economics, University of Alberta (2010)
28. Iannario, M.: Modelling shelter choices in a class of mixture models for ordinal responses. Stat Meth. Appl. **21**, 1–22 (2012)
29. Iannario, M., Simone, R.: Zero-inflated ordinal data models with application to sport (in)activity. In: Abbruzzo, A., Brentari, E., Chiodi, M., Piacentino, D. (eds.) Book of Short Papers SIS 2018, pp. 89–96 (2018)
30. Johnson, N.L., Kotz, S.: Discrete Distributions. Wiley, New York (1969)
31. Karlis, D., Meligkotsidou, L.: Finite mixtures of multivariate Poisson distributions with application. J. Stat. Plan. Infer. **137**, 1942–1960 (2007)
32. Karlis, D., Ntzoufras, I.: Analysis of sports data by using bivariate Poisson models. Statistician **52**, 381–393 (2003)
33. Karlis, D., Ntzoufras, I.: Bivariate Poisson and diagonal inflated bivariate Poisson regression models. J. Stat. Softw. **14**, 1–36 (2005)
34. Kocherlakota, S., Kocherlakota, K.: Bivariate Discrete Distributions. Marcel Dekker, New York (1992)
35. Lambert, D.: Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics **34**, 1–14 (1992)
36. Lera-López, F., Rapún-Gárate, M.: The demand for sport: sport consumption and participation models. J. Sport Manage. **21**, 103–122 (2007)
37. Li, C.S.: Identifiability of zero-inflated Poisson models. Brazilian J. Probab. Stat. **26**, 306–312 (2012)
38. Wang, C.L.: Handbook of Research on the Impact of Fandom in Society and Consumerism. IGI Global, Hershey (2019)
39. Manski, C.F., Molinari, F.: Rounding probabilistic expectations in surveys. J. Bus. Econ. Stat. **28**, 219–231 (2010)
40. Munkin, M.K., Trivedi, P.K.: Simulated maximum likelihood estimation of multivariate mixed-Poisson regression models, with application. Econometrics J. **2**, 29–48 (1999)
41. Panahi, S., Tremblay, A.: Sedentariness and health: is sedentary behavior more than just physical inactivity? Front. Public Health **6**, 258 (2018)
42. Ridout, M., Hinde, J., Demétrio, G.: A score test for testing a zero-inflated Poisson regression model against zero-inflated negative Binomial alternatives. Biometrics **57**, 219–223 (2001)
43. Thas, O., Rayner, J.C.W.: Smooth tests for the zero-inflated Poisson distribution. Biometrics **61**, 808–815 (2005)
44. Thivel, D., Tremblay, A., Genin, P.M., Panahi, S., Rivière, D., Duclos, M.: Physical activity, inactivity, and sedentary behaviors: definitions and implications in occupational health. Front. Public Health **6**, 288 (2018)
45. van den Broek, J.: A score test for zero inflation in a Poisson distribution. Biometrics **51**, 738–743 (1995)
46. Walhin, J.F.: Bivariate ZIP models. Biometrical J. **43**, 147–160 (2001)
47. Wang, P.: A bivariate zero-inflated negative binomial regression model for count data with excess zeros. Econ. Lett. **78**, 373–378 (2003)
48. Wang, K., Lee, A.H., Yau, K.K.W., Carrivick, P.J.W.: A bivariate zero-inflated Poisson regression model to analyze occupational injuries. Accid. Anal. Prev. **35**, 625–629 (2003)

Network-Based Dimensionality Reduction for Textual Datasets

Michelangelo Misuraca, Germana Scepi, and Maria Spano

Abstract There is an increasing interest in developing statistical tools for extracting information from textual datasets. In a text mining framework, a knowledge discovery process typically implies the reduction of the vocabulary dimensionality, via a feature selection or a feature extraction approach. Here we propose a strategy designed to reduce the dimensionality of textual datasets through a network-based procedure. Network tools allow performing the reduction taking into account the association relations among terms used in the texts. The effectiveness of this strategy is shown by analysing a set of tweets about the recent COVID-19 global pandemic.

Keywords Vector space model · Network analysis · Community detection

1 Introduction

The ever increasing progress of information technology, joined with the growth of the Internet, has fastened the transition from analog to digital in data communication and storage. This relentless revolution necessarily involved the statistical reasoning and the wide range of methods traditionally used to analyse data. All the applicative domains that Statistics can support in searching trends, revealing patterns or explaining relations, have experimented an over-availability of data. This huge volume of data has required the development of mining techniques with the aim of discovering and extracting knowledge, since often a little part of the collected data is relevant and informative to understand a given phenomenon. This scenario is even more common when the phenomena are investigated by analysing textual datasets.

M. Misuraca
DiScAG, University of Calabria, Rende, Italy
e-mail: michelangelo.misuraca@unical.it

G. Scepi · M. Spano (✉)
DiSeS, University of Naples Federico II, Naples, Italy
e-mail: maria.spano@unina.it

G. Scepi
e-mail: germana.scepi@unina.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings
in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_12

Texts may express a wide and rich range of information, but this information is encoded in a form difficult to be processed. In order to recognise the basic components of a text belonging to a given collection—commonly referred to as *words*—it is necessary to perform different sequential operations aiming at obtaining a structured representation of the enclosed textual content. Because of the linguistic variety, to have an algebraic formalisation of texts the entire vocabulary of the text collection has to be taken into account, even if only a subset of this vocabulary is actually used in a single text. Due to this peculiar setting, textual datasets have a high-dimensionality and high sparsity. This means that analysing textual data not only incurs unaffordable memory requirements but also high computational costs. Moreover, the high-dimensionality affects the generalisability of the results, making problematic any method that requires statistical significance. To avoid these problems, textual datasets can be summarised by finding "narrower" representations that in some sense are close to the original ones, i.e. dimensionality is reduced.

Dimensionality reduction has been addressed in various ways in text mining and natural language processing, mainly by filtering from the dimensions related to the vocabulary the meaningless ones or by combining the dimensions in new compact but informative entities. Both approaches have strengths and weaknesses, sharing the information encoding and the data structures used to organise it. In this paper, we propose a strategy based on an alternative viewpoint. Starting from a peculiar data structure, we use a network-based approach for reducing the dimensionality of huge datasets. The advantage of applying this strategy consists in recovering the contexts lost due of text encoding, retaining at the same time a higher readability.

This work is structured as follows. In Sect. 2, dimensionality reduction issue is briefly introduced and the reference literature is reviewed. Section 3 describes the problem of dimensionality reduction from a network perspective and defines the proposed strategy. In Sect. 4, the effectiveness of the proposal is showed by analysing a set of tweets about COVID-19 global pandemic. Final remarks and possible future developments of the strategy are discussed in Sect. 5.

2 Background and Related Work

Quantitatively analysing the information enclosed in a set of texts is challenging, because a textual content written in natural language does not follow a defined data model. It is then necessary to prepare texts via a multi-stage process for extracting a set of structured data that can be handled with statistical techniques. Initially, texts are parsed and tokenised to decompose their content into the basic meaning components, i.e. the employed terms. The peculiar coding used in this step is commonly known as *bag of words*, inasmuch as texts are viewed as a multiset of terms, disregarding each grammatical and syntactical roles but keeping multiplicity.

Typically, the *vector space model* formalised by Salton, Wong and Yang [38] is used to represent texts in an algebraic fashion. Applying this model, the unstructured content of texts can be converted into a set of structured data. Let consider a set

$\mathcal{D}$ containing n texts (namely, a *corpus*) and a set $\mathcal{T}$ listing the different p terms used by the texts in $\mathcal{D}$ (namely, a *vocabulary*). Each text d_i ($i = 1, \ldots, n$) can be represented as a vector $\mathbf{d_i} = \{w_{i1}\ w_{i2}\ \ldots\ w_{ij}\ \ldots\ w_{ip}\}$ in a p-dimensional vector space $\mathcal{R}^p$ spanned by the terms of the vocabulary. When a term j ($j = 1, \ldots, p$) occurs in the text i, the corresponding value w_{ij} in the vector is non-zero. The latter quantity can be interpreted as a weight measuring the importance of the term in the text, i.e. how much that term contributes in expounding that textual content. Various *term-weighting* schemes have been proposed in the literature [37], taking into account the relative importance of the terms with respect to each single text (e.g. *boolean* scheme, *term frequency* scheme) or also incorporating the discrimination power of the terms with respect to the entire corpus (e.g. term frequency plus inverse document frequency—known as *tf-idf*) [see 31].

Once texts have been pre-treated and converted into vectors, and terms have been weighted according to a given scheme, it is possible to embed the data into different types of structures. The juxtaposition of the n different p-dimensional text-vectors allows building a particular contingency table $\mathbf{F}$ called *lexical table* [24]. When a term-weighting scheme based on the raw frequencies of terms is considered—i.e. merely counting the number of occurrences of each term in each text—the row marginals of $\mathbf{F}$ can be interpreted as the texts' lengths with regard to the employed terms whereas the column marginals represent the terms' occurrences in all the texts, i.e. the frequency distribution of the terms in the corpus vocabulary.

One of the critical issues in the analysis of large corpora is the high-dimensionality. Typically, lexical tables are large and very sparse matrix, since each single term in the vocabulary represents a dimension in the vector space of texts even if it has not be used to express part of the specific textual content. This noise in the data has to be reduced because it significantly increases the computational complexity and may leads to unreliable results in the subsequent analyses performed to discover association relationships between the texts. During the pre-treatment step, it is possible to carry on various transformation to reduce language variability. *Normalisation* allows resolving trivial cases of variability, operating on case sensitivity, spelling and special characters (e.g. accented letters in Romance languages). *Stemming* and *lemmatisation* allow instead reducing variability at a morphological level, substituting inflected terms with their roots and their canonical forms, respectively. Even if these standard operations can partially help to cope the dimensionality drawback, it is necessary to perform specific procedures aiming at reducing the high-dimensionality. In a text mining framework, similarly to other applicative domains in which high-dimensionality may occur, *feature selection* and *feature extraction* are usually considered to deal with dimensionality reduction [26].

Feature selection aims at finding a subset of the original vocabulary, focusing on the contribution of a term in drawing the textual content. One of the main advantages of this approach is that the selected terms retain the original meaning and provide easily interpretable results. The most common solution is using a *stop-list* to prune the corpus vocabulary. Typically, in a stop-list we can find non-informative terms with general and weaker lexical meaning (e.g. the most common terms used in the language and in the specific analysed domain). Other methods aims at filtering

the content-bearing terms in the corpus [7]. The simplest criterion for selecting terms is based on *document frequency*, i.e. the number of corpus' texts in which a term occurs, keeping only the terms with a frequency greater than a predetermined threshold. A well-known alternative is the *tf-idf*—used for term-weighting—which jointly considers the term frequency and the *inverse document frequency* [39]. More recently, some approaches based on *term variance* have been proposed [e.g. 16, 27], computing the variance of each term in the corpus and carrying on the selection according to a given threshold. Other complex methods, such as *term strength* [44], *term contribution* [28] and *entropy-based ranking* [13], involve text similarity to determine if a term has to be saved or deleted from the vocabulary. All the above-described alternatives can be used when no additional information on the texts is available, concerning for example the authorship, the document type or the subjects enclosed in the texts. In a supervised framework, when text categories are known a priori and used in the analytical process, it is possible to select terms with metrics like χ^2, *gain ratio* or *information gain* [15, 45].

Differently from feature selection methods, feature extraction aims at minimising the amount of terms required to describe a large corpus. The extraction process allow deriving a set of new constructs through some functional mapping, considering a linear combination of the original terms. The relevant information can be represented in a vector space with a lower dimensionality. An obvious drawback of feature extraction is that data compression leads to new constructs that may have an unclear lexical or semantic meaning, thus the results are often harder to read and interpret. Factorial techniques like *lexical correspondences analysis* [LCA: 24] and *latent semantic analysis* [LSA: 14] are widely used in text analysis. Although developed in different domains and for different tasks, LCA and LSA share a common algebraic framework. The lexical table **F** is decomposed in both cases via a *generalised singular value decomposition* [19]. The difference between LCA and LSA relies on how rows (the texts) and columns (the terms) are analysed, introducing proper orthonormalising constraints [4]. In the first case, the constraints are based on the marginal distributions of **F**, introducing a weighting scheme and a weighted-Euclidean metric based on the χ^2 statistics. The effect of these constraints is that all the terms (normalising their number of occurrences) and to all the texts (normalising their lengths) has the same importance in the analysis. In the second case, unitary weights and classic Euclidean metric are used to constrain both terms and texts, giving more importance to the most occurring terms and to the longest texts. Starting from LCA and LSA, several related techniques have been developed [e.g. 11, 29, 42]. Some authors argued that the reduction of dimensionality performed by singular value decomposition contradict physical reality, since term frequencies are non-negative where instead the linear combinations may have positive and negative coefficients [5, 35, 47]. To obtain a low-rank space with additive combinations of the original terms, a *non-negative matrix factorisation* can be used [NMF: 34]. A non-linear perspective to the problem of reducing dimensionality is offered by the so-called *manifold learning*, a research issue started in 2000 with the work of Tenenbaum, de Silva and Langford on *isometric feature mapping* [41] and the work of Roweis and Saul on *local linear embedding* [36]. The rationale behind these techniques is that the minimal distance along some

(unknown) surface parameterising the textual dataset may offer a better indication than Euclidean distance in a linear standpoint of the problem [25].

The dimensionality reduction problem can also be seen from a probabilistic standpoint by considering the so-called *topic modelling* [12]. This approach encompasses a family of generative statistical models used to highlight semantic patterns reflecting the underlying (not observable) topics of a corpus. These topics are easily interpretable because each topic is characterised by the terms in the corpus most strongly associated with. Nevertheless, topic labelling is not an easy task. The automatic assignment of a unique name able to identify each topic can be coped in several ways [e.g. 2, 22]. The original idea of topic modelling has been directly derived from LSA in 1999, when Hofmann introduced the *probabilistic-LSA* [pLSA 20]. Differently from techniques based on singular value decomposition, pLSA considers a probabilistic mixture decomposition of the lexical table [21]. The probability that a term occurs in a text is modelled as a mixture of conditionally independent multinomial distributions. *Latent dirichlet allocation* [LDA: 8] is a Bayesian model that overcomes some pLSA drawbacks (e.g. the large number of parameters to take into account), using a Dirichlet distributional law to model the probability of term appearance in a text.

3 A Network-Based Strategy for Dimensionality Reduction

Feature selection and feature extraction are both performed on the lexical table, following the bag of word logic. These means that the vector space model does not consider the context in which each term is used. It is possible to get back part of the structural and semantic information hold by the texts in the corpus by constructing a squared $terms \times terms$ co-occurrence matrix. Each element of this matrix represents the number of times two terms co-occur in the corpus. In a network analysis framework, the co-occurrence matrix can be seen as an adjacency matrix and graphically represented as an undirected weighted graph, where the nodes are the different terms used in the corpus and the ties pair the co-occurring terms. The adjacency matrix is usually displayed in a non-metric space, using topological representations. Nevertheless, some author has shown that it is possible to effectively project a graph in a Euclidean space [e.g. 23].

Since textual graphs allow visualising both single terms and subsets of terms frequently co-occurring together, it is possible to perform dimensionality reduction starting from the underlying data structure. Compared with feature extraction methods, a network-based strategy has the advantage of preserving the lexical and semantic meaning of the terms in each step. Moreover, due to high-dimensionality, Euclidean distance loses its readability and interpretability [1]. This condition—known as *curse of dimensionality*—has peculiar effects on several statistical methods [3]. The use of network-based strategy may avoid this problem, since term relationships are expressed as similarities rather than distances.

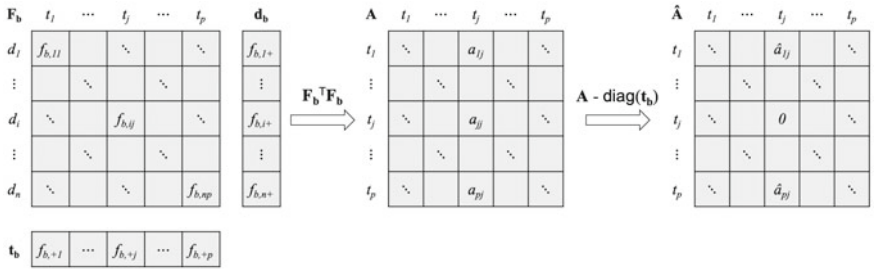


Fig. 1 Textual datasets: from binary lexical table to co-occurrence matrix

Starting from the lexical table $\mathbf{F}$ [see Sect. 2], where each element f_{ij} represents the number of occurrences of the jth term in the ith text. Let encode $\mathbf{F}$ with a boolean term-weighting scheme to obtain a binary lexical table $\mathbf{F_b}$. The generic element $f_{b,ij}$ is equal to 1 when the term j occurs at least once in text i, 0 otherwise. The text marginal distribution $\mathbf{d_b} = \mathbf{F_b 1_p} = \{f_{b,1+} \dots f_{b,n+}\}$ states how many terms of the vocabulary each text embodies, whereas the term marginal distribution $\mathbf{t_b} = \mathbf{F_b^\top 1_n} = \{f_{b,+1} \dots f_{b,+p}\}$ states how many texts enclose each term ($|\mathbf{1_n}| = n$ and $|\mathbf{1_p}| = p$, respectively). From matrix $\mathbf{F_b}$, we derive the p-dimensional $terms \times terms$ co-occurrence matrix $\mathbf{A}$ by the matrix product $\mathbf{A} = \mathbf{F_b}^\top \mathbf{F_b}$. The generic element $a_{jj'}$ is the number of texts in which the term j and the term j' co-occur ($j \neq j'$), whereas the diagonal element a_{jj} (known as *loops*) counts how many texts contain the term j ($a_{jj} \equiv f_{b,+j}$). Since we are not interested in loops, the adjacency matrix $\hat{\mathbf{A}}$ with null diagonal elements can be easily obtained as $\hat{\mathbf{A}} = \mathbf{A} - \mathsf{diag}(\mathbf{t_b})$. Figure 1 schematically depicts the different data structures above described.

According to graph theory, $\hat{\mathbf{A}}$ is an adjacency matrix that can be depicted as an undirected weighted graph $\mathsf{G}\{\mathcal{V}, \mathcal{E}, \mathcal{W}\}$, where $\mathcal{V}$ is the non-empty, finite set of the p terms, $\mathcal{E}$ is the set of l ties pairing the different terms and $\mathcal{W}$ is a set of weights on $\mathcal{E}$ expressing the level of co-occurrence between terms. If the interest is visualising term association with respect to diverse criteria, the weighting set can be substituted by other similarity measures, such as the *equivalence index* [10].

3.1 Dimensionality Reduction as a Community Detection Problem

To reduce dimensionality in a textual dataset, terms used together in the corpus can be grouped taking into account the graph structure associated with $\hat{\mathbf{A}}$.

In a network analysis framework, the task of grouping subsets of nodes belonging to a large graph is commonly known as *community detection* [32]. According to the reference literature, there is not a shared definition of community, even if it is well known that most real-world networks display this kind of structure [17]. Neverthe-

less, community detection from a theoretical point of view is not very different from clustering: homogeneous nodes are partitioned into the same group, while heterogeneous nodes should be kept in different groups. Thus, communities can be seen as disjoint subgraphs $\mathsf{g}_k \in \mathsf{G}$ ($k = 1, \ldots, K$), with high inner cohesion and low outer connection with other subgraphs [43]. Several algorithms have been proposed to perform a community detection procedure. Many authors proposed a taxonomy of community detection methods based on the different criteria. According to Tang and Liu [40], a rough classification can be done distinguishing methods focused on nodes properties, on group properties or on network properties. Network-centric procedures, in particular, look at graph topology with the aim of partitioning the nodes into a set of disjoint groups, typically by optimising a given criterion.

In a study dating back to 2004, Newman and Girvan introduced a new measure called *modularity* to quantify the strength of a graph partition [33]. The logic of modularity is that a random graph is not expected to have a community structure. The possible existence of communities is revealed by comparing the actual density of ties in a subgraph and the expected density of ties in the subgraph under the hypothesis that the nodes are attached regardless of a community structure. This expected ties' density depends on the chosen null model, i.e. the original graph keeping some of its structural properties but without communities. Modularity calibrates the quality of community partitions thus can be used as an objective measure to maximise. Given the co-occurrence matrix $\hat{\mathbf{A}}$, the modularity for a K-dimensional partition is:

$$Q = \frac{1}{2l} \sum_{k \in K} \sum_{j, j' \in k} \left[\hat{a}_{jj'} - \frac{\delta_j \delta_{j'}}{2l} \right] = \sum_{k \in K} \left[\frac{l_k}{l} - \left(\frac{\delta_k}{2l} \right)^2 \right] \tag{1}$$

where l is the total number of ties in the graph, l_k is the number of inner ties of the community g_k, δ_j is the *degree* of a term j—i.e. how many terms are linked to j—and δ_k is the sum of the terms' degrees in the community g_k. In case of weighed ties, Eq. 1 is modified replacing l by the sum of weights associated to the different ties and δ by the terms' *strengths*, i.e. the weighed degrees. It is also possible to consider a modularity matrix $\mathbf{M}$, with a generic element $m_{jj'} = \hat{a}_{jj'} - (\delta_j \delta_{j'}/2l)$. According to this formalisation, modularity in Eq. 1 can be rewritten as:

$$Q = \frac{1}{2l} \sum_k \mathbf{s}_k \mathbf{M} \mathbf{s}_k = \frac{1}{2l} \mathrm{tr}(\mathbf{S}^\mathsf{T} \mathbf{M} \mathbf{S}) \tag{2}$$

where $\mathbf{S}$ is a community indicator matrix with $s_{jk} = 1$ if a term j belongs to a community g_k. Modularity may assumes values in a range $[-1; 1]$, with 0 indicating that the community division is not better than random and 1 indicating instead a strong community structure. From an empirical standpoint, however, it has been observed that Q usually falls in a $[0.3; 0.7]$ subinterval.

In a recent comparative study on the performances of modularity-based algorithms and other type algorithms for community detection [46], Yang, Algesheimer

and Tessone showed that the *multilevel* algorithm proposed by Blondel et al. [9]—also known as *Louvain* algorithm—can be effectively applied in the case of very large networks, outperforming the other alternatives both regarding accuracy and computing time. The multilevel algorithm performs a greedy agglomerative community detection using an optimisation criterion based on modularity. In particular, it performs a 2-fold procedure with a greedy optimisation aiming at finding local maxima of modularity followed by a hierarchical refinement (see below for the pseudo-code).

Algorithm 1: Multilevel algorithm [9]

Data: a p-dimensional adjacency matrix $\hat{\mathbf{A}} \rightarrow \mathsf{G}\{\mathcal{V}, \mathcal{E}, \mathcal{W}\}$
Result: $\mathsf{G} = \{\bigcup_k \mathsf{g}_k\}$, $\mathsf{g}_k \cap \mathsf{g}_{k'} = \varnothing \rightarrow$ a p-dimensional membership vector $\mathbf{c}$

```
1  begin
2      repeat
3          PHASE1 (Graph partitioning via greedy modularity optimisation)
4          assign a different community to each node
5          while local maximum reached do
6              forall the nodes v_j do
7                  move v_j from its community to a neighbouring community
8                  calculate the modularity gain ΔQ
9                  if modularity gain is positive then
10                     move v_j to the neighbouring community with greater ΔQ
11                 else
12                     no change
13         PHASE2 (Communities' agglomeration in super-nodes)
14         every community g_k forms a super-node v_k
15         l_kk' = Σ{ties between g_k and g_k'}, where l_kk' is the tie between v_k and v_k'
16     until no improvement is possible
```

The greedy optimisation consists in moving each node in one of its neighbouring community to obtain the maximal modularity gain. The modularity gain obtained by moving a node j to a community k can be formalised as:

$$\Delta Q_{k,j} = \left[\frac{l_{j,k}}{l} - \frac{\delta_k \delta_j}{2l^2}\right] \tag{3}$$

Nodes can be moved several times and the phase stops only when a local maximum is obtained, i.e. when moving nodes does not lead to a positive modularity gain. The hierarchical refinement consists instead in building a meta-graph whose nodes are the communities found in the previous phase and ties represent the sum of connections between the communities. The algorithm stops only when no improvement can be obtained by any of the two phases, having a partition with the highest achieved

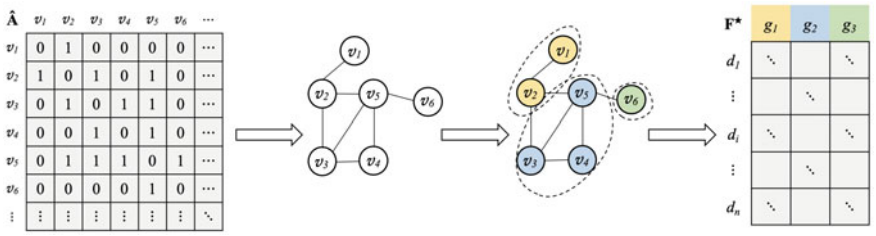


Fig. 2 Network-based dimensionality reduction

modularity. Differently from other graph partitioning techniques, a modularity-based community detection does not require any prior decision on K, performing an automatic procedure.

3.2 The Proposed Strategy

In this paper, we propose the use of the multilevel algorithm as a tool to reduce the high-dimensionality of term space. After transforming the lexical table $\mathbf{F}$ into an adjacency matrix $\hat{\mathbf{A}}$, the community detection is performed obtaining from the set of p terms a lower set of K communities. Differently from the constructs obtained with feature extraction methods, communities are easier to interpret. Each community of term can be seen as a distinct concept enclosed in the corpus with a core semantic meaning. At the end of the process, it is possible to build from the p-dimensional community membership vector $\mathbf{c}$ a complete disjunctive $terms \times communities$ matrix $\mathbf{C}$, whose generic element c_{jk} assumes 1 or 0 wether the term j belongs or not belongs to the community g_k, respectively. The result of the dimensionality reduction process can be represented as a $texts \times communities$ matrix $\mathbf{F}^{\star} \equiv (\mathbf{F}^{\mathsf{T}}\mathbf{C})\mathbf{D}_K{}^{-1}$, where $\mathbf{D}_K{}^{-1}$ is the diagonal matrix obtained from the column marginal distribution of $\mathbf{C}$. Each cell of $\mathbf{F}^{\star}$ contains the proportion of terms belonging to each community detected across the different texts of the corpus. Figure 2 schematically shows the logic of the proposed network-based dimensionality reduction. Starting from these peculiar lower-dimensional data structures, the matrices $\mathbf{C}$ and $\mathbf{F}^{\star}$, it is possible to carry on further analyses both on term and text side [e.g. 30].

4 A Case Study on COVID-19 Tweets

Starting from December 2019, a new infectious disease spread all over the world. The *severe acute respiratory syndrome coronavirus 2* (SARS-CoV-2)—commonly known as COVID-19—appeared in the Chinese mainland and in few weeks became a global pandemic. Italy was directly interested for the first time at the end of January

2020, but the state of emergency caused by COVID-19 started at the end of February 2020, when a cluster of cases was detected in Lombardy, a region in the North of Italy. Due to the rapid spread of the disease and the occurrence of new clusters in different areas, the government decided to take drastic measures putting in lockdown the Northern regions interested by the first clusters on 1st March and the whole country on 9th March, placing more than 60 million people in quarantine. The exceptional nature of this measure had a great impact on everyday life, because of the stop on all non-essential activities and the restrictions on travels across Italy and abroad. The widespread use of social networks like Facebook and Twitter allowed a rapid circulation of news more than traditional media but also an improvement of misinformation, triggering an over-production of comments and opinions and causing an actual *infodemic*. Here in the following, an application of the proposed strategy on a set of short texts related to COVID-19 is presented, showing the advantages of using a network-based dimensionality reduction.

We considered the 40wita dataset [6], containing more than 2 million tweets written in Italian about the COVID-19 outbreak. The dataset is part of a wider project developed by some researchers of the University of Turin, aiming at building a massive collection of tweets in Italian.[1] Tweets were selected from the main set by the 1st March 2020 using a list of 43 different keywords that includes both trivial terms and hashtags related to COVID-19 (e.g. *covid19*, *coronavirus*) and others became popular in Italy during the emergency (e.g. *#iorestoacasa*, "I stay at home", or *#andratuttobene*, "everything will be fine"). We decided to focus only on the 100,928 tweets published between 9th and 10th March, the two days in which the Italian government started the lockdown and the shutdown of the whole country. Data pre-processing was performed in two steps. Firstly, we stripped URLs, usernames, hashtags, emoticons, normalising the tweets by removing special characters and any delimiter different from blank. Secondly, we performed on cleaned tweets a POS tagging and a lemmatisation. We decided to retain only nouns and adjectives for the analysis. Nouns can be seen as something or someone an individual can speak about or think of. Adjectives represent the characteristics that qualify an entity represented by a noun. Because of the peculiar structure of short texts as the tweets, terms with a different POS tag have been discarded since they are less efficient in capturing the essence of a topic. Moreover, we firstly selected terms occurring in at least two tweets and with more than 3 characters from the vocabulary (since in Italian there are few significant nouns and adjectives with a length of two or three characters), then we deleted terms occurring less than 5 times. The latter threshold has been empirically chosen following the common practice in text analyses. At the end of the pre-treatment process, we obtained a $terms \times documents$ table $\mathbf{F}$ with 9,296 rows and 99,911 columns, and consequently a $terms \times terms$ co-occurrence table $\hat{\mathbf{A}}$ with 9,296 rows and columns.

[1] http://twita.di.unito.it.

Table 1 Network metrics for different co-occurrence thresholds

τ	% of terms	ADARP	Avg. degree	CC	Density	Diameter
1	99.49	2.185	72.856	0.176	0.008	5
2	90.31	2.296	44.607	0.170	0.005	5
3	77.76	2.308	35.314	0.164	0.005	6
4	65.99	2.283	30.761	0.161	0.005	6
5	57.57	2.272	27.796	0.159	0.005	7
6	50.16	2.253	26.038	0.157	0.006	7
7	45.07	2.232	24.453	0.154	0.006	5
8	41.20	2223	22.909	0.151	0.006	6
9	38.03	2.211	21.684	0.149	0.006	6
10	35.38	2.197	20.585	0.146	0.006	5
11	33.00	2.186	19.659	0.143	0.006	5
12	31.14	2.188	18.781	0.141	0.006	5

4.1 Dimensionality Reduction and Concepts' Identification

We performed the community detection procedure on $\hat{\mathbf{A}}$ to identify the different concepts embodied in the dataset. In order to prune the network from isolated and peripheral terms, a threshold τ on term-term co-occurrences was set. This threshold allows considering the network core, i.e. the denser and more significant part of the collection vocabulary, and at the same time speed up the detection process because remove rare collocations. Table 1 shows how network structure changes with different threshold values. Following [18], we calculated several metrics to evaluate the behaviour of the whole network: *average distance among reachable pairs* (ADARP), *average degree*, *clustering coefficient* (CC), *density* and *diameter*.

As we can see, filtering and deleting isolated and peripheral terms do not have a substantial impact on the network structure. At different threshold values, the diameter and the density remained quite stable and the other metrics showed small variations. We decided to consider $\tau = 6$, retaining the 50.16% of the terms listed in $\hat{\mathbf{A}}$. By applying the *Louvain* algorithm on this sub-network, we detected 219 different concepts. As stated above, this term network's partition is automatically determined without any prior decision on the K parameter. The modularity value obtained in the community detection ($Q = 0.541$) supported the effectiveness of the strategy.

Figure 3 shows in different colours the communities/concepts detected in the network (some colours were reused due to the large number of communities). It is possible to see that some concepts are described by a huge number of terms (e.g. the community depicted in light blue in the centre of the network and the community depicted in yellow on the left-top side of the network). Many other concepts are described by few terms (e.g. peripheral communities around the larger commu-

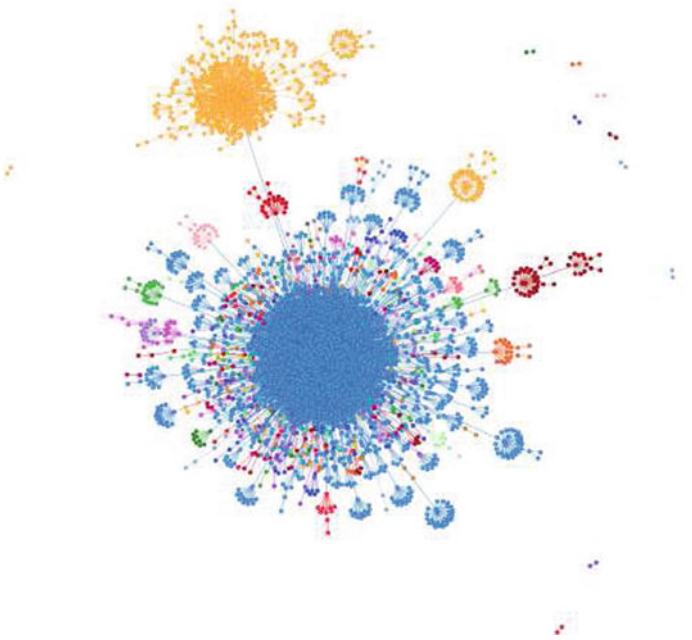


Fig. 3 Communities/concepts detected on the term network

Table 2 First eight concepts and corresponding terms

Concept	Nodes	Terms
C1	569	*quarantena, sforzo, esempio, riposo, film, Netflix, playstation, sacrificio, indispensabile, cucina, pigiama, smartworking, cena, tutorial, …*
C2	10	*motivo, comprovato, necessitá, salute, datore, lavoro, futile, …*
C3	17	*tuta, protettivo, mascherina, guanto, lattice, introvabile, igienizzante, monouso, sciarpa, …*
C4	20	*virus, espansione, letale, fobia, ignoranza, selezione, naturale, antidoto, …*
C6	21	*medico, infermiere, eroe, guardia, notturno, specializzando, stremato, paramedico, …*
C7	20	*scorta, latte, fila, coda, supermercato, notturno, assalto, ammassare, chilometrico, code, carestia, …*
C8	7	*cancellare, volo, malpensa, ryanair, british, airways, easyjet, …*

nity of the network), whereas some of them contain only two terms (e.g. isolated communities on the right-top side of the network).

To better highlight the detected concepts, an example of the terms enclosed in the first eight communities (considering the order number of the communities obtained by the algorithm) is showed in the table below (Table 2).

It is interesting to note that the detected concepts relate to different aspects of the Italian lockdown. The concept **C1** contains terms mainly referred to home activities that the Italians started with the quarantine, and to how important is staying at home even though it is a great sacrifice. **C2** refers to the reasons of necessity for leaving home during the quarantine: health causes, work, or other demonstrable needs (i.e. purchase of essential goods). The concept **C3** concerns the lack of medical equipments and individual protection devices (DPI) at the beginning of the lockdown. Concepts **C4** and **C7** are mainly related to people's concerns with respect to the epidemic on public health and at the same time with respect to the effects of the epidemic on the productive and economic system (i.e. fear of food shortage). As said above, together with these primary concepts, it was possible to detect some other simple concepts composed of only two terms. In this case, the strategy helped identifying proper names of public figures (i.e. *Vittorio Feltri*, *Alberto Angela*, *Bebe Vio*), place names (i.e. *Cologno Monzese*, *Hong Kong*, *Regno Unito*), collocations/multiwords (i.e. *asilo nido*, *carro armato*, *talk show*, *salto mortale*), not detected during the text preprocessing and clearly useful for subsequent analyses.

5 Conclusions and Final Remarks

In this paper, we proposed and discussed an alternative strategy based on networks to cope the problem of high dimensionality in textual datasets. Differently from other approaches operating on $documents \times terms$ matrices, the network-based strategy allows extracting high-level structures from a collection of documents by detecting communities of terms. Each community can be seen as a new feature that retains the meaning of the single terms, and it can be seen as a relevant concept/topic for the domains referring to the collection. The strategy is suitable when we deal with short texts, as in the case study on COVID-19, but it can also be applied to other kind of longer documents. One of the primary advantages of this approach, compared with other proposals discussed in the literature (see Sect. 2), is that dimensionality is reduced by considering the co-occurrence of terms, taking into account the association relationships and hence saving the context in which terms are used. As a consequence, the original meaning of the features is saved in each step of the analysis, making easy for the researchers to understand how the approach reduce dimensionality and interpret the results. Moreover, differently from other approaches like topic modelling, the number of topics is automatically determined by the algorithm without using any prior information on the texts, and it is not necessary to set up different parameters related to the distributions used in the analysis. Some evidence of the latter claim can be found in [30]. The reduction operated by the network-based strategy can also be seen as a first stage of more complex analytic processes. Since terms are linked each others due to the co-occurrence structure, it is possible to detect both complex structures and simple relations, like collocations and multiwords.

Several improvement of the proposal here presented can be considered. Future developments will be devoted to automatically set the co-occurrence threshold in

the community detection step, in-depth evaluating the co-occurrence distribution and its behaviour. Moreover, alternative similarity indices for measuring the relation strength among terms will be evaluated, trying to enclose additional information on the terms and produce more informative structures for the analytic steps.

Acknowledgements The research presented in this work was financially supported by University of Calabria (Fondo UNICAL ex60%—2018).

References

1. Aggarwal, C.C., Hinneburg, A., Keim, D.A.: On the surprising behavior of distance metrics in high dimensional space. In: Van den Bussche, J., Vianu, V. (eds.) Database Theory – ICDT 2001, pp. 420–434 (2001)
2. Allahyari, M., Pouriyeh, S., Kochut, K., Arabnia, H.R.: A knowledge based topic modeling approach for automatic topic labelin. Int. J. Adv. Comput. Sci. Appl. **8**(9), 335–349 (2017)
3. Balbi, S.: Beyond the curse of multidimensionality: high dimensional clustering in text mining. Ital. J. Appl. Stat. **22**(1), 53–63 (2010)
4. Balbi, S., Misuraca, M.: Procrustes techniques for text mining. In: Zani, S., Cerioli, A., Riani, M., Vichi, M. (eds.) Data Analysis, Classification and the Forward Search, pp. 227–234. Springer (2006)
5. Barman, P.C., Iqbal, N., Lee, S.Y.: Non-negative matrix factorization based text mining: feature extraction and classification. In: King, I., Wang, J., Chan, L.W., Wang, D. (eds.) Neural Information Processing, pp. 703–712. Springer (2006)
6. Basile, V., Caselli, T.: 40twita 1.0: a collection of italian tweets during the COVID-19 pandemic (2020). http://twita.di.unito.it/dataset/40wita
7. Bharti, K.K., Singh, P.K.: A survey on filter techniques for feature selection in text mining. In: Babu, B.V., Nagar, A.K., Deep, K., Pant, M., Bansal, J.C., Ray, K., Gupta, U. (eds.) Proceedings of the 2nd International Conference on Soft Computing for Problem Solving, pp. 1545–1559 (2014)
8. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
9. Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. J. Stat. Mech. **5**(10) (2008)
10. Callon, M., Courtial, J.P., Laville, F.: Co-word analysis as a tool for describing the network of interactions between basic and technological research: the case of polymer chemistry. Scientometrics **22**(1), 155–205 (1991)
11. Choulakian, V., Kasparian, S., Miyake, M., Akama, H., Makoshi, N., Nakagawa, M.: A statistical analysis of synoptic gospels. In: Viprey, J.R. (ed.) Proceedings of the 8th International Conference on Textual Data, pp. 281–288. Presses Universitaires de Franche-Comté (2006)
12. Crain, S.P., Zhou, K., Yang, S.H., Zha, H.: Dimensionality reduction and topic modeling: from latent semantic indexing to latent Dirichlet allocation and beyond. In: Aggarwal, C.C., Zhai, C. (eds.) Mining Text Data, pp. 129–161. Springer (2012)
13. Dash, M., Liu, H.: Feature selection for clustering. In: Terano, T., Liu, H., Chen, A.L.P. (eds.) Proceedings of the 4th Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 110–121. Springer (2000)
14. Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R.: Indexing by latent semantic analysis. J. Amer. Soc. Inf. Sci. **41**(6), 391–407 (1990)
15. Deng, Z.H., Tang, S.W., Yang, D.Q., Zhang, M., Li, L.Y., Xie, K.Q.: A comparative study on feature weight in text categorization. In: Yu, J.X., Lin, X., Lu, H., Zhang, Y. (eds.) Advanced Web Technologies and Applications, pp. 588–597. Springer (2004)

16. Dhillon, I., Kogan, J., Nicholas, C.: Feature selection and document clustering. In: Berry, M.W. (ed.) Survey of Text Mining. Clustering, Classification and Retrieval, pp. 73–100. Springer (2004)
17. Fortunato, S.: Community detection in graphs. Phys. Rep. **486**, 75–174 (2010)
18. Fronzetti Colladon, A., Gloor, P.: Measuring the impact of spammers on e-mail and twitter networks. Int. J. Inf. Manage. **48**, 254–262 (2019)
19. Greenacre, M.: Theory & applications of Correspondence Analysis. Academic, London (1983)
20. Hofmann, T.: Probabilistic latent semantic indexing. In: Gey, F., Hearst, M., Tong, R. (eds.) Proceedings of the 22nd Annals International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 50–57. ACM (1999)
21. Hofmann, T.: Unsupervised learning by probabilistic latent semantic analysis. Mach. Learn. **42**(1–2), 177–196 (2001)
22. Lau, J., Grieser, K., Newman, D., Baldwin, T.: Automatic labelling of topic models. In: Lin, D., Matsumoto, Y., Mihalcea, R. (eds.) Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, vol. 1, pp. 1536–1545. ACL (2011)
23. Lebart, L.: Contiguity analysis and classification. In: Gaul, W., Opitz, O., Schader, M. (eds.) Data Analysis. Scientific Modeling and Practical Application, pp. 233–244. Springer (2000)
24. Lebart, L., Salem, A., Berry, L.: Exploring Textual Data. Kluwer, Dordrecht (1988)
25. Lee, J.A., Verleysen, M. (eds.): Nonlinear Dimensionality Reduction. Springer, New York (2007)
26. Liu, H., Motoda, H. (eds.): Feature Extraction, Construction and Selection. Springer, New York (1998)
27. Liu, L., Kang, J., Yu, J., Wang, Z.: A comparative study on unsupervised feature selection methods for text clustering. In: Proceedings of the 2005 IEEE International Conference on Natural Language Processing and Knowledge Engineering, pp. 597–601. IEEE (2005)
28. Liu, T., Liu, S., Chen, Z., Ma, W.: An evaluation on feature selection for text clustering. In: Fawcett, T., Mishra, N. (eds.) Proceedings 20th International Conference on Machine Learning, pp. 488–495. AAAI (2003)
29. Misuraca, M., Scepi, G., Grassia, M.G.: Extracting and classifying keywords in textual data analysis. Ital. J. Appl. Stat. **17**(4), 517–528 (2005)
30. Misuraca, M., Scepi, G., Spano, M.: A network-based concept extraction for managing customer requests in a social media care context. Int. J. Inf. Manage. **51**(101956) (2020)
31. Misuraca, M., Spano, M.: Unsupervised analytic strategies to explore large document collections. In: Iezzi, D.F., Mayaffre, D., Misuraca, M. (eds.) Text Analytics. Advances and Challenges, pp. 17–28. Springer (2020)
32. Newman, M.E.J.: Networks: An Introduction. Oxford University Press, Oxford (2010)
33. Newman, M.E.J., Girvan, M.: Finding and evaluating community structure in networks. Phys. Rev. E **69**(026113) (2004)
34. Paatero, P., Tapper, U.: Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. Environmetrics **5**, 111–126 (1994)
35. Pauca, V.P., Shahnaz, f., Berry, M.W., Piemmons, R.J.: Text mining using non-negative matrix factorizations. In: Berry, M.W., Dayal, U., Kamath, C., Skillikorn, D. (eds.) Proceedings of the 2004 SIAM International Conference on Data Mining, pp. 452–456. SIAM (2004)
36. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. Science **290**, 2323–2326 (2000)
37. Salton, G., Buckley, C.: Term-weighting approaches in automatic text retrieval. Inf. Process. Manag. **24**(5), 513–523 (1988)
38. Salton, G., Wong, A., Yang, C.S.: A vector space model for automatic indexing. Commun. ACM **18**(11), 613–620 (1975)
39. Spärck Jones, K.: A statistical interpretation of term specificity and its application in retrieval. J. Doc. **28**(1), 11–21 (1972)
40. Tang, L., Liu, H.: Graph mining applications to social network analysis. In: Aggarwal, C.C., Wang, H. (eds.) Managing and Mining Graph Data, pp. 487–513. Springer (2010)

41. Tenenbaum, J., de Silva, V., Langford, J.: A global geometric framework for nonlinear dimensionality reduction. Science **290**, 2319–2323 (2000)
42. Valle-Lisboa, J.C., Mizraji, E.: The uncovering of hidden structures by latent semantic analysis. Inf. Sci. **177**(19), 4122–4147 (2007)
43. Wasserman, S., Faust, K.: Social Network Analysis. Cambridge University Press, New York (1994)
44. Wilbur, W.J., Sirotkin, K.: The automatic identification of stop words. J. Inf. Sci. **18**(1), 45–55 (1992)
45. Yang, Y., Pedersen, J.O.: A comparative study on feature selection in text categorization. In: Fisher, D.H. (ed.) Proceedings of the 14th International Conference on Machine Learning, pp. 412–420. ACM (1997)
46. Yang, Z., Algesheimer, R., Tessone, C.J.: A comparative analysis of community detection algorithms on artificial networks. Sci. Rep. **6**(30750) (2016)
47. Zurada, J.M., Ensari, T., Asl, E.H., Chorowski, J.: Nonnegative matrix factorization and its application to pattern analysis and text mining. In: Ganzha, M., Maciaszek, L., Paprzycki, M. (eds.) 2013 Federated Conference on Computer Science and Information Systems, pp. 11–16. IEEE (2013)

Assessing the Performance of the Italian Translations of Modified MEIM, EIS and FESM Scales to Measure Ethnic Identity: A Case Study

Antonino Mario Oliveri and Gabriella Polizzi

Abstract Measuring the ethnic identity of linguistic minorities is a research problem which can be tackled departing from a clear operational definition of the construct. This paper will present the performance of the Italian translations of various scales widely used in the relevant literature, which have been modified for the aims of this study and used in research conducted in 2016 in the Arbereshe Municipalities of Piana degli Albanesi and Santa Cristina Gela (Province of Palermo). These scales consist of modifications of the *Multigroup Ethnic Identity Measure* (MEIM), *Ethnic Identity Scale* (EIS) and *Familial Ethnic Socialization Measure* (FESM). The psychometric properties were analysed for all these scales in terms of reliability and unidimensionality, using Classical Test Theory (CTT) and Item Response Theory (IRT). The latter proved useful in suggesting further improvements regarding the scales.

Keywords Ethnic identity · IRT · Arbereshe · MEIM · EIS · FESM

1 Ethnic Identity: A Conceptualization

At a time when Western societies are affected by massive migratory flows comprising subjects with ethnic identities which integrate with varying degrees of difficulty into host societies, the issue of ethnic identity is particularly relevant: on the one hand, it constitutes a protective factor against disadvantage and maladjustment, and, on the other, a possible obstacle to integration into the host society. As demonstrated by Berry [2], the outcomes of the encounter between the culture of origin and the

A. M. Oliveri (✉)
Dipartimento Culture E Società, Università Degli Studi Di Palermo, Palermo, Italy
e-mail: antoninomario.oliveri@unipa.it

Viale Delle Scienze, Edificio 15, 90128 Palermo, Italy

G. Polizzi
Facoltà Di Scienze Dell'Uomo E Della Società, Università Degli Studi Di Enna "Kore", Enna, Italy

Cittadella Universitaria, Via Salvatore Mazza N. 1, 94100 Enna, Italy

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_13

host culture can be multiple and they may take four different forms, namely, *assimilation*, *integration*, *separation* and *marginalization.* The category of *assimilation* is used when people leave their former culture due to interaction with other cultures. *Integration* refers to the possible maintenance of one's own culture of origin, even in the presence of interaction with the host culture. *Separation* consists in maintaining one's own culture of origin, whilst not being willing to interact with other cultures. Finally, *marginalization* occurs in cases when the subject has little interest in their culture and interaction with other cultures.

Whereas *ethnicity* has been considered an *ascribed status*, i.e. a permanent characteristic linked to the culture of the country of origin, *ethnic identity* has been conceptualized as the way "individuals perceive themselves within an environment as they categorize and compare themselves to others of the same or a different ethnicity. It is the closeness or distance one feels from one's own ethnicity or from other ethnicities, as one tries to fit into the society. As such, it can differ among migrants of the same origin, or be comparable among migrants of different ethnic backgrounds" [5], p. 4).

Ethnic groups are conceived as communities identifying themselves in terms of similarities and differences with respect to a series of traits or markers [3]. Schermerhorn ([14], p. 12) has defined an ethnic group as "a collectivity within a larger society having real or putative common ancestry, memories of a shared historical past, and a cultural focus on one or more symbolic elements defined as the epitome of their peoplehood. Examples of such symbolic elements are: kinship patterns, physical contiguity (as in localism or sectionalism), religious affiliation, language or dialect forms, tribal affiliation, nationality, phenotypical features, or any combination of these. A necessary accompaniment is some consciousness of kind among members of the group". And according to Phinney et al. ([12], p. 169), ethnic identity consists of "a feeling of belonging to one's group, a clear understanding of the meaning of one's membership, positive attitudes towards the group, familiarity with its history and culture, and involvement in its practices".

Phinney and Ong [11] have pointed out that two constructs have been identified in the literature as key components of ethnic identity, i.e. *exploration* and *commitment*; both play an important role in the dynamic process of ethnic identity formation. *Exploration* consists in seeking information and experiences which are capable of affecting ethnic identity formation, such as "reading and talking to people, learning cultural practices, and attending cultural events" [11], p. 272). *Exploration* may continue throughout one's life to provide knowledge and understanding about one's ethnic sense of belonging, to which the construct of *commitment* is referred and conceptualised in terms of a person's strong attachment and investment in a group [13]. Specifically, the deeper the exploration process, the more stable the commitment may be since the latter is based on a more stable sense of self (developed by virtue of the exploration itself) and a subsequently achieved identity.

The two constructs of *exploration* and *commitment* lie at the base of *Multigroup Ethnic Identity Measure-Revised* (MEIM-R), as proposed by Phinney and Ong [11]. The aim of MEIM-R is to improve the original version of MEIM [10], which considered exploration and commitment (i.e. individuals' achievement), ethnic behaviors (i.e. degree of participation in cultural activities), and affirmation and belonging

(i.e. degree of positive feelings toward their ethnic group) as the points of departure for assessing individuals' degree of *ethnic identity achievement* on a 14-item single scale, rather than considering ethnic identity as a multidimensional construct.

Slightly different conceptualizations of ethnic identity affect the ways by which it is measured. An example can be found in the *Ethnic Identity Scale* (EIS) by Umaña-Taylor et al. [17]. This is based on three factors believed to influence the process of ethnic identity formation, namely: *exploration*, *resolution*, and *affirmation*. *Exploration* is the same as Phinney and Ong's concept. *Resolution* means the degree to which individuals have solved what their ethnic identity means to them. The *resolution* construct is based on the Erikson's idea that "the culmination of such a period of exploration will lead the individual to 'reconcile his conception of himself and his community's recognition of him' ([6], p. 120)" ([17], p. 11). *Affirmation* is the positive or negative affect which individuals associate with that resolution. Thus, whereas *resolution* is the awareness of belonging to a certain group, *affirmation* refers to the degree to which individuals feel positively or negatively integrated into the host society.

Individual ethnic identity has been conceptualized as being strongly influenced by the construct of *family ethnic socialization*. This is the theoretical base of the *Familial Ethnic Socialization Measure* (FESM), originally proposed by Umaña-Taylor [15], subsequently revised by Umaña-Taylor et al. [17]. There will be further discussion regarding FESM below.

2 Measuring Ethnic Identity: A Case-Study

With the aim of measuring the ethnic identity of the Arbereshe populations living in Sicily, a questionnaire was constructed and administered to the families of the students of the "Skanderbeg" comprehensive school in the Municipalities of Piana degli Albanesi and Santa Cristina Gela in the Province of Palermo between April and May 2016. The Arbereshe communities are composed of descendants of the historical Albanians who settled in Sicily in the fifteenth century to escape the Ottoman domination in the Balkans. As long-term immigrants, the Sicilian Arbereshe people constitute a case study of special interest regarding the possible persistence of an ethnic identity hundreds of years after their successful migration.

The research focus was twofold: one strand consisted in measuring ethnic identity with reference to all the relevant theoretical dimensions of the concept; another strand analysed the psychometric characteristics of the scales included in the questionnaire, which derive in part from various widely-used measurement instruments, namely MEIM [10], EIS [17] and FESM [16]. This paper will report results relating to the latter research objective.

The MEIM scale originally consisted of twenty items, including six items to measure orientation towards other ethnic groups. However, it was amended and a reduced version was proposed by Roberts et al. [13], who estimated a two-factor

structure (factor 1: *affirmation, belonging and commitment*; factor 2: *ethnic exploration and behaviors*); items 13 and 14, which were meant as replicating with reverse polarity items 1 and 3, were eliminated. The scale is reported in Table 1.

Starting with the revised version by Roberts et al. [13], the authors of this paper noticed an overlap among several items: item 3 "I have a clear sense of my ethnic background and what it means for me" and item 7 "I understand pretty well what my ethnic group membership means to me"; item 6 "I have a strong sense of belonging to my own ethnic group" and item 11 "I feel a strong attachment towards my own ethnic group"; item 5 "I am happy that I am a member of the group I belong to" and item 12 "I feel good about my cultural or ethnic background".

These overlapping items were therefore merged and replaced by new formulations. The adapted and revised MEIM scale is reported in Table 1.

The original item 4 "I think a lot about how my life will be affected by my ethnic group membership" was replaced by "Have Arbereshe culture and lifestyle had a positive influence on your life?" due to the fact that this new formulation was functional to comparing the Arbereshe culture with concurrent/coexisting cultures, such as Sicilian, Albanian, and Italian, which were capable of generating multiple identities.

An objective of this research was, therefore, to test the psychometric characteristics of the reduced 9-item MEIM scale. As previously stated, efforts have been successfully made in this same direction by Phinney and Ong [11], who introduced a revised version of the MEIM scale, named MEIM-R, which consists of just six items measuring the two factors of exploration and commitment.

The fact that in this study the questionnaire was administered in Italian means that the properties of the modified scales and their translations into Italian were tested simultaneously.

The EIS scale by Umaña-Taylor et al. [17] comprises seventeen items. The scale was developed by being based on factor analysis and residual analysis in order to maximize performance. However, the EIS scale, which is reported in Table 2, seems to be affected by item semantic overlapping even more than the MEIM. The items that seem to overlap more conspicuously are items 1, 7, 9, 16; items 7 and 13; items 5, 8 and 15; items 6 and 11; items 3, 12, 14 and 17. As a consequence, it was decided to use a reduced scale, thereby merging the overlapping items. The output consisted of a much simpler 7-item EIS scale, which is reported in Table 2.

Various items on the EIS scale show the same information content as some items of the MEIM scale. For example, EIS items 3, 12, 14 and 17 share the same information with MEIM items 3 and 7. In other words, in this study some items were included in the revised 9-item MEIM scale and the revised 7-item EIS scale. The new formulations partially modified those reported in the original scales.

Ethnic identity develops in individuals over years, starting from the first socialization experiences in childhood, passing into adolescence and even into adulthood [10]. The formation of ethnic identity depends on various factors of: an individual nature (such as the need to belong), a relational nature (such as the relationships which are established in families, with friends or at school), and a social nature (such as stereotypes and possible discrimination to which one can be subjected). However,

Table 1 12-Item MEIM scale and 9-Item adapted and revised MEIM scale: lists of items

Original 12-Item MEIM scale		9-Item adapted and revised MEIM scale	
Item number	Item text	Item number	Item text
Item 1	I have spent time to find out more about my ethnic group, such as its history, traditions, and customs	Item 1 (MEIM 1)	I have spent time to find out more about the Arbereshe ethnicity, such as its history, traditions, and customs
Item 2	I am active in organizations or social groups that include mostly members of my own ethnic group	Item 2 (MEIM 2)	Are you active in organizations or social groups that include mostly Arbereshe people?
Item 3	I have a clear sense of my ethnic background and what it means for me	Item 3 (MEIM 3, MEIM 7)	I have a clear sense of my Arbereshe ethnicity and what it means to me
Item 4	I think a lot how my life will be affected by my ethnic group membership	Item 4 (MEIM 4)	Have Arbereshe culture and lifestyle had a positive influence on your life?
Item 5	I am happy that I am a member of the group I belong to	Item 5 (MEIM 5, MEIM 12)	I am happy that I am Arbereshe
Item 6	I have a strong sense of belonging to my own ethnic group	Item 6 (MEIM 6, MEIM 11)	I have a strong sense of belonging to the Arbereshe ethnic group
Item 7	I understand pretty well what my ethnic group membership means to me		
Item 8	To learn more about my ethnic background, I have often talked to other people about my ethnic group	Item 7 (MEIM 8)	To learn more about my ethnic background, I have often talked to other people about my Arbereshe ethnicity
Item 9	I have a lot of pride in my ethnic group and its accomplishments	Item 8 (MEIM 9)	I am proud to be Arbereshe
Item 10	I participate in cultural practices of my own group, such as special food, music, or customs	Item 9 (MEIM 10)	I have experienced things which reflect the Arbereshe ethnicity, such as eating food, listening to music, and watching films
Item 11	I feel a strong attachment towards my own ethnic group		

(continued)

Table 1 (continued)

Original 12-Item MEIM scale		9-Item adapted and revised MEIM scale	
Item number	Item text	Item number	Item text
Item 12	I feel good about my cultural or ethnic background		

Source Roberts et al. [13], p. 319

the role of family assumes a very important role in forming ethnic identity. The stronger that parents feel they belong to a group, the more likely they will try to socialize their children as well into this group [4].

The Familial Ethnic Socialization Measure (FESM) was first introduced by Umaña-Taylor [15] and later revised in 2004 [17]. It consists of 12 polytomic items and it is reported in Table 3.

The FESM scale was also modified in the study presented in this paper. Questions were administered to young people, to adults and to elderly people. However, all of them were asked twice: the first time regarded what was/had been the socialization that they were receiving/had experienced as children; the second time concerned the socialization that they were imparting to their children in their role of parents (where the respondents were parents). The FESM scale was therefore duplicated into a FESM-1 scale (socialization experienced as a child) and a FESM-2 scale (socialization given as a parent). Thus, it was thought it would be possible to investigate intergenerational ethnic socialization: whether there existed significant relationships between that which people had learned from parents as a child and that which they were transmitting to their children as parents.

The scores of FESM-1 and FESM-2 items were limited in the range 1-4, being reversed between the two scales in order to avoid a memory effect in the responses to the scale which had been subsequently administered (in the role of parents).

3 Results

Having administered the questionnaires, 207 were returned although only 178 were included in this study; the latter were based on the eligibility criteria of possessing Italian citizenship and a self-assessed ability to speak and understand Arberisht, that is, the language of the Arbereshe people. An analysis of the psychometric characteristics of the four scales, which were constructed as previously described, elicited results of note.

Table 2 17-Item EIS scale and 7-Item adapted and revised EIS scale: lists of items

Original 17-Item EIS scale		7-Item adapted and revised EIS scale	
Item number	Item text	Item number	Item text
Item 1	My feelings about my ethnicity are mostly negative	Item 1 (EIS1, EIS 7, EIS 9, EIS 16)	Are your feelings about the Arbereshe ethnicity mostly negative?
Item 2	I have not participated in any activities that would teach me about my ethnicity	Item 2 (EIS 2, EIS 15)	I have not participated in any activities that would teach me about my ethnicity
Item 3	I am clear about what my ethnicity means to me	Item 3 (EIS 3, EIS 12, EIS 14, EIS 17)	I have a clear sense of my Arbereshe ethnicity and what it means to me
Item 4	I have experienced things that reflect my ethnicity, such as eating food, listening to music, and watching movies	Item 4 (EIS 4)	I have experienced things which reflect the Arbereshe ethnicity, such as eating food, listening to music, and watching films
Item 5	I have attended events that have helped me learn more about my ethnicity	Item 5 (EIS 5, EIS 8, EIS 15)	I have attended events that have helped me learn more about the Arbereshe ethnicity
Item 6	I have read books/magazines/newspapers or other materials that have taught me about my ethnicity	Item 6 (EIS 6, EIS 11)	I have learned something about the Arbereshe ethnicity by doing things such as reading books/magazines/newspapers, searching over the Internet or keeping up to date with current events
Item 7	I feel negatively about my ethnicity		
Item 8	I have participated in activities that have exposed me to my ethnicity		
Item 9	I wish I were of a different ethnicity		
Item 10	I am not happy with my ethnicity	Item 7 (EIS 10, EIS 13)	Would you like to be of an ethnicity other than the Arbereshe?
Item 11	I have learned about my ethnicity by doing things such as reading (books, magazines, newspapers), searching the internet, or keeping up with current events		

(continued)

Table 2 (continued)

Original 17-Item EIS scale		7-Item adapted and revised EIS scale	
Item number	Item text	Item number	Item text
Item 12	I understand how I feel about my ethnicity		
Item 13	If I could choose, I would prefer to be of a different ethnicity		
Item 14	I know what my ethnicity means to me		
Item 15	I have participated in activities that have taught me about my ethnicity		
Item 16	I dislike my ethnicity		
Item 17	I have a clear sense of what my ethnicity means to me		

Source Umaña-Taylor et al. [17], p. 38

Table 3 FESM scale: list of items

Item number	Item text
Item 1	My family teaches me about my ethnic cultural background
Item 2	My family encourages me to respect the cultural values and beliefs of our ethnic/cultural background
Item 3	My family participates in activities that are specific to my ethnic group
Item 4	Our home is decorated with things that reflect my ethnic/cultural background
Item 5	The people who my family hangs out with the most are people who share the same ethnic background as my family
Item 6	My family teaches me about the values and beliefs of our ethnic/cultural background
Item 7	My family talks about how important is to know about my ethnic/cultural background
Item 8	My family celebrates holidays that are specific to my ethnic/cultural background
Item 9	My family teaches me about the history of my ethnic/cultural background
Item 10	My family listens to music sung by artists from my ethnic/cultural background
Item 11	My family attends things such as concerts, plays, festivals, or other events that represent my ethnic/cultural background
Item 12	My family feels a strong attachment to our ethnic/cultural background

Source https://umana-taylorlab.gse.harvard.edu/familial-ethnic-socialization-measure

3.1 The 9-Item MEIM Scale

The "classical" analysis of tests (Classical Test Theory—CTT) [7] was carried out on the 9-item MEIM scale although only seven items were included after the exclusion of items 4 and 2, which assumed a dichotomous form in the questionnaire, unlike the other 4-category items. The scale performed well in terms of reliability, as shown by a value of Cronbach's alpha equal to 0.83. All items correlated satisfyingly with the overall scale scores, with correlations never below 0.50. The authors of this paper held that reliability would not be significantly improved by excluding items.

In addition to CTT, Item Response Analysis was performed, in order to assess the unidimensionality of the scale [8]. Item Response Theory permitted the reinserting of the two dichotomous items MEIM 4 and MEIM 2 into analysis, and an assessment of the performance of the complete 9-item scale. A Partial Credit Model (PCM) [9] was fitted to the data. PCMs model the probability that each individual chooses one response category rather than another for a given item, irrespective of whether such item is dichotomous or polytomic.

The 9-item PCM achieved convergence after 73 iterations, under the constraint of item and person convergence limits which were set at very low values (0.0005 and 0.005 respectively) to enhance estimate precision. The data were fitted in a satisfying manner, with an approximation of the Chi-square statistic equal to 28.9 with 18 degrees of freedom, and p-value of 0.05. The 7-item PCM (which did not include dichotomous items) produced a Chi-square statistic equal to 23.2 with 14 degrees of freedom, and a p-value of 0.06. It was not considered necessary to exclude dichotomous items in order not to reduce the semantic content of the scale.

It is common opinion that the *Person Separation Index* is an effective measure of scale reliability [1]. In the case of the 9-item MEIM, the index assumed a rather high value, equal to 0.865. Scale items performed generally well with three items, however, whose fit was worse than the others: MEIM 3-7-13, MEIM 6-11 and MEIM 9 with respective p-values equal to 0.03; 0.05; 0.03. An inappropriate order of response categories also occurred in the case of MEIM 6-11, as shown in Fig. 1.

Figure 1 shows that there are no regions on the continuum with a probability of selecting category 3 ("slightly disagree"), that is higher than with the other categories.

3.2 The 7-Item EIS Scale

Before analysis, the scores of items 1, 2, and 7 with negative polarity were reversed. The 7-item EIS scale, as modified by the authors of this contribution, revealed a value for Cronbach's alpha index equal to 0.58, which cannot be considered as fully satisfactory. In general, items perform poorly in terms of item-scale correlation (correlations range between 0.03 and 0.56). However, the value of the alpha index increased after the elimination of specific items (EIS 2-15 "I have not participated in any activities that would teach me something about the Arbereshe ethnicity" and

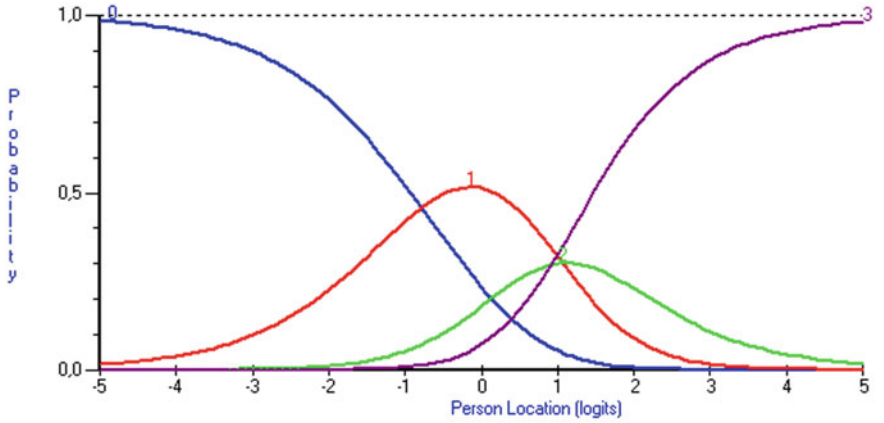


Fig. 1 Category Probability Curves for item MEIM 6-11

EIS 1-7-9-16 "I have a clear sense of my Arbereshe ethnicity and what it means to me").

The IRT analysis produced similar results. A Rating Scale Model (RSM), which was considered appropriate for the case of polytomous items with same number of response categories, was fitted to the data. The model converged in just 11 iterations despite the limits of item and person convergence being the same as for 9-item MEIM analysis. The Person Separation Index was equal to 0.68, which cannot be considered a particularly high value. The model did not fit the data well: the approximation of the Chi-square statistic was equal to 48.84 with 14 degrees of freedom and p-value close to zero. Moreover, the generality of items fitted the data poorly, and the p-value was greater than 0.05 in only four cases. Items EIS 5-8-15, EIS 11-6, and EIS 3-12-14-17 failed to comply with the correct ordering of response categories, as shown in Fig. 2.

The elimination of the worse performing items did not prove capable of raising the model fitting above a Chi-square equal to 28.42 and a p-value equal to 0.005.

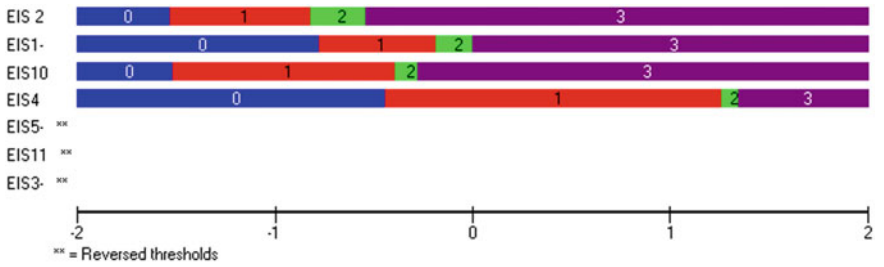


Fig. 2 Threshold map for 7-item EIS items

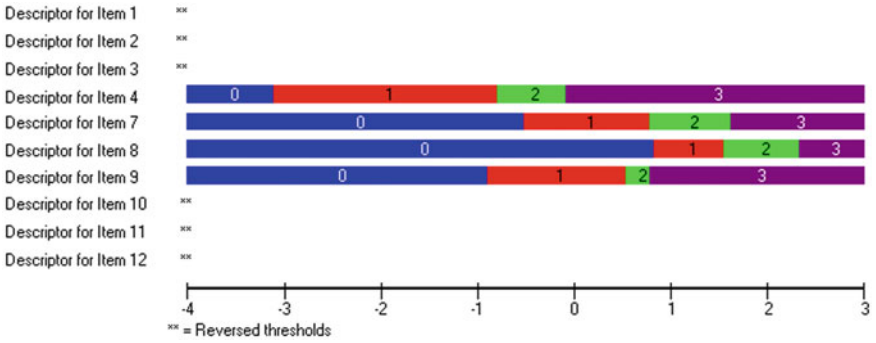


Fig. 3 Threshold map for FESM-1 items

3.3 The FESM-1 Scale

The FESM-1 scale relates to the ethnic socialization experienced as a child. The related Cronbach's alpha index equaled 0.90, and only one item correlated less than 0.5 with the scale: "My family used to attend concerts, plays, festivals, or other events that represented the Arbereshe ethnic background".

The IRT analysis, performed via a Rating Scale Model, provided conflicting results. Although the Person Separation Index was equal to 0.904 (which was consistent with the value of Cronbach's alpha index), the model converged quickly in only 14 iterations but did not fit the data sufficiently well. The Chi-square statistic was equal to 65.41 with 24 degrees of freedom and a p-value close to zero. However, not many items seemed to perform poorly: only item 5 ("The people with whom my family dealt most were people with the same Arbereshe origin") and item 6 ("My family taught me Arbereshe values and beliefs") reported a p-value close to zero. Excluding both items from the scale improved the performance of the model: convergence was obtained in 12 iterations, the Person Separation Index still had a sufficiently high value being equal to 0.887; the Chi-square statistic dropped to 18.54 with 20 degrees of freedom and a p-value of 0.55. Several items were affected by reversed thresholds, as shown in Figs. 3 and 4. For all such items, the third response category ("slightly disagree") was never more likely than the others in any region of the continuum. This seems to point the way to aggregating this response category to others regarding subsequent uses of the instrument.

3.4 The FESM-2 Scale

The FESM-2 scale, concerning socialization being imparted in the role of parents, reported a Cronbach's alpha index equal to 0.92. As in the case of the FESM-1 scale, the FESM-2 only correlated one item with the scale less than 0.5: item 4 "My

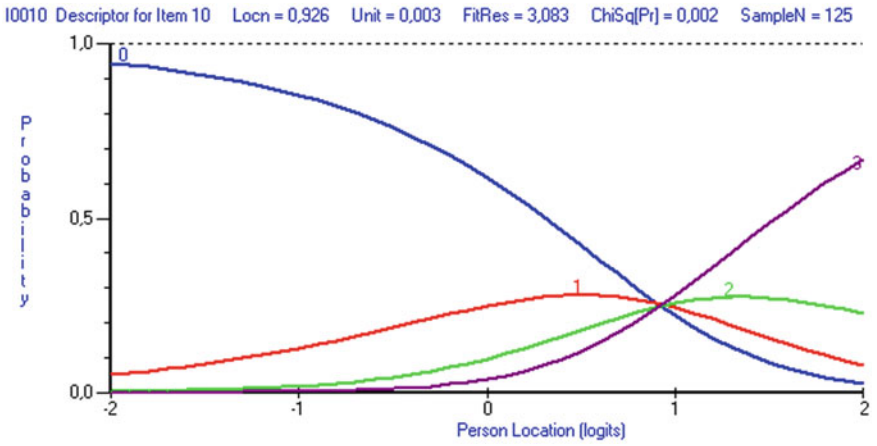


Fig. 4 Category probability curves for FESM-1 item 10

house is furnished with objects that reflect our Arbereshe origin". An IRT analysis was performed over the FESM-2 scale and estimates converged rapidly in just 8 iterations, despite the persistence of the usual restrictive conditions of 0.0005 and 0.005 convergence limits for items and persons. As was the case with FESM-1, FESM-2 also demonstrated a very high Person Separation Index (equal to 0.933) albeit with a poor fitting of the rating model and a p-value practically equal to zero. Regarding the analysis of the performance of the items, it was evident that several fitted the scale poorly, with p-values close to zero. However, only item 10 suffered from reversed thresholds, relating to both the second and the third response categories: "true in part" and "marginally true". This suggests a simplification of the item structure from polytomic to dichotomous in subsequent studies.

The progressive exclusion of the four items whose fit on the scale was the worst (items 4, 10, 5, 11) led to improve goodness of fit, which was considered acceptable with a Chi-square statistic equal to 27.29, 16 degrees of freedom, and a p-value equal to 0.04. The reliability of the scale was still sufficiently high, as is demonstrated by a Person Separation Index equal to 0.935.

Excluding item 5 from the scale meant replicating what had already been done for the FESM-1 scale, and this can be considered indicative that this item is problematic within the FESM. The other items concern "My house is furnished with objects that reflect our Arbereshe origin" (item 4), "In my family we listen to music sung or played by Arbereshe artists" (item 10), "My family attends concerts, plays, festivals or other events representing the Arbereshe ethnic origin" (item 11). Despite these items being clearly related to the notion of ethnic socialization, they can be considered as just potential to the Arbereshe culture, given that they gradually lost any importance to the date of data collection.

4 Conclusions

As is commonly known, social measurement scales perform differently within different populations of interest. And this study supports this idea, given that the use of the generalist scales of measurement of ethnic identity required revision prior to use. Indeed, MEIM, EIS and FESM were subject to rigorous tests in one of the least favourable settings, that is, measuring the ethnic identity of a population consisting of the descendants of people who had settled in Southern Italy five hundred years previously. The Arbereshe people are today fully integrated into Italian society, and it might well be questioned whether an Arbereshe ethnic identity still exists and can be fruitfully detected.

The authors of this study would also contend that its results support the idea that scales need to be adapted to the particular environment in which they are used. Adaptations may well consist in aggregating overlapping questions, as well as reformulating items. Also, translating items into a language other than the original is a well-known technical problem, which, however, all researchers have to deal with.

This study used revised MEIM, EIS and FESM scales. Specifically, the MEIM and the EIS scales were first simplified by merging semantically overlapping items, and then translated into Italian. The FESM scale was instead duplicated into the FESM-1 scale and the FESM-2 scale, due to the fact that interviewees were asked to report on ethnic socialization experienced as children and that imparted as parents. Thus, item formulations were also slightly modified in the case of the FESM-1 and FESM-2 scales.

The new scales led to partially contradictory results. The revised 9-item MEIM scale, FESM-1 and FESM-2 displayed high internal consistency; however, the revised 7-item EIS scale was held to be rather unreliable, failing to reach unidimensionality.

The use of FESM-1 and FESM-2 scales provided results worthy of further investigation. Although the reliability of both scales was very high, the IRT analysis produced issues of unidimensionality. This, therefore, suggested the exclusion of several items which were not considered functional to the adaptation of the probabilistic model. Reducing the number of items of the FESM-1 scale from 12 to 10, as suggested by IRT analysis, does not seem to dramatically affect the structure of the scale; on the other hand, the revised FESM-2 scale requires the inclusion of just 8 out of the original 12 items. A content analysis of the poorly fitting items reveals that they may not assist in detecting how ethnic socialization today occurs among descendants of very long-term migrants.

However, it could be noticed that all these considerations are substantially data driven. This would suggest particular caution in drawing conclusions and, at the very least, that it is advisable to reinvestigate the performance of measurement scales, particularly on other populations of descendants of historical migrants. IRT revealed that item response categories tended to overlap for some items. This suggests that in such cases response categories might be collapsed in order to facilitate the task for interviewees. A feasible solution could consist in using dichotomous

"agree-disagree" item structures rather than polytomic structures, which were used extensively in this study.

References

1. Andrich, D.: An index of person separation in latent trait theory, the traditional KR-20 Index, and the Guttman scale response pattern. Educ. Res. Perspect. **9**(1), 95–104 (1982)
2. Berry, J.W.: Acculturation: living successfully in two cultures. Int. J. Intercult. Rel. **29**, 697–712 (2005)
3. Berthoud, R., Modood, T., Smith, P.: Introduction. In: Modood, T., Berthoud, R., Lakey, J., Nazroo, J., Smith, P., Virdee, S., Beishon, S. (eds.) Ethnic Minorities in Britain Diversity and Disadvantage, pp. 1–18. Policy Studies Institute, London (1997)
4. Boykin, A.W., Toms, F.D.: Black child socialization: a conceptual framework. In: McAdoo, H.P., McAdoo, J.L. (eds.) Black Children: Social, Educational, and Parental Environments, pp. 33–51. Sage, Newbury Park (1985)
5. Constant, A., Zimmermann, K.F.: Measuring ethnic identity and its impact on economic behavior. IZA Discuss. Pap. **3063**, 1–23 (2007)
6. Erikson, E.H.: Identity and the life cycle. Psychol. Issues **1**, 1–171 (1959)
7. Fan, X.: Item response theory and classical test theory: an empirical comparison of their item/person statistics. Educ. Psychol. Meas. **58**(3), 357–381 (1998)
8. Fisher, G.W., Molenaar, I.W.: Rasch models. In: Foundations, Recent Developments, and Applications. Springer, Heidelberg (1995)
9. Masters, G.N.: A Rasch model for partial credit scoring. Psychometrika **47**(2), 149–174 (1982)
10. Phinney, J.S.: The multigroup ethnic identity measure: a new scale for use with diverse groups. J. Adolescent. Res. **7**, 156–176 (1992)
11. Phinney, J.S., Ong, A.D.: Conceptualization and measurement of ethnic identity: current status and future directions. J. Couns. Psychol. **54**(3), 271–281 (2007)
12. Phinney, J.S., DuPont, S., Espinosa, C., Revill, J., Sanders, K.: Ethnic identity and American identification among ethnic minority youths. In: Bouvy, A., van de Vijver, F.J.R., Boski, P., Schmitz, P. (eds.) Journeys Into Cross-Cultural Psychology, pp. 167–183. Swets & Zeitlinger, Berwyn (1994)
13. Roberts, R.E., Phinney, J.S., Masse, L.C., Chen, Y.R., Roberts, C.R., Romero, A.: The structure of ethnic identity of young adolescents from diverse ethnocultural groups. J. Early Adolescence **19**, 301–322 (1999)
14. Schermerhorn, R.A.: Comparative Ethnic Relations: A Framework for Theory and Research. Random House, New York (1970)
15. Umaña-Taylor, A.J.: Ethnic identity development among Mexican-origin Latino adolescents living in the U.S. Unpublished doctoral dissertation, University of Missouri, Columbia (2001)
16. Umaña-Taylor, A.J., Fine, M.A.: Methodological implications of grouping Latino adolescents into one collective ethnic group. Hispanic J. Behav. Sci. **23**, 347–362 (2001)
17. Umaña-Taylor, A.J., Yazedjian, A., Bámaca-Gómez, M.: Developing the ethnic identity scale using Eriksonian and social identity perspectives. Int. J. Th. Res. **4**(1), 9–38 (2004)

Towards Global Monitoring: Equating the Food Insecurity Experience Scale (FIES) and Food Insecurity Scales in Latin America

Federica Onori, Sara Viviani, and Pierpaolo Brutti

Abstract In order to face food insecurity as a global phenomenon, it is essential to rely on measurement tools that guarantee comparability across countries. Although the official indicators adopted by the United Nations in the context of the Sustainable Development Goals (SDGs) and based on the Food Insecurity Experience Scale (FIES) already embeds cross-country comparability, other experiential scales of food insecurity currently employ national thresholds and issues of comparability thus arise. In this work we address comparability of food insecurity experience-based scales by presenting two different studies. The first one involves the FIES and three national scales (ELCSA, EMSA and EBIA) currently included in national surveys in Guatemala, Ecuador, Mexico and Brazil. The second study concerns the adult and children versions of these national scales. Different methods from the equating practice of the educational testing field are explored: classical and based on the Item Response Theory (IRT).

Keywords Food insecurity · Item response theory (IRT) · Test equating · SDGs · Experience-based scales

1 Introduction

Food security is a subject of indisputable relevance, being it conceived as a basic human right since 1948, as stated in Article 25 of the Universal Declaration of Human Rights: "Everyone has the right to a standard of living adequate for the health

F. Onori (✉) · P. Brutti
Sapienza University of Rome, Rome, Italy
e-mail: onori.federica@gmail.com

P. Brutti
e-mail: pierpaolo.brutti@uniroma1.it

S. Viviani
Food and Agriculture Organization of the United Nations, Rome, Italy
e-mail: sara.viviani@fao.org

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_14

and well-being of himself and of his family, including food, clothing, housing and medical care" [2]. However, food security is a complex and multifaceted concept whose terminology has long been affected by a variety of sectors and disciplines strictly related to it (e.g. agriculture, nutrition, economy, public policy, etc.) [7, 20]. A consensus around the definition of food security was finally reached during the World Food Summit in 1996 when it was formalized as follows: "Food security exists when all people, at all times, have physical and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life" [14].[1] Grounding on this definition, the conceptualization and operationalization of food security emerge as that of a multidimensional phenomenon made up of four different, hierarchically ordered dimensions: availability, access, utilization and stability [5]. As a consequence, no single indicator can be successfully designated to return a thorough picture of the phenomenon, but a suite of indicators exists, each monitoring specific aspects of food security at different levels of the observation: national, regional, households and individual [20]. Among all possible aspects related to food insecurity, the dimension of *access to food* is given nowadays high-priority, being acknowledged among the 17 Sustainable Development Goals (SDGs) of the 2030 Agenda for Sustainable Development adopted by the United Nations. Access to food is in fact the subject of Target 2.1 [24], which states:

> By 2030, end hunger and ensure access by all people, in particular the poor and people in vulnerable situations, including infants, to safe, nutritious and sufficient food all year round.

Although food security is now a well-established concept within the scientific community, its definition changed throughout the last century and so did the tools employed to measure the phenomenon [7, 20]. A brief summary of the main steps will enable to fully appreciate the novelties brought about by the measurement tools developed since the '80s. During the 1940s and for some decades on, the issue of food security was completely identified with that of having enough provisions to cover the needs of the population and, therefore the "food problem" was mainly dealt with in terms of country-level supplies [13]. Nevertheless, this formulation could not catch the aspect, yet observable, of malnutrition and famines in countries that did not suffer from food supply at national level [8], signal that a change in approaching the measurement of food insecurity was required, in particular towards considering the point of view of people's *access to food*. To mark this change in prospective, the expression *household food insecurity* began to be used. Since then, other shifts pertained to the definition of food insecurity as for what we use today. A very fundamental one was in the 1990s when interest moved from dietary energy adequacy to *experience of food insecurity* and livelihood conditions, which involved, among others, also social, nutrition and psychological considerations. Food insecurity has in fact been recognized as a "managed process", described by means of a spectrum of behaviours and coping strategies that can reveal the level of severity of a food access condition [25]. Although specific attitudes and coping strategies might

[1] This definition was further refined in 2001 [15], when food access was not only conceived in terms of affordability and physical access, but also in terms of removal of *social barriers*.

change from country to country, there is a general consensus in the scientific community about the common pattern of behaviours that characterize food insecurity with very minor differences across cultures [9]. To this regard, ethnographic and societal studies established that, in case of increasing lack of money or other resources, a common pattern of experiences and behaviours manifests in order to cope with shortage of food [25]: at first, psychological concern arises since people start worrying about having enough food; then, a change in the diet occurs by decreasing the quality and variety of the consumed food in order to face a concrete limited access to food; and, in case of more severe food shortages, people would diminish the quantity of consumed food by reducing meals' size and then by even skipping meals, potentially up to experiencing hunger. The steps just described are commonly referred to as the three *domains* of resource-constrained access to food: psychological concern, decrease of food quality, decrease of food quantity and hunger.[2]

Mirroring these shifts in the paradigm (from global and national to households and individuals; from food supplies to livelihood conditions; and from objective to subjective measures [7]), a number of indicators have been proposed to measure food insecurity, like measures of adequacy of food consumption, prevalence of undernourishment, dietary diversity score, etc. Among all, *experience-based food insecurity scales* found a place of relevance, having proved to be a valid and reliable tool for measuring food insecurity in its access dimension, encompassing the current definition of the phenomenon while adopting a behavioural perspective [7]. As the name suggests, experience-based food insecurity scales measure access to food from a behavioural perspective, building on a set of items that directly ask people about their own personal experiences and behaviours related to the three domains of access to food [20]. The very first experience-based food insecurity scale was the Household Food Security Survey Module (HFSSM), applied yearly in the United States of America since 1995 for monitoring purposes [19]. As a matter of fact, the HFSSM pioneered in this field and several countries in Latin America followed this example by developing their own national scales to be included in national surveys for periodical monitoring. In 2004, Brazil included the *Brazilian Scale of Food Insecurity* (EBIA) into national Brazilian surveys; Haiti, Guatemala and Ecuador, among other countries, developed the *Latin American and Caribbean Food Security Scale* (Escala Latinoamericana y Caribeña de Seguridad Alimentaria—ELCSA); and in 2008 Mexico developed its adaptation of the ELCSA, called *Mexican Food Security Scale* (EMSA). Peculiar to these scales is the availability of two different survey modules, one for households with children and one for households without children and made up of a different number of items. Finally, beside these country-specific applications of the experience-based approach to measuring food insecurity, in 2013

[2] The aim of this first part of the work was mainly to provide a general framework for the topic and clarify that the expression "food insecurity" technically refers to a multitude of aspects that relate to different *dimensions*. However, in order to avoid confusion and enable an agile treatise of the subject, hereafter "food insecurity" will specifically be meant at the individual or household level and interpreted as the set of the restrictions in accessing food due to limited resources (or, equivalently, *resources-constrained access to food*). This choice will also facilitate conceiving food insecurity as a measurable construct.

the Food and Agriculture Organization of the United Nations (FAO) launched the Voices of the Hungry project (VoH) and developed the Food Insecurity Experience Scale (FIES) conceived as a global adaptation of HFSSM and ELCSA [16]. The FIES is based on people's responses to only 8 dichotomous items and, by means of an ad-hoc methodology that grounds on the Item Response Theory (IRT), and more specifically on the Rasch model, it is the first food insecurity measurement system based on experiences that generates *formally comparable* measures of food insecurity across countries. As such, it is one of the official measurement tool for monitoring progresses toward Target 2.1 of the SDGs, being the scale used to compute the related Indicator 2.1.1, (*Prevalence of food insecurity at moderate and severe levels based on FIES*) [3, 16].

Although the national and regional scales proved to be adequate tools for measuring and monitoring access to food within each country [12, 29], the need for a global monitoring, such as that sought in the context of the SDGs, raised the issue of comparing results from applications of different scales in different countries [5]. In fact, despite sharing a common evolution, each national scale uses specific thresholds to measure prevalences of food insecurity for *nominally* the same level of severity. Moreover, comparability issues also arises in the context of each national/regional scale, when considering the adult and the children-referenced versions of the same scale. This work aims at filling this gap by addressing comparability issues in the context of the experience-based food insecurity scales within a statistical framework. Specifically, methods and techniques from the statistical field of the educational testing are applied, with the goal of computing thresholds on the different scales that can be considered as "equivalent". Both classical and Item Response Theory (IRT)-based techniques are employed and two different comparability studies are presented:

1. Comparison between the FIES and national food insecurity scales in use in some countries in Latin America. Equating analyses are conducted between FIES and ELCSA in Guatemala; FIES and ELCSA in Ecuador; FIES and EMSA in Mexico; and between FIES and EBIA in Brazil.
2. Comparison between household and children-referenced scales within each national context. This analysis is conducted for ELCSA in Guatemala, ELCSA in Ecuador, EMSA in Mexico and EBIA in Brazil.

Data used for the analysis were collected by the single countries and are available for downloading on the internet at the websites of the Statistical National Institutes of Guatemala, Ecuador, Mexico and Brazil. Implementation of the equating methods was performed on the free software R (http://www.r-project.org) using, among others, the packages *RM.weights* [6], *equate* [1] and *plink* [30]. The remaining of the paper is organized as follows: Sect. 2 describes more in depth the features of the experience-based food insecurity scale; Sect. 3 presents the data and is devoted to describe the pillars and the methods of the Test Equating; Sect. 4 presents the main results; and Sect. 5 concludes with some remarks and possible directions for future works.

2 Experience-Based Food Insecurity Scale

As already mentioned, the FIES is strongly based on the ELCSA, which in turn represents a common ancestor for other scales in use in Latin America (EMSA, EBIA, etc.). As a consequence, all these scales largely share the same cognitive content of the items, which constitutes the promising ground on which addressing comparability. Nevertheless, the FIES and the national scales show important differences. First of all, national scales measure food insecurity at the household level, while the FIES produces national measures of food insecurity at the *individual level*. Secondly, national scales have a reference period of 3 months, while the FIES refers to the 12 months previous to the interview. Thirdly, and perhaps most importantly, national scales compute prevalences of food insecurity following a *deterministic* methodology based on raw scores (number of affirmative responses) and use discrete thresholds (expressed in terms of raw scores) for computing prevalences of food insecurity at different levels of severity. On the other hand, VoH methodology for the FIES is *probabilistic* in nature in that it fits the Rasch model to the data, models access to food by means of a probabilistic distribution and computes prevalences of food insecurity using thresholds on the continuum latent trait.

2.1 National and Regional Scales of Food Insecurity: ELCSA, EMSA, EBIA

The survey modules on which ELCSA, EMSA and EBIA are built have strong similarities [12, 29]. They all account for the three domains of the access dimension of food insecurity discussed in the previous section, aim at measuring food insecurity at the *household* level and all adopt the same reference period of 3 months previous to the day of the module administration. As far as the methodology is concerned, ELCSA, EMSA and EBIA agree on a similar procedure that can be summarized in few steps [12, 29]:

1. Computation of a *raw score* for each household: by counting the number of items affirmed by that household. Raw scores represent an *ordinal measure* of food insecurity: the highest the raw score, the more severe the level of food insecurity.
2. Computation of prevalences of food insecurity at three levels of severity: mild, moderate and severe. Prevalences are computed as percentages of households in the sample that scored within a certain range expressed in terms of raw scores and with different thresholds depending on whether children live in the household or not (Table 1).
3. Data validation. Homogeneity of the items comprising the scale is assessed by fitting the Rasch model to the data.

Moreover, each national scale makes use of two different versions of the survey module, distinguishing between households with children (i.e. people under the age

Table 1 Classifications of food insecurity using national scales (ELCSA, EMSA and EBIA) and corresponding ranges of the raw scores for households with and without children

Scale	Food insecurity level	Households without children	Households with children
ELCSA	Mild	1–3	1–5
	Moderate	4–6	6–10
	Severe	7–8	11–15
EMSA	Mild	1–2	1–3
	Moderate	3–4	4–7
	Severe	5–6	8–12
EBIA	Mild	1–3	1–5
	Moderate	4–6	6–10
	Severe	7–8	11–15

of 18 years) and households without children. The first group of survey modules is usually made up of 6 to 9 household-referenced items and, for the sake of simplicity, the scale obtained from this set of items will be referred to, in this work, as the *Adult* scale. The second one integrates the first one by adding from 6 to 7 extra children-referenced questions and the scale obtained from this set of items will be referred to as the *Children* scale. The two survey modules thus encompass a different number of items and, from each of them, a scale is built that uses different thresholds to compute prevalences of food insecurity that should be meant to reflect the same level of severity. Prevalences derived from the two scales are then considered jointly in order to derive national prevalences of food insecurity.

It is worth highlighting that, as reported in Table 1, the thresholds used to compute categories of food insecurity that *nominally* reflect the same level of severity (mild, moderate or severe), are country (or regional)-specific. As a matter of fact, these thresholds were not chosen in order to assure comparability among countries (no matter how geographically close to each other they might be) nor in light of clear statistical properties, but according to opinions of experts from the nutrition and social sciences fields. The same consideration holds for the thresholds chosen for the household referenced-scale and the children-referenced scale within each national context. As a consequence, there is no clear guarantee that, for example, a raw score of 7 truly reflects the same level of severity in Mexico and Brazil, or that, applying ELCSA in Guatemala, 7 and 11 can be considered as equivalent scores in households without and with children, respectively.

2.2 The Food Insecurity Experience Scale (FIES)

Inspired by Target 2.1 of the SDGs, the Voices of the Hungry (VoH) project of the Food and Agriculture Organization developed the Food Insecurity Experience Scale

Table 2 FIES Survey Module (FIES-SM) for individuals and with a reference period of 12 months

Items	Abbreviations
During the last 12 months, was there a time when, because of lack of money or other resources:	
1. You were worried you would not have enough food to eat?	WORRIED
2. You were unable to eat healthy and nutritious food?	HEALTY
3. You ate only a few kinds of foods?	FEWFOOD
4. You had to skip a meal?	SKIPMEAL
5. You ate less than you thought you should?	ATELESS
6. Your household ran out of food?	RUNOUT
7. You were hungry but did not eat?	HUNGRY
8. You went without eating for a whole day?	WHLDAY

(FIES), designed to have cross-cultural equivalence and validity in both developing and developed countries, aiming at producing comparable prevalences of food insecurity at various levels of severity [16]. As reported in Table 2, the FIES Survey Module is made up of 8 dichotomous items accounting for the three domains of access to food. Since 2014, the FIES Survey Module (FIES-SM) is part of the Gallup World Poll (GWP) Survey, from Gallup Inc. [28], a survey that is repeated every year in over 150 countries and administered to a sample of adult individuals (aged 15 or more) representative of the national population. This has practically allowed to reach countries that do not have a national measurement system for food insecurity, yet. In accordance with the characteristics of the GWP, the version of the FIES-SM here considered refers to a period of 12 months prior to the survey administration and investigates food insecurity at the level of adult individuals (people aged older than 15 years), which represents a first difference between FIES and the national scales.

However, the main difference between the two is in the methodology used [16]. The Voh methodology developed for the FIES employs a probabilistic model not only as a validation tool (for assessing homogeneity of the items in the scale), but also for computing measurements of food insecurity. In fact, food insecurity is treated as a *latent trait* whose measurement is achieved by means of some "observables" (the items' answers) and a probabilistic model that links the two. The Rasch model (also known as the one-parameter logistic model or 1PL model) is one of the most simple model that can serve this purpose while, at the same time, assuring a set of favourable measurement properties [17, 26]. It was proposed in the context of educational testing, where the purpose is generally to score students based on a set of questions (items) and, according to this model, the probability of a respondent to correctly answering the $j-$th item is modelled as a logistic function of the distance between two parameters, one representing the item's severity (b_j) and one representing the respondent's ability (θ):

$$P_j(\theta) = P(X_j = 1|\theta; b_j) = \frac{\exp(\theta - b_j)}{1 + \exp(\theta - b_j)}. \quad (1)$$

The Rasch model provides a sound statistical framework to assess the suitability of a set of items for scale construction and comparing performance of scales. Basic assumptions are unidimensionality, local independence, monotonicity, equal discriminating power of the items and logistic shape of the Item Response Functions (IRFs). Moreover, it has several interesting properties for which it earned its success among social science measurement models, like sufficiency of the raw score, independence between items and examinees' parameters, and invariance property [18]. In the context of food insecurity, the item's severity can be interpreted as the severity of the restrictions in food access represented by each item while the ability parameters are to be meant as the overall severity of the restrictions in accessing food that the respondent had to face (in light of her answers to the items in the survey module). From the point of view of the Stevenson's classification of scales [27], this way of measuring food insecurity guarantees the construction of an *interval* scale as opposed to the *ordinal* scale obtained from the methodology employed, for instance, for ELCSA, EMSA and EBIA and named *deterministic* as opposed to the *probabilistic* developed for the FIES. Moreover, prevalences obtained by means of the FIES are guaranteed to be comparable across countries, thanks to the implementation of an *equating step* for which estimates of the model parameters obtained in a single application of the scale are adjusted on the FIES Global Standard scale, a set of item parameters serving as a reference metric and based on application of the FIES in all countries that were covered by the GWP survey in 2014, 2015 and 2016 [16] (Fig. 1). Finally, each respondent is assigned a probabilistic distribution of his/her food insecurity along the latent trait, depending on his/her raw score. This distribution is Gaussian with mean equal to the adjusted (to the Global Standard) respondent parameter and standard deviation equal to the adjusted measurement error for that raw score. As a last step, this mixture of distributions is used to compute the percentage of population whose severity is beyond a fixed threshold on the latent trait, calculated as a weighted sum across raw scores, with weights reflecting the proportions of raw scores in the sample. While theoretically it is possible to compute percentages of population beyond each and every value on the continuum, the VoH methodology suggests the computation of two prevalence rates corresponding to choosing thresholds on the Global Standard metric set at the severity of items ATELESS (−0.25) and WHLDAY (1.83) (Fig. 1). The resulted indicators of food insecurity take the name of *Prevalence of Experienced Food Insecurity at moderate or severe levels* ($FI_{Mod+Sev}$) and *Prevalence of Experienced Food Insecurity at severe levels* (FI_{Sev}), respectively. However, in order for these quantities to be valid and reliable measurements of food insecurity, a validation step must be undertaken in each and every application of the scale. This is commonly performed by computing goodness-of-fit statistics of the Rasch model (e.g. Infit, Outfit and Rasch reliability statistics) that assess the good behaviour of the items and by performing a Principal Component Analysis (PCA) on the residuals to investigate the existence of a second latent trait. For more details

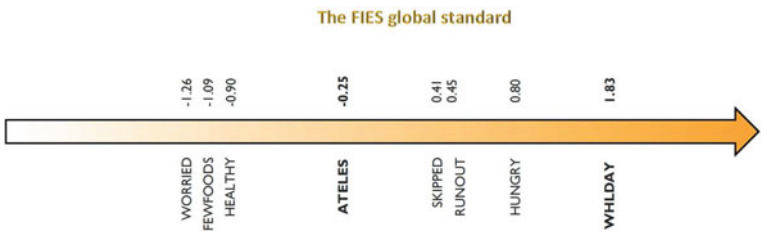


Fig. 1 The FIES global standard

on the usage of the Rasch model as a measuring tool for food insecurity we refer the reader to [23], while for more insights in the VoH methodology for the FIES we refer to [16].

3 Data and Methods

3.1 Data

Data referred to Guatemala were collected in the *Encuesta Nacional de Condiciones de Vida* (ENCOVI) conducted by the *Instituto Nacional de Estadística* (INE) in 2014 and the sample used included 11433 households. Data referred to Ecuador were collected in the *Encuesta Nacional de Empleo y Desempleo* (ENCOVI) conducted by the *Instituto Nacional de Estadísticas y Censos* (INEC) in 2016 and the sample used included 16716 households. Data referring to Mexico were collected in the *Encuesta Nacional de Ingresos y Gastos de los Hogares* (ENIGH) conducted by the *Instituto Nacional de Geografia e Estatística* in 2014 and the sample used included 19479 households. Data referring to Brazil were collected in the *Pesquisa Nacional de Amostra de Domicílios* (PNAD) conducted by the *Instituto Nacional de Geografia e Estatística* (IBGE) in 2013 and the sample used included 116543 households. All samples were representative of the corresponding national populations.

3.2 Test Equating

Scores deriving from tests usually have an important role in the decision making process that brings to excluding some candidates for a job or scholarship position, or adopting specific public policy strategies in order to take action on a public relevant issue. Evidently, this requires that tests to be administered in multiple occasions, as it is the case for the admission college tests that are held in specific *test dates* during the year. Therefore, a crucial consideration arises: if the same questions were

included in the tests, students that already took the test would have an advantage and the test would rather measure the degree of exposure of students to past tests than their ability on some specific subject. At the same time, it is important that all students take the "same" test, in order to fairly compare performances and make decisions accordingly. This issue is commonly addressed by administering on every test date a different version of the same test, called *test form*, that is built according to certain content and statistical *test specifications*. Nonetheless, minor differences might still occur among different test forms, one resulting slightly more difficult than the others. Therefore, in order to evenly score students that took multiple test forms and establish if a poorer performance is due to a less skillful respondent and not to a more difficult test, a procedure is needed to make tests comparable. This procedure is called *equating* and it is formally defined as the statistical process that is used to adjust for differences in difficulty between tests forms built to be similar in content and difficulty, so that scores can be used interchangeably [11, 21]. Every test equating should meet some fundamental equating requirements and needs the specification of both a data collection design and of one or more methods to estimate an equating function. All these aspects will be discussed in the remaining of this section.

3.2.1 Equating Requirements

Equating scores on two test forms X and Y must meet some requirements that assure that the equating to be meaningful and useful (i.e. equated scores can be used interchangeably). The following five requirements are globally considered of primary importance for an equating to be run, although they would better be considered as general guidelines than easily verifiable conditions:

- **Equal Construct Requirement** Tests that measure different constructs should not be equated.
- **Equal Reliability Requirement** Tests that measure the same construct but differ in reliability should not be equated.
- **Symmetry Requirement** Equating function that equate scores on X to scores on Y should be the inverse of the equating function that equate scores on Y to scores on X.
- **Equity Requirement** For the examinee should be a matter of indifference which test will be used.
- **Population Invariance Requirement** Equating function used to equate scores on X and scores on Y should be *population invariant* in that the choice of a specific sub-population used to compute the equating function should not matter.

It might be the case that the two tests to be equated do not satisfy all five requirements. For example, they could differ in length and statistical specifications, with consequences on the "Equal Reliability requirement" and "Equity requirement". In fact, a longer test would be in general more reliable and, if a poorly skillful examinee had to be scored, he or she would have more chance to score higher if administered the

shortest test. The aforementioned requirements assure that scores derived from tests that do meet all of them can be used interchangeably while, if they do not all strictly hold, the exercise would rather be addressed as a weaker analysis of comparability named *linking* [11, 21].

3.2.2 Equating Designs

There are basically two ways in which data collection designs can account for differences in the difficulty of two or more test forms in test equating, namely either by the use of "common examinees" or the use of "common items" [11, 21]. In the first case the same group of examinees (or two random samples of examinees from the same target population) take both tests. In this case, any difference in the scores is attributable to differences in the test forms. Examples of this category are the "Single-Group" (SG) and the "Equivalent-Groups" (EG) designs. In the second case, a set A of common items called *anchor test* is included in both test forms in order to account for such differences. Therefore, any difference between scores on the anchor test is due to differences among examinees. Data designs that use this method are called "Non-Equivalent groups with Anchor Test" (NEAT) designs.

3.2.3 Equating Methods

Several equating methods have been proposed and applied to equate observed scores on equatable tests. In this section an overview of the most common and popular methods is provided, starting from the observed-score methods of mean, linear and equipercentile equating and ending up with the true score equating in the context of IRT. All these methods have been implemented for the two comparability studies between experience-based food insecurity scales that are presented in this work.

Observed-Score Equating Methods

Let X and Y be two tests (or two forms of the same test) scored correct/incorrect (1/0). Scores on test X and Y will be denoted as random variables $\mathbf{X}$ and $\mathbf{Y}$ with possible values, respectively $x_k, (k = 0, \ldots, K)$ and y_l (for $l = 0, \ldots, L$), where K and L are the lengths of tests X and Y, respectively. We denote the score probabilities of $\mathbf{X}$ and $\mathbf{Y}$ by $r_k = P(\mathbf{X} = \mathbf{x}_k)$ and $s_l = P(\mathbf{Y} = \mathbf{y}_l)$. The cumulative distribution functions of $\mathbf{X}$ and $\mathbf{Y}$ are denoted by $F(x) = P(\mathbf{X} \leq x)$ and $G(x) = P(\mathbf{Y} \leq y)$ and the moments are $\mu_X = \mathbf{E}(\mathbf{X})$, $\sigma_X = SD(\mathbf{X})$ and $\mu_Y = \mathbf{E}(\mathbf{Y})$, $\sigma_Y = SD(\mathbf{Y})$, respectively.

Mean Equating

In mean equating, test form X is assumed to differ from test form Y by a constant amount along the scale. For example, if form X is 2 points easier than form Y for

low-ranking examinees, the same will hold for high-ranking examinees. In mean equating, two scores on different forms are considered equivalent (and set equal) if they are the same (signed) distance from their respective means, that is

$$x - \mu_X = y - \mu_Y. \tag{2}$$

Then, solving for y, the score on test Y that is equivalent to a score x on test X, and called $m_Y(x)$, is

$$m_Y(x) = y = x - \mu_X + \mu_Y. \tag{3}$$

Clearly, mean equating allows for the means to differ in the two test forms.

Linear Equating

In linear equating, difference in difficulty between the two tests is not constraint to remain constant but can vary along the score scale. In this equating method, scores are considered equivalent and set equal if they are an equal (signed) distance from their means in standard deviation units, that is the two standardized deviation scores (z-scores) on the two forms are set equal

$$\frac{x - \mu_X}{\sigma_X} = \frac{y - \mu_Y}{\sigma_Y} \tag{4}$$

from which the score on test Y equivalent to a score x on test X, and that is called $l_Y(x)$, is

$$l_Y(x) = y = \frac{\sigma_Y}{\sigma_X} x + \left[\mu_Y - \frac{\sigma_Y}{\sigma_X}\mu_X\right]. \tag{5}$$

where $\frac{\sigma_Y}{\sigma_X}$ can be recognized as the *slope* and $\mu_Y - \frac{\sigma_Y}{\sigma_X}\mu_X$ as the *intercept* of the linear equating transformation. Linear equating allows for both means and scale units to differ in the two test forms.

Equipercentile Equating

In the equipercentile equating method a curve is used to describe differences between scores in the two forms. Basic criterion for the equipercentile equating transformation is that the distribution of the scores on Form X converted to the Form Y scale is equal to the distribution of scores on Form Y. Scores on the two forms are considered -and set- equivalent if they have the same *percentile rank*. We adopt here the definition of equipercentile equating function given by Braun and Holland in [4]. Let's consider the random variables $\mathbf{X}$ and $\mathbf{Y}$ representing the scores on forms X and Y and F and G their cumulative distribution functions. We call e_Y the symmetric equating function converting Form X scores into scores on Form Y scale and $G^\star$ the cumulative distribution function of $e_Y(\mathbf{X})$, that is the cdf of the scores on Form X converted to the Form Y scale. Function e_Y is the *equipercentile equating function* if $G^\star = G$.

According to the definition of Braun and Holland, if **X** and **Y** are continuous random variables, then

$$e_Y(x) = G^{-1}[F(x)], \tag{6}$$

is an equipercentile equating function, where G^{-1} is the inverse of G. This definition meets the "Symmetric requirement" and, given a Form X score, its equivalent on the Form Y scale is defined as the score having the same percentage of examinees at or below it.

IRT-Based Equating Methods: The IRT-True Score Equating (IRT-TS)

Equating different forms of the same test using the IRT-True Score equating (IRT-TS) is a three steps process [10]:

1. *Estimation*: Fit an IRT-model to the data for both tests.
   This step consists in assessing goodness-of-fit of a specific IRT model and estimating item parameters for both forms. In the case of the Rasch model, in light of the sufficient statistics property, estimates of the item severities do not depend on the group of examinees and therefore the IRT-TS based on the Rasch model can be claimed to meet the "Population Invariance requirement", since it produces results that are sample-independent.
2. *Linking*: Put parameters' estimate on a common metric through a linear transformation based on a set A of common items.
   In this second step, a linear transformation is used to bring parameter estimates to a common IRT scale. In fact, "if an IRT model fits the data, then any linear transformation of the θ-scale also fits the data, provided that the item parameters are transformed as well" [21]. Let consider Form X made up of J dichotomously scored items administered to N examinees and let consider the Rasch model to fit the data. Then, if P and Q are Rasch scales that differ by a linear transformation, the item severities b_j, $j \in \{1, \ldots, J\}$ are related as follows

   $$b_{jQ} = Ab_{jP} + B, \qquad j \in \{1, \ldots, J\}$$

   and the same relationship holds for the ability parameters. A useful way to express the constants A and B is through the mean and standard deviation of the item parameters in both scales

   $$A = \frac{\sigma(b_Q)}{\sigma(b_P)}, \qquad B = \mu(b_Q) - A\mu(b_P).$$

   In equating two different forms of the same test with a set of common items administered to non-equivalent groups, it is possible to exploit this linear relationship through the so called *Mean/Sigma transformation method* [21], which uses means and standard deviations of item parameter's estimates of only those

items in the anchor test. More specifically, given Form X and Form Y with a set A of common items, estimates of the difficulty parameters for items in the set A in the two calibrations are linked via a linear transformation and used to compute the coefficients A and B of this transformation. Once the transformation is estimated, it can be applied to transform the ability parameters on one Form to the corresponding parameters on the other Form, thus enabling comparability between the two Forms.

3. *Equating*: Get equivalent expected raw scores through the Test Characteristic Curves of the two tests (TCC).
   Once the metrics of the two Forms are put on the same scale (that can be either the scale of one of them or a third scale) it is finally possible to compare performance of examinees taking the two Forms. However, as it often happens with standardized tests, reported scores could be expressed in terms of raw scores and, if this is the case, a further step is needed. Within the framework of IRT, it is possible to mathematically relate ability estimates to specific *true scores* on each test form. The *IRT- True Score equating* method computes equivalent true scores in the two forms and considers them, as it is common in the practice of equating, as equivalent observed scores [22]. Given Form X and Form Y two test forms measuring the same ability θ with respectively n_X and n_Y items and given both item and ability estimates are on the same scale through a linear transformation, the estimated true scores on the two forms are related to θ by the so called *Test Characteristic Curve* as follows

$$T_X = \sum_{j=1}^{n_X} \hat{P}_j(\theta), \qquad T_Y = \sum_{i=1}^{n_Y} \hat{P}_i(\theta)$$

   where T_X (respectively T_Y) is the estimated true score for Test X (Y) and $\hat{P}_i(\theta)$ ($\hat{P}_j(\theta)$) is the estimated probability function for item j (i) (Fig. 2). Through the Test Characteristic Curve, an ability θ can thus be transformed into an estimated

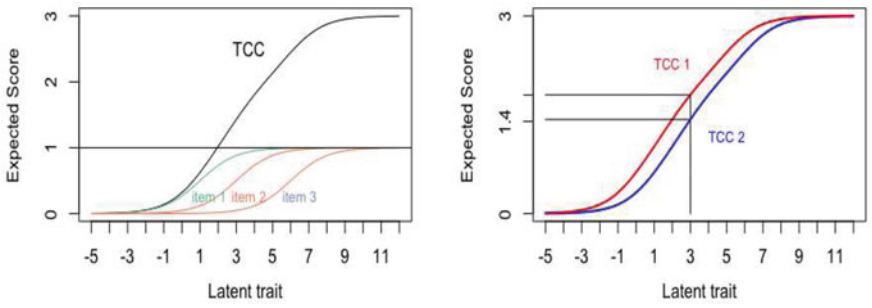


Fig. 2 Test Characteristic Curve (TCC) referred to a test of three items (left) and a pictorial description of the IRT-True Score (IRT-TS) equating method (right)

true score on the test form and, provided ability parameters on the two forms are put on the same metric, true scores corresponding to the same θ are considered equivalent.

4 Results

As a preliminary step to both equating studies, the Rasch model has been fitted to all eight datasets (an Adult scale and a Children scale in the four countries), and a validation step was performed to confirm the good behaviour of the scale. In all eight applications it was possible to observe an overall good fit of the model. Assumptions of equal discrimination of the items was certainly met, thanks to item Infit statistics entirely in the range of (0.7, 1.3), confirming the strength and consistency of the association of each item with the underlying latent trait (compare [16, 23]). Moreover, Outfit statistics were never as high as to warn misbehaviour due to highly unexpected response patterns, assessing the good performance of the items. Assumptions of conditional independence and unidimensionality of the items were assessed through computation of conditional correlations among each pair of items and submission of the correlation matrix to principal component factor analysis (PCA). All pairwise residual correlation were, in absolute value, smaller than 0.4 thus confirming that all correlations among items result from their common association with the latent trait. PCA performed on the matrix of residual correlations showed the presence of only one main dimension that, due to the cognitive content of the items, can thus be recognized as the *food access* dimension that the scales aim at measuring. Finally, overall model fit is assessed by Rasch reliability statistics (proportion of total variation in true severity in the sample that is accounted for by the model), ranging between 0.65 (Mexico) and 0.79 (Guatemala) for the Adult scale and between 0.80 (Mexico) and 0.86 (Guatemala) for the Children scale, confirming a good overall discriminatory power for all scales. Sporadic departures from this irreproachable behaviour could only be attested for one or two items in the Children scale (like a residual correlation of 0.6 between two item of the ELCSA in Guatemala) that however never compromised the good performance of the overall scale.

4.1 First Study: Equating FIES and National Scales

The aim of this comparability study is to find raw scores on the national scales EBIA, EMSA and ELCSA that can be considered *equivalent* to the continuum FIES global thresholds used to compute the two indicators $FI_{Mod+Sev}$ and FI_{Sev}, namely -0.25 and 1.83. However, it is worth noticing that, since VoH methodology uses thresholds on the continuum while national scales methodology uses discrete thresholds, the equivalent raw score will almost never exactly produce the same prevalence obtained with the VoH thresholds.

As it is currently set up and implemented, the FIES Module refers to adults (people aged 15 or above). Therefore, in order to meet the "Equal Construct requirement", the modules of the national scales administered to households *without* children have here been considered. Technically, the FIES Survey Module and the survey modules of the national scales (households without children) will thus serve the role of *test forms* of the same test that are to be equated. This was ultimately made possible in light of the common history that brought to the development of these scales (i.e. FIES, ELCSA, EMSA, EBIA), which assures that, despite some differences such as the level of the measurement and the reference time (see Sect. 2), the survey modules used to collect data have very strong similarities and share the same dichotomous structure (possible answers are "Yes/No").

This first study was carried out by implementing the following methods:

1. **IRT True Score** (IRT-TS) equating.
2. **Linking** via a linear transformation applied to ability parameters.
3. **Minimization** of the difference between prevalences of food insecurity.

The IRT-TS equating method was implemented in the context of the NEAT equating design. In this work, the set A of common items was computed according to an iterative procedure that starts with all items considered as in common (apart from the ones classified as unique a priori) and then discards one item at a time beginning from the one that exceeds the tolerance threshold of 0.5 the most. Algorithm ends when a set A of items all within this threshold is found. Item WHLDAY was considered as *unique* a priori in all four equating analyses due to its different cognitive content in the considered scales: more severe in the FIES since it refers to "not eating for a whole day", and less severe in the national scales where it reports on members of the household that either only ate once *or* went without eating for a whole day. The Standard Error of Equating (SEE) for the IRT True-Score equating method was estimated using 1000 bootstrap replications [21]. The second and third methods can be considered as either variations of the IRT-TS or techniques that might sound particularly reasonable in the present context. They were explored for investigation purposes and the obtained scores won't be claimed to be "equivalent", but rather "corresponding" scores. In fact, the second method (Linking) consists in considering the linear transformation obtained in the second step of the IRT-TS method and applying it to the estimated ability parameters of the Rasch model. Once ability parameters are adjusted to the Global Standard metric, the raw score corresponding to the ability parameters that are closer to the two VoH thresholds are considered as *corresponding* raw score. On the other hand, the third method (Minimizing) consists in computing prevalences of food insecurity at the household level applying the FIES methodology to the data used for the national scales and comparing the prevalences so obtained with the percentages of population scoring from a certain raw score on. The two raw scores that realize the minimum distance with the two VoH global thresholds (in terms of prevalences) are considered as the *corresponding* raw scores in accordance to this method.

Results from the first comparability study are summarized in Tables 3 and 4, which report the raw scores on the national scales that are computed equivalent to the VoH

Table 3 Equated raw scores on the national scales corresponding to the VoH threshold for $FI_{Mod+Sev}$ (−0.25 on the global standard)

Food insecurity	Internal monitoring	IRT-TS rasch (SEE)	Linking	Min. Diff.
ELCSA (Guatemala)	4	3.3 (0.19)	3	4
ELCSA (Ecuador)	4	4.2 (0.14)	4	4
EMSA (Mexico)	3	2.0 (0.23)	2	2
EBIA (Brazil)	4	4.0 (0.09)	4	5

Table 4 Equated raw scores on the national scales corresponding to the VoH threshold for FI_{Sev} (1.83 on the global standard)

Food insecurity scales	Internal monitoring	IRT-TS rasch (SEE)	Linking	Min. Diff.
ELCSA (Guatemala)	7	7.8 (0.18)	8	8
ELCSA (Ecuador)	7	7.1 (0.18)	7	8
EMSA (Mexico)	5	6.0 (0.26)	6	6
EBIA (Brazil)	6	7.9 (0.07)	8	8

global thresholds used for the indicators $FI_{Mod+Sev}$ and FI_{Sev}, respectively. Table 3 shows that the threshold used for computing $FI_{Mod+Sev}$ and corresponding to the severity of item ATELESS on the Global Standard metric (i.e. −0.25) might reflect a *less severe* condition of food insecurity compared to the one measured by the national scales for the moderate category of food insecurity. In fact, all the equated raw scores are either equal to or around one point less than the thresholds currently used by ELCSA, EMSA and EBIA for this category of food insecurity. On the contrary, the threshold used for FI_{Sev} and corresponding to the severity of item WHLDAY on the Global Standard metric (i.e. 1.83) generally reflects a *more severe* condition of food insecurity than the one captured by the national scales for the severe level of food insecurity, Table 4 reporting equated raw scores that are either equal to or one point higher than the national thresholds currently in use for this category.

4.2 Second Study: Comparing Household- and Children-Referenced Item Scales

This second analysis aims at comparing the Adult and Children scales within each national context. To this purpose, we implemented the Single Group (SG) data collection design by considering the scores obtained by the households with children

on both survey modules (the one containing only adult and household-referenced questions and the one containing also children-referenced questions). This equating design is usually not easily implemented, since it requires the same group of respondents to be administered two different forms of the same test, resulting in an expensive and time-consuming procedure. However, we here could exploit the fact that the survey module for the Children scale is simply an extended version of the module for the Adult scale (see Sect. 2) and as such, we can "imagine" to administer the adult referenced survey to households with children just by dropping children-referenced items. With regards to the equating requirements (see Sect. 3.2.1), it is worth noticing that the two survey modules have different length and, as such, the obtained scales could have different reliability which, in turn, could potentially challenge the "Equal Reliability Requirement". Equating of the Adult and Children scales in the four countries was carried out through implementation of four equating methods: IRT True Score equating with the Rasch model, Mean, Linear and Equipercentile equating methods [21]. The first method is IRT-based while the other four are classical methods of equating, that do not rely on any model to fit the data but only on the observed raw scores.

Tables 5 and 6 show raw scores on the Children scale that are computed equivalent to raw scores on the Adult scales that are used as *lower* thresholds for the moderate and severe categories of food insecurity. Results suggest that the national thresholds currently used for nominally the same levels of severity could reflect different degrees of the severity of access to food. This is particularly evident when looking at the most severe category of food insecurity, where raw scores on the Children scale that are computed equivalent to the lower thresholds on the Adult scale are around one point *higher* than the thresholds currently in use for households with children for ELCSA in Guatemala and EMSA in Mexico and between one and two points *lower* for EBIA (Tables 5 and 6, column "Severe"). On the other hand, the corresponding raw scores for moderate food insecurity mainly align with the thresholds currently in use for this category (Tables 5 and 6, column "Moderate"). Interestingly,

Table 5 Raw scores on the Children scale corresponding to raw scores 4 and 7 on the Adult scale (lower thresholds for moderate and severe food insecurity, respectively) and related Standard Error of Equating (SEE). Left: ELCSA (Guatemala). Right: ELCSA (Ecuador)

Equating method	Moderate (SEE)	Severe (SEE)
IRT-TS	6.2 (0.09)	12.1 (0.10)
Mean	6.6 (0.07)	12.2 (0.07)
Linear	6.5 (0.07)	11.7 (0.11)
Equip	6.3 (0.09)	11.3 (0.15)
IRT-TS	5.8 (0.09)	12.1 (0.11)
Mean	6.4 (0.05)	11.6 (0.05)
Linear	6.2 (0.07)	11.0 (0.10)
Equip	5.8 (0.13)	11.2 (0.16)

Table 6 Left: Raw scores on the Children scale corresponding to raw scores 3 and 5 on the Adult scale (lower thresholds for moderate and severe food insecurity, respectively) and related Standard Error of Equating (SEE), EMSA (Mexico). Right: Raw scores on the Children scale corresponding to raw scores 4 and 7 on the Adult scale (lower thresholds for moderate and severe food insecurity, respectively) and related Standard Error of Equating (SEE), EBIA (Brazil)

Equating method	Moderate (SEE)	Severe (SEE)
IRT-TS	4.8 (0.12)	8.7 (0.13)
Mean	5.5 (0.05)	9.5 (0.05)
Linear	5.1 (0.07)	8.6 (0.10)
Equip	4.8 (0.13)	8.1 (0.14)
IRT-TS	4.8 (0.08)	8.7 (0.09)
Mean	5.5 (0.04)	9.0 (0.04)
Linear	5.5 (0.04)	8.8 (0.07)
Equip	4.7 (0.05)	8.3 (0.14)

minor differences emerge between the behaviour of the equated scores for ELCSA in Guatemala and Ecuador, possibly due to specific features of the phenomenon in the two countries, confirming the importance of an equating analysis even between different applications of the same scale. Finally, it is noteworthy that, among all implemented methods, the Equipercentile equating method is the one whose results mostly resemble the current adopted thresholds.

5 Conclusions

The present work presented two studies investigating comparability between experiential scales of food insecurity. The first study aimed at addressing comparability between the FIES and three national scales (ELCSA, EMSA and EBIA) in Guatemala, Ecuador, Mexico and Brazil. Results show that, in general, the VoH threshold used for computing the indicator *Prevalence of Experienced Food Insecurity at moderate or severe levels* ($FI_{Mod+Sev}$) and corresponding to the severity of item ATELESS on the Global Standard (-0.25) seems to reflect a *less severe* level of food insecurity than that described by the thresholds used by the national scales (and expressed in terms of raw scores) for the same level of severity. On the other hand, the VoH threshold used for computing the indicator *Prevalence of Experienced Food Insecurity at severe levels* (FI_{Sev}) and corresponding to the severity of item WHLDAY on the Global Standard (1.83) seems to reflect a *more severe* condition than the one measured through the national thresholds for this level of severity. The relevance of such a result for the practice of food insecurity measurement is self-evident. In fact, the possibility of comparing prevalences of food insecurity derived from applying different scales represents an important step in the direction of realizing a more reliable monitoring of the global progresses towards the goal of food security for

all people worldwide (as expressed by Target 2.1 of the Sustainable Development Goals) and, as such, it is expected to gain increasing attention by practitioners and decision makers in the field.

Additionally, a second study investigated the issue of comparability between food insecurity scales referred to households without children (Adult scale) and households with children (Children scale), within each national context. Results show that the national thresholds currently used to compute prevalences of food insecurity at *nominally* the same level of severity among households with and without children might not always represent the same degree of the restrictions on food access. This seems especially evident for the most severe category of food insecurity, where current thresholds on the Children scale are lower than those computed as equivalent to the thresholds used on the Adult scale for this category.

Future studies are expected to shed light on possible reasons and additional aspects of this topic. A more detailed characterization of the phenomenon of food insecurity across countries as well as in households with and without children (especially from a social and economic point of view) will likely better motivate and clarify the distinctive behaviour of different food insecurity scales. Furthermore, similar analyses to be conducted on other experiential scales of food insecurity might contribute to reach a deeper knowledge of the phenomenon. Significant examples being the HFSSM in North America, as well as applications of ELCSA in countries of the Latin America beyond the ones here considered.

References

1. Albano, A.D., et al.: equate: an R package for observed-score linking and equating. J. Stat. Softw. **74**(8), 1–36 (2016)
2. Assembly, U.N.G.: Universal Declaration of Human Rights, vol. 3381. Department of State, United States of America (1949)
3. Ballard, T.J., Kepple, A.W., Cafiero, C.: The food insecurity experience scale: development of a global standard for monitoring hunger worldwide. Rome, Italy. FAO (2013)
4. Braun, H.I.: Observed-score test equating: A mathematical analysis of some ETS equating procedures. Test equating (1982)
5. Cafiero, C.: What do we really know about food security? Technical report, National Bureau of Economic Research (2013)
6. Cafiero, C., Viviani, S., Nord, M.: Rm. weights: weighted Rasch modeling and extensions using conditional maximum likelihood (2018). https://CRAN.R-project.org/package=RM.weights
7. Cafiero, C., et al.: Validity and reliability of food security measures. Ann. N. Y. Acad. Sci. **1331**(1), 230–248 (2014)
8. Carlson, S.J., et al.: Measuring food insecurity and hunger in the United States. J. Nutr. **129**(2), 510S-516S (1999)
9. Coates, J., et al.: Commonalities in the experience of household food insecurity across cultures: what are measures missing? J. Nutr. **136**(5), 1438S-1448S (2006)
10. Cook, L.L., et al.: IRT equating methods. Educ. Meas.: Issues Pract. (1991)
11. Dorans, N.J., Pommerich, M., Holland, P.W.: Linking and aligning scores and scales. Springer Science & Business Media (2007)
12. Científico de la ELCSA, C.: Escala Latinoamericana y Caribeña de Seguridad Alimentaria (ELCSA): Manual de uso y aplicaciones. Rome, Italy. FAO (2012)

13. FAO.1946: World Food Survey. Washington, DC. FAO (1946)
14. FAO.1996: World food summit. Rome, Italy. FAO (1996)
15. FAO.2001: The state of food insecurity in the world 2001. Rome, Italy. FAO (2001)
16. FAO.2016: Methods for estimating comparable rates of food insecurity experienced by adults throughout the world. Rome, Italy. FAO (2016)
17. Fischer, G.H., Molenaar, I.W.: Rasch models: foundations, recent developments, and applications. Springer Science & Business Media (2012)
18. Hambleton, R.K., et al.: Fundamentals of Item Response Theory, vol. 2. Sage (1991)
19. Hamilton, W.L., et al.: Household food security in the United States in 1995. Summary Report of the Food Security Measurement Project, vol. 20 (1997)
20. Jones, A.D., et al.: What are we assessing when we measure food security? a compendium and review of current metrics. Adv. Nutr. **4**(5), 481–505 (2013)
21. Kolen, M.J., Brennan, R.L.: Test equating, scaling, and linking: methods and practices. Springer Science & Business Media (2014)
22. Lord, F.M., Wingersky, M.S.: Comparison of IRT observed-score and true-score 'equatings'. ETS Res. Rep. Ser. **1983**(2), i–33 (1983)
23. Nord, M.: Assessing potential technical enhancements to the U.S. household food security measures. U.S. Department of Agriculture, Economic Research Service **TB-1936** (2012)
24. Organization, W.H., et al.: Sustainable Development Goals (2016)
25. Radimer, K.L., et al.: Understanding hunger and developing indicators to assess it in women and children. J. Nutr. Educ. **24**(1), 36S-44S (1992)
26. Rasch, G.: Probabilistic Models for Some Intelligence and Achievement Tests. Danish Institute for Educational Research, Copenhagen (1960)
27. Stevens, S.S., et al.: On the theory of scales of measurement (1946)
28. Tortora, R.D., Srinivasan, R., Esipova, N.: The gallup world poll. In: Survey Methods in Multinational, Multiregional, and Multicultural Contexts, pp. 535–543 (2010)
29. Villagómez-Ornelas, P., et al.: Validez estadística de la Escala Mexicana de Seguridad Alimentaria y la Escala Latinoamericana y Caribeña de Seguridad Alimentaria. Salud Pública de México **56** (2014)
30. Weeks, J.P., et al.: Plink: an R package for linking mixed-format tests using IRT-based methods. J. Stat. Softw. **35**(12), 1–33 (2010)

Position Weighted Decision Trees for Ranking Data

Antonella Plaia, Simona Buscemi, and Mariangela Sciandra

Abstract Preference data represent a particular type of ranking data where a group of people gives their preferences over a set of alternatives. Within this framework, distance-based decision trees represent a non-parametric tool for identifying the profiles of subjects giving a similar ranking. This paper aims at detecting, in the framework of (complete and incomplete) ranking data, the impact of the differently structured weighted distances for building decision trees. By means of simulations, we will compute the impact of higher/lower homogeneity in groups and different weighting structures both on splitting and on consensus ranking. The distances that will be used satisfy Kemeny's axioms and, accordingly, a modified version of the rank correlation coefficient τ_x, proposed by Emond and Mason, will be proposed and used for rank aggregation and class label in the tree leaves.

Keywords Weighted distances · Ranking · Kemeny · Consensus · Trees

1 Introduction

Distances between rankings and the rank aggregation problem have received a growing consideration in the past few years. Ranking and classifying are two simplified cognitive processes useful for people to handle many aspects of their life. When some subjects are asked to indicate their preferences over a set of alternatives (items), ranking data are called preference data. A critical issue involving rankings concerns the aggregation of the preferences in order to identify a compromise or a "consensus".

A. Plaia (✉) · S. Buscemi · M. Sciandra
Dipartimento di Scienze Economiche, Aziendali e Statistiche, University of Palermo, Viale delle Scienze, Edificio 13, 90128 Palermo, Italy
e-mail: antonella.plaia@unipa.it

S. Buscemi
e-mail: simona.buscemi@unipa.it

M. Sciandra
e-mail: mariangela.sciandra@unipa.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_15

Different approaches have been proposed in the literature to cope with this problem, but probably the most popular is the one related to distances/correlations. In order to get homogeneous groups of subjects having similar preferences, it is natural to measure the spread between rankings through dissimilarity or distance measures among them. In this sense, a consensus is defined to be the ranking that is the closest (i.e. it shows the minimum distance) to the whole set of preferences. A distance d between two rankings is a non-negative value, ranging in 0–Dmax, where 0 is the distance between a ranking and itself. Another possible way for measuring (dis)-agreement between rankings is in terms of a correlation coefficient: rankings in full agreement are assigned a correlation of +1, those in full disagreement are assigned a correlation of –1, and all others lie in between. A distance measure d can be transformed into a correlation coefficient c (and vice-versa) using the linear transformation $c = 1 - \frac{2d}{D_{max}}$.

One significant issue of interest in literature is: how profiles of people expressing similar preferences can be identified in terms of the subject-specific characteristics? In order to answer this question, different solutions have been proposed: distance-based tree models [17], distance-based multivariate trees for ranking [1], log-linearized Bradley-Terry models [10] and a semi-parametric approach for recursive partitioning of Bradley-Terry models for incorporating subject covariates [23].

Decision trees are a competitive tool compared with other state-of-the-art methods, in terms of predictive accuracy and, not less relevant, they are generally considered as being more comprehensible and interpretable than most other model classes [5]. Piccarreta [19] proposes binary trees for dissimilarity data, listing preference data too as a possible practical application, but the paper does not consider dissimilarities that are specific to preference data. The class of distance-based ranking models states that the probability of observing a specific ranking depends on the distance between the observed ranking and the modal ranking, i.e. "they assume a modal ranking and the probability of observing a ranking is inversely proportional to its distance from the modal ranking. The closer to the modal ranking, the more frequent the ranking is observed" [17]. Further development in the distance-based approach is again due to Lee et al. [17], who developed a distance-based tree model introducing a weighted distance, where the weights are linked to items in the preference ranking.

The possibility of using weights can take into account crucial concepts, totally ignored by classical non-weighted distances: item relevance and positional information. As a matter of fact, in some real situations, it might be relevant to give more importance to changes in the top-positions of rankings, or it might be more important to emphasize changes occurring in highly-relevant items rather than changes in less essential items. Distance measures with either position or element weights are well discussed by Kumar and Vassilvitskii [16]. The purpose of this paper is to investigate the effect of higher/lower homomeneity in groups and different weighting vectors on the tree. Particular attention is given to the weighted Kemeny distance and to the consensus ranking process for assigning a suitable label to the leaves of the tree. The stopping criterion for detecting the optimum tree is a properly modified τ_x [11].

The rest of the paper is organized as follows: Sect. 2 introduces different metrics between rankings, their properties and their weighted extension; Sect. 3 introduces

the weighted correlation coefficient; after a brief view on decision trees and the rank aggregation problem in the tree leaves in Sect. 4, in Sect. 5 we perform our analysis through a simulation study and, in the end, a short conclusion is presented (Sect. 6).

2 Distances Between Rankings

Ranking data arise when a group of n individuals (experts, voters, raters etc.) shows their preferences on a finite set of items (k different alternatives of objects, like movies, activities and so on). If the k items are ranked in k distinguishable ranks, a complete ranking or linear ordering is achieved [3]. A ranking π is, in this case, one of the $k!$ possible permutations of k elements, containing the preferences given by the judge to the k items. When some items receive the same preference, then a tied ranking or a weak ordering is obtained. In real situations, many times it happens that not all items are ranked: partial rankings, when judges are asked to rank only a subset of the whole set of items, and incomplete rankings, when judges can freely choose to rank only some items (a partial/incomplete ranking can be easily converted to a weak ordering by assigning a common-last position to unranked items). In order to get homogeneous groups of subjects having similar preferences, it is natural to measure the spread between rankings through dissimilarity or distance measures among them. Such a measure has to meet the natural properties of a distance function:

- Reflexivity: $d(a, a) = 0$,
- Positivity: $d(a, b) > 0$ if $a \neq b$,
- Symmetry: $d(a, b) = d(b, a)$,
- Triangle inequality: $d(a, b) \leq d(a, c) + d(c, b)$ (in case of a distance).

Moreover, a desirable property of any distance is its invariance toward a renumbering of the elements: the so-called label invariance or equivariance [5].

Within the metrics proposed in the literature to compute distances between rankings, the Kemeny distance will be here considered [15]. The Kemeny distance (d_K) between two rankings π and σ is a city-block distance defined as:

$$d_K(\pi, \sigma) = \frac{1}{2} \sum_{r=1}^{k} \sum_{s=1}^{k} |a_{rs} - b_{rs}| \tag{1}$$

where a_{rs} and b_{rs} are the generic elements of the $k \times k$ score matrices associated to π and σ respectively, assuming value equal to 1 if the item r is preferred to or tied with the item s, –1 if the item s is preferred to the item r and 0 if $r = s$.

Considering the usual relation between a distance d and its corresponding correlation coefficient $\tau = 1 - 2d/D_{max}$, where D_{max} is the maximum distance [11], d_K is in a one-to-one correspondence to the rank correlation coefficient τ_x proposed by Emond and Mason [11], defined as:

$$\tau_x(\pi, \sigma) = \frac{\sum_{r=1}^{k} \sum_{s=1}^{k} a'_{rs} b'_{rs}}{k(k-1)}. \quad (2)$$

where a'_{rs} and b'_{rs} are the generic elements of the $k \times k$ score matrices associated to a and b respectively, assuming now value equal to 1 if the item r is preferred to or tied with the item s, –1 if the s-th item is preferred to r-th item and 0 if $r = s$.

2.1 Weighted Distances

Kumar and Vassilvitskii (2010) [16] introduced two aspects essential for many applications involving distances between rankings: positional weights and element weights. In short, (i) the importance given to swapping elements near the head of a ranking could be higher than the same attributed to elements belonging to the tail of the list or (ii) swapping elements similar between themselves should be less penalized than swapping elements which are not similar.

As Henzgen and Hüllermeier [13] say, weighted versions of rank correlation measures have been studied in many fields other than statistics. For example, "in information retrieval, important documents are supposed to appear in the top, and a swap of important documents should incur a higher penalty than a swap of unimportant ones". In the context of the web, comparing the query results from the different search engines, the distance should emphasize the difference of the top elements more than the bottom ones, since people may be interested in the first few items [6]. A short review of the solutions proposed in the literature to cope with this issue can be found in [26].

In this paper, we deal with the case (i) and consider the weighted version of the Kemeny metric. For measuring the weighted distances, the non-increasing weights vector $w = (w_1, w_2, \ldots, w_{k-1})$ constrained to $\sum_{p=1}^{k-1} w_p = 1$ is used, where w_p is the weight given to position p in the ranking.

Given two generic rankings of k elements, π and σ, the Weighted Kemeny distance was provided by [12] as follows:

$$d_K^w(\sigma, \pi) = \frac{1}{2} \left[\sum_{\substack{r,s=1 \\ r<s}}^{k} w_r |a_{rs}^{(\sigma)} - b_{rs}^{(\sigma)}| + \sum_{\substack{r,s=1 \\ r<s}}^{k} w_r |b_{rs}^{(\pi)} - a_{rs}^{(\pi)}| \right], \quad (3)$$

where (σ) states to follow the σ ranking and (π), similarly, orders according to π. More specifically, $b_{rs}^{(\sigma)}$ is the score matrix of the ranking π reordered according to σ, $a_{rs}^{(\pi)}$ is the score matrix of the ranking σ reordered according to π and $a_{rs}^{(\sigma)} = b_{rs}^{(\pi)}$ is the score matrix of the linear order $1, 2, \ldots, k$ (see [20] for more details).

Let's see a small example that shows how to compute the weighted Kemeny distance. Given the ranking vectors $\sigma = (3, 1, 2, 4)$ and $\pi = (1, 4, 2, 3)$, we can easily define their orderings: $(\sigma) = (2, 3, 1, 4)$ and $(\pi) = (1, 3, 4, 2)$. Then we have:

$\sigma^{(\sigma)} = (1, 2, 3, 4)$, $\sigma^{(\pi)} = (3, 2, 4, 1)$, $\pi^{(\sigma)} = (4, 2, 1, 3)$ and $\pi^{(\pi)} = (1, 2, 3, 4)$. The score matrices of $\sigma^{(\sigma)}$, $\pi^{(\sigma)}$, $\sigma^{(\pi)}$ and $\pi^{(\pi)}$ are respectively:

$$\sigma^{(\sigma)} = \begin{bmatrix} 0 & +1 & +1 & +1 \\ & 0 & +1 & +1 \\ & & 0 & +1 \\ & & & 0 \end{bmatrix} \quad \pi^{(\sigma)} = \begin{bmatrix} 0 & -1 & +1 & -1 \\ & 0 & +1 & +1 \\ & & 0 & +1 \\ & & & 0 \end{bmatrix}$$

$$\pi^{(\pi)} = \begin{bmatrix} 0 & +1 & +1 & +1 \\ & 0 & +1 & +1 \\ & & 0 & +1 \\ & & & 0 \end{bmatrix} \quad \sigma^{(\pi)} = \begin{bmatrix} 0 & -1 & -1 & -1 \\ & 0 & -1 & +1 \\ & & 0 & +1 \\ & & & 0 \end{bmatrix}$$

Through simple algebric operations it can be easily found that:

$$|\sigma^{(\sigma)} - \pi^{(\sigma)}| = \begin{bmatrix} 0 & 2 & 0 & 2 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix} \quad |\pi^{(\pi)} - \sigma^{(\pi)}| = \begin{bmatrix} 0 & 2 & 2 & 2 \\ & 0 & 2 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix}$$

Now we can compute the weighted Kemeny distance. Let's assume equal weights, $w = (1/3, 1/3, 1/3)$, the distance between a and b is:

$$d_K^w(\sigma, \pi) = \frac{1}{2}\left[(2+2+2+2+2)\frac{1}{3} + (2)\frac{1}{3} + (0)\frac{1}{3}\right] = \frac{1}{2}\left(\frac{12}{3}\right) = 2$$

If we change the vector of weights, with a decreasing structure $w = (2/3, 1/3, 0)$, the value of the distance is:

$$d_K^w(\sigma, \pi) = \frac{1}{2}\left[(2+2+2+2+2)\frac{2}{3} + (2)\frac{1}{3} + (0)0\right] = \frac{1}{2}\left(\frac{22}{3}\right) = \frac{11}{3}.$$

3 A Suitable Rank Correlation Coefficient

In this paper we will investigate the behavior of the consensus measure suitable for position weighted rankings proposed in [21] in the frame of decision trees. It represents an extension of τ_x proposed by Emond and Mason [11] that handles linear rankings when the position occupied by the items is relevant. It is defined as:

$$\tau_x^w(a, b) = \frac{\sum_{r<s}^{k} a_{rs}^{(\sigma_1)} b_{rs}^{(\sigma_1)} w_r + \sum_{r<s}^{k} a_{rs}^{(\sigma_2)} b_{rs}^{(\sigma_2)} w_r}{Max[d_K^w(a, b)]}, \tag{4}$$

where the denominator represents the maximum value for the Kemeny weighted distances [12], equal to:

$$Max[d_K^w(a,b)] = 2\sum_{r=1}^{k-1}(k-r)w_r. \tag{5}$$

Note that the constraint $\sum_{i=1}^{k-1} w_r = 1$ is no more necessary. For linear orderings (i.e. without ties), a^{σ_1} and b^{σ_2} represent the natural ascending orderings. Let $a^{\sigma_1} = b^{\sigma_2} = o$, with $o = 1, 2, \ldots, k$, the new rank correlation coefficient reduces to:

$$\tau_x^w(a,b) = \frac{\sum_{r<s} o_{rs}(b_{rs}^{(\sigma 1)} + a_{rs}^{(\sigma 2)})w_r}{Max[d_K^w(a,b)]},$$

where the score matrix O fo the ranking o is:

$$A_{rs} = \begin{bmatrix} 0 & 1 & 1 & \ldots & 1 & 1 \\ -1 & 0 & 1 & \ldots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & -1 & -1 & \ldots & -1 & 0 \end{bmatrix}$$

In [21] the following properties are analitycally proved:

- the correspondence between the weighted rank correlation coefficient and the weighted Kemeny distance holds, through the linear transformation $\tau = 1 - 2d/D_{max}$, whatever is the weighting vector assigned to the items' positions;
- the weighted rank correlation coefficient satisfies the main property of a generic correlation coefficient, i.e. it takes values between -1 and 1;
- when equal importance is assigned to the items' position, i.e. $w_r = \frac{1}{(k-1)}$ $\forall r = 1, 2, \ldots, k-1$, the weighted rank correlation coefficient is equivalent to the rank correlation coefficient defined by Emond & Mason (2002).

The weighted correlation coefficient can be used to deal with a consensus ranking problem: given a n x k matrix $\mathbf{X}$, whose generic s-th row represents the ranking associated to the s-th judge, the purpose is to identify that ranking (b) (a candidate within the universe of the permutations of k elements) that best represents the average consensus of the subjects involved (i.e. the matrix $\mathbf{X}$). Considering that there is a one-to-one correspondence between a rank correlation coefficient and a distance, the solution ranking is reached by minimizing the average distance or, similarly, maximizing the average rank correlation:

$$\sum_{s=1}^{n} d(\mathbf{X}, b) = min,$$

$$\sum_{s=1}^{n} \tau_x^w(\mathbf{X}, b) = max. \tag{6}$$

4 Decision Trees for Preference Data

The weighted Kemeny distance measure, introduced by García-Lapresta and Pérez-Román [12] and reported in Eq. (3), together with the corresponding correlation coefficient (4) proposed by Plaia et al. [21] represent the main ingredient for the definition of our proposal, which consists in extending classical splitting criteria, used in decision tree modelling, in order to include other measures that can be used to evaluate impurity when faced with preference data.

4.1 Splitting Criterion and Impurity Function for Preference Data

The main problem in extending univariate regression trees [4] to the multivariate response case deals with generalizing the definition of the partitioning metric. In order to avoid the problems linked to the multivariate nature of the ranking data, in this work the vectors of preferences will be treated as a unique multivariate entity. Specifically, when the k items are completely ranked, the ranking process can be seen as a permutation function from $\{1, \ldots, k\}$ onto $\{1, \ldots, k\}$. Hence, if a label is assigned to each permutation defined from a set of distinct items, each ranking vector can be identified through a label that, used as a response variable, allows to apply the classical univariate classification tree methodologies. The recursive binary partitioning process used by the CART methodology, starting from the root node, produces a nested sequence of subtrees

$$T_k = \{root - node\} \subset \cdots \subset T_0 = \{full - tree\}$$

obtained, at each step, through a splitting criterion which works with the maximization of $\Delta i(s, t)$, i.e.: the reduction in the impurity resulting from the split s in node t. In a line:

$$\Delta i(s, t) = i(t) - p_L i(t_L) - p_R i(t_R)$$

where p_L and p_R are the proportions of units in node t falling in the left child node t_L and in the right child node t_R, respectively, at the s-th split.

When the impurity has to be evaluated working with ranking data, the impurity function can be properly modified as follows:

$$i(t) = \sum d(a, b), \tag{7}$$

where d is the chosen distance measure between the orderings a and b and the summation is extended to all the couples of rankings in the node t. A decrease in node impurity at each step will be evaluated according to all covariates and respective split points. Piccarreta [19] proposes an impurity function based on dissimilarities, without proposing specific distances for preference data. To allow the impurity function to take into account the ordinal nature of rankings, we propose to use the weighted Kemeny distance (Eq. 3) in Eq. (7).

According to the CART methodology [4], tree optimization implies cutting off insignificant nodes through a standard pruning process based on some cost-complexity criteria that will lead to obtaining the optimum tree size.

4.2 Rank Aggregation in the Leaves

Once the tree has grown, the definition of a consensus ranking is necessary for assigning a class label or class ranking to each node, i.e. the ranking in the best agreement with all the rankings in the node [8]. Among the several consensus ranking measures proposed in the literature, the median ranking approach will be used in the examples presented in the next section. The median is defined as the ranking corresponding to the minimum sum of the distances of all rankings from it or, equivalently, it is the ranking that maximizes the average τ_x^w rank correlation coefficient between itself and the other rankings belonging to the set of rankings (see Eqs. (4) and (6)) [9]. To compute the median ranking, we use the adapted versions of the Branch and Bound algorithm, proposed by Plaia et al. [21] for positional weighted ranking data.

The ranking having the highest correlation with the entire set of the n rankings in a leaf is the consensus ranking, i.e. the most representative ordering among all possible solutions, and therefore the "class label" to assign to the leaf.

R language [22] has been used for the implementation of the decision trees above described; in particular, a proper splitting function has been written as a user-defined splitting function in the "rpart" package [24].

5 Simulation Study

We are interested in evaluating the effect both on the splits and on the leaf labels of different weighting vectors w. For this reason, following D'Ambrosio and Heiser [9], we considered a predictor space (X_1 and X_2), with $X_1 \sim U(0, 10)$ and $X_2 \sim U(0, 6)$, which was partitioned into five regions as shown in Fig. 1. The number of datapoints (of 4 items) within each sub-partition was determined by: (i) randomly drawing from a normal distribution $N \sim (10, 2)$, (ii) dividing them by their summation, (iii) multiplying by the true sample size ($N = 200$, $N = 500$). The rankings (datapoints) within each sub-partition were generated from a Mallows Model [18], which is one of the first probability models proposed for rankings and frequently used in

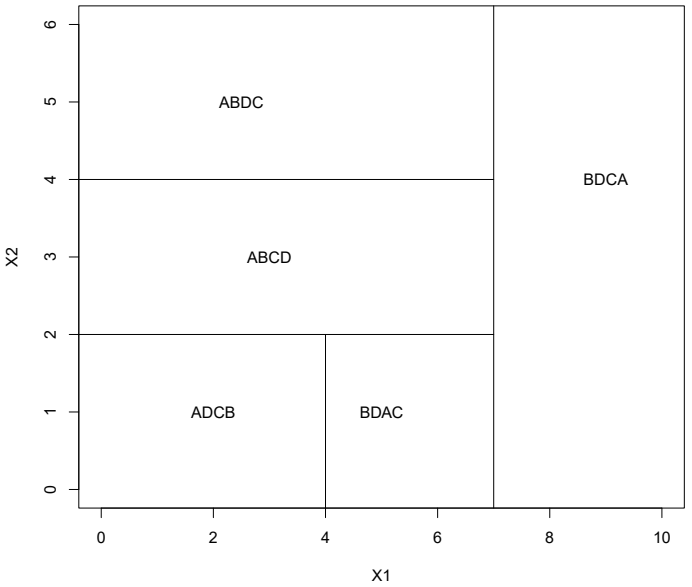


Fig. 1 Theoretical partition of the predictor space (X1, X2) with 4 items

both theoretical and applied studies. It is an exponential model defined by a central permutation σ_0 and a dispersion parameter θ. When $\theta > 0$, σ_0 represents the mode of the distribution, i.e., the permutation that has the highest probability. The σ_0 values for our simulation studies are shown in Fig. 1. The probability of any other ranking decays exponentially with increasing distance to the central permutation. The dispersion parameter controls the steepness of this decline. Assuming that σ is a generic ranking, the probability for this ranking is given by

$$\Pr(\sigma) = \frac{\exp(-\theta d(\sigma\sigma_0^{-1}))}{\psi(\theta)}, \tag{8}$$

where d is a ranking distance measure and $\psi(\theta)$ is a normalization constant.

Considering two levels for the sample size (100 and 300) and three different levels of noise (low with $\theta = 50$, medium with $\theta = 2$ and high with $\theta = 1$), the experimental design counts $2 \times 3 = 6$ different scenarios. For each scenario 100 datasets are generated. Figure 2 shows one of the 6 datasets considered in the simulation study, obtained with $\theta = 50$ and $n = 300$.

For each dataset five different weighting vectors are considered:
$w_1 = (1/3, 1/3, 1/3)$, $w_2 = (3/6, 2/6, 1/6)$, $w_3 = (1/2, 1/2, 0)$, $w_4 = (2/3, 1/3, 0)$ and $w_5 = (1, 0, 0)$.

With reference to the data in Figs. 2, 3 reports two of the five trees obtained (for one of the 100 datasets corresponding to this scenario): in particular, the plot on the left shows the tree corresponding to w_1, which perfectly recreates the original

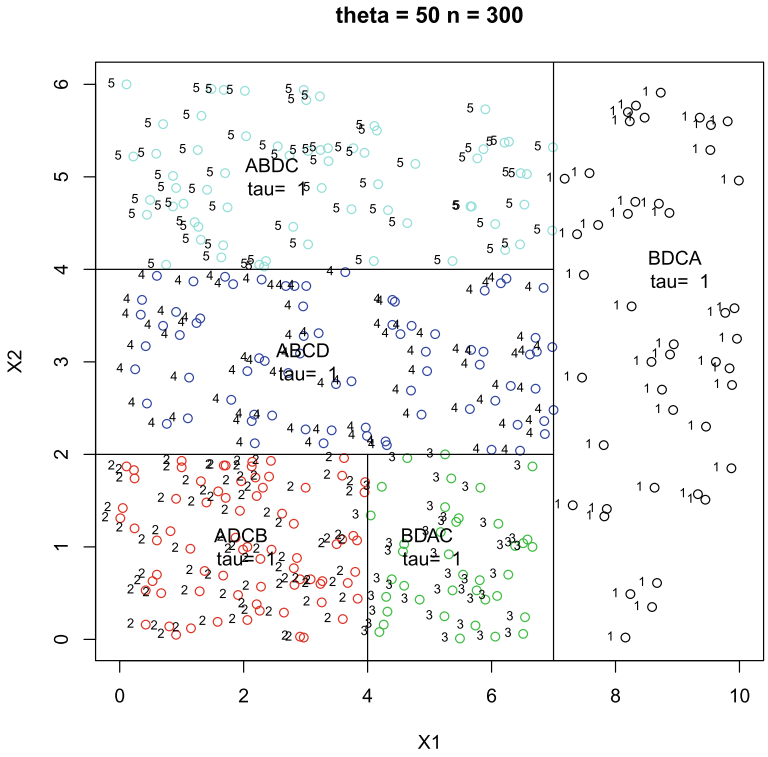


Fig. 2 Generation of homogeneous groups of ranking from the theoretical partition with: $X1 \sim U(0, 10)$ and $X2 \sim U(0, 6)$, $\theta = 50$ and $n = 300$

partition of the predictor space (Fig. 2); the plot on the right corresponds to w_3 and, as expected, does not perform the two splits $X \lessgtr 4$ and $X \lessgtr 7$ (the couples of rankings below each of the splits in the left tree do not differ for the first two positions). In fact, $w_3 = (1/2, 1/2, 0)$ means that we are not interested in positions 3 and 4, i.e. the τ_x^w between two rankings that differ only for the positions 3 and 4 is maximum $(= 1)$.

For each dataset generated with $n = 300$ and $\theta = (1, 2, 50)$, τ_x^w was measured at the root of the tree (white boxplots) and after applying the decision trees (average τ_x^w in the leaves - red boxplots) with all five weighting vectors, and the distributions obtained over the 100 datasets are shown in Figs. 4, 5 and 6. The higher the homogeneity among rankings, the better the ability of the decision trees to detect groups of similar rankings according to different values of $X1$ and $X2$. When rankings have an initial high homogeneity ($\theta = 50$, see Fig. 6) the trees show an almost perfect performance for all kinds of position weights, except for the case of w_5: assigning a weight only to the first position penalizes the performance of the trees. This penalization reduces as the homogeneity tends to decrease ($\theta = 1$, see Fig. 4).

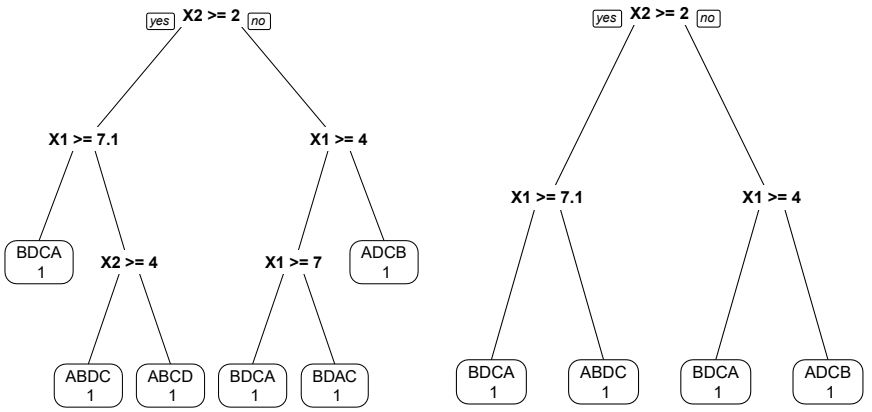


Fig. 3 Decision tree models for weighted rankings with weights vector w_1 (left) and w_3 (right)

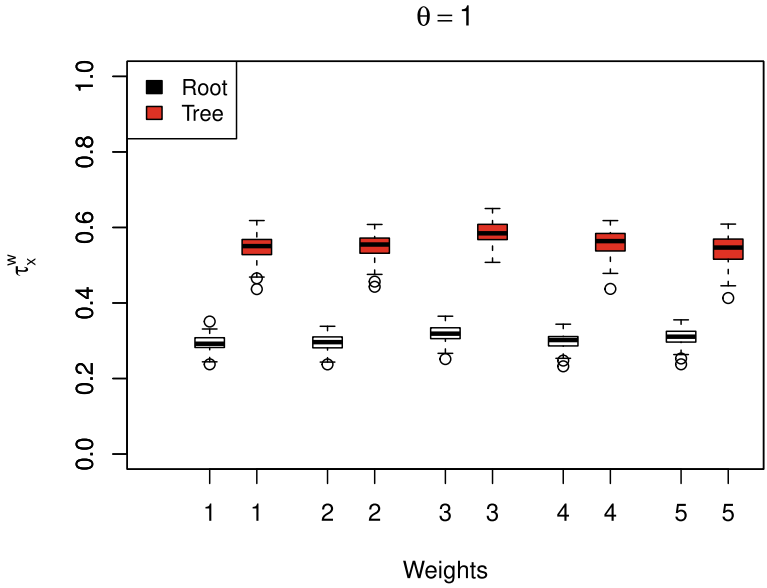


Fig. 4 Distribution of τ_x^w over 100 datasets both in the root and overall tree ($\theta = 1$ and $n = 300$)

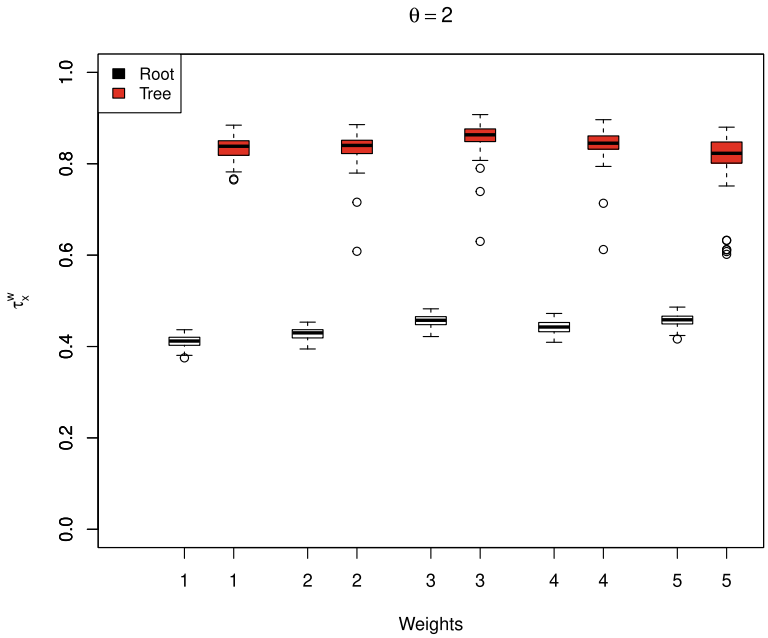


Fig. 5 Distribution of τ_x^w over 100 datasets both in the root and overall tree ($\theta = 2$ and $n = 300$)

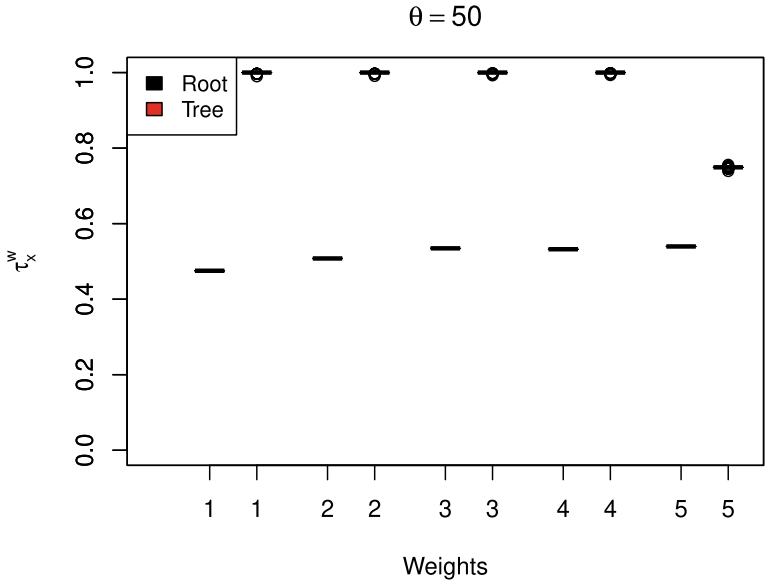


Fig. 6 Distribution of τ_x^w over 100 datasets both in the root and overall tree ($\theta = 50$ and $n = 300$)

6 Conclusion

In this paper, we have focused on distance-based decision trees for ranking data, when the positions occupied by the items are relevant. We have proposed the weighted Kemeny distance as impurity function and the relative proper weighted correlation coefficient τ_x in order to achieve the consensus measure in the terminal nodes. The kemeny distances is considered because, as it is known, it is the only distance satisfying the Kemeny's axioms. The methodology found to be capable of identifying correctly homogeneous groups of rankings when more than one position is taken into account, i.e. the profiles of subjects having similar preferences falling in the same leaf of the tree. Through some simulations we evaluated how splitting rules could be affected by the amount of homogeneity in the groups; we showed that the higher the homogeneity among rankings, the better the ability of the decision trees to detect groups of similar rankings. Particular attention has been addressed to the weighting systems and their effects on splitting and on consensus ranking. The implementation of a faster algorithm for the rpart package [24] could lead to work faster in the presence of large number of items. Further developments could be, hence, a replication of the same analyses with an increasing number of items.

References

1. D'Ambrosio, A.: Tree based methods for data editing and preference rankings. Ph.D. thesis, Universitá degli Studi di Napoli "Federico II" (2007)
2. D'Ambrosio, A., Heiser, W.J.: A recursive partitioning method for the prediction of preference rankings based upon Kemeny distances. Psychometrika **81**(3), 774–794 (2016)
3. Cook, W.D.: Distance based and ad hoc consensus models in ordinal preference ranking. Eur. J. Oper. Res. **172**, 369–385 (2006)
4. Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J.: Classification and regression trees. Chapman & Hall (1984)
5. Cheng, W., and Hühn, J., and Hüllermeier, E.: Decision tree and instance-based learning for label ranking. In: Proceedings of the 26th Annual International Conference on Machine Learning, pp. 161–168 (2009)
6. Chen, J., Li, Y., Feng, L.: On the equivalence of weighted metrics for distance measures between two rankings. J. Inf. Comput. Sci. **11**(13), 4477–4485 (2014)
7. D'Ambrosio, A., Amodio, S., Mazzeo, G.: ConsRank: compute the median ranking(s) according to the Kemeny's axiomatic approach. R package version 2.0.1 (2017). https://CRAN.R-project.org/package=ConsRank
8. D'ambrosio, A., Amodio, S., Iorio, C.: Two algorithms for finding optimal solutions of the Kemeny rank aggregation problem for full rankings. Electron. J. Appl. Stat. Anal. **8**(2), 198–213 (2015)
9. D'Ambrosio, A., Heiser, W.J.: A recursive partitioning method for the prediction of preference rankings based upon Kemeny distances. Psychometrika **81**(3), 774–794 (2016)
10. Dittrich, R., Hatzinger, R., Katzenbeisser, W.: Modelling the effect of subject-specific covariates in paired comparison studies with an application to university rankings. J. R. Stat. Soc. C (Appl. Stat.) **47**(4), 511–525 (1998)
11. Emond, E.J., Mason, D.W.: A new rank correlation coefficient with application to the concensus ranking problem. J. Multi-Criteria Decis. Anal. **11**, 17–28 (2002)

12. García-Lapresta, J.L., Pérez-Román, D.: Consensus measures generated by weighted Kemeny distances on weak orders. In: Proceedings of the 10th International Conference on Intelligent Systems Design and Applications, Cairo (2010)
13. Henzgen, S., Hüllermeier, E.: Weighted rank correlation: a flexible approach based on fuzzy order relations. Jt. Eur. Conf. Mach. Learn. Knowl. Discov. Databases 422–437 (2015)
14. Irurozki, E., Calvo, B., Lozano, J.A.: PerMallows: an R package for mallows and generalized mallows models. J. Stat. Softw. **71**(12), 1–30 (2016). https://doi.org/10.18637/jss.v071.i12
15. Kemeny, J.G., Snell, J.L.: Preference rankings an axiomatic approach. MIT Press (1962)
16. Kumar, R., Vassilvitskii, S.: Generalized Distances Between Rankings. In: Proceedings of the 19th International Conference on World Wide Web, WWW '10, pp. 571–580. ACM, New York, NY, USA (2010)
17. Lee, P.H., Yu, P.L.H.: Distance-based tree models for ranking data. Comput. Stat. Data Anal. **54**(6), 1672–1682 (2010)
18. Mallows, C.L.: Non-null ranking models. Biometrika **44**(1–2), 114–130 (1957)
19. Piccarreta, R.: Binary trees for dissimilarity data. Comput. Stat. Data Anal. **54**(6), 1516–1524 (2010)
20. Plaia, A., Sciandra, M.: Weighted distance-based trees for ranking data. In: Advances in Data Analysis and Classification, pp. 1–18. Springer, Berlin (2017) https://doi.org/10.1007/s11634-017-0306-x
21. Plaia, A., Buscemi, S., Sciandra, M.: A new position weight correlation coefficient for consensus ranking process without ties. Stat **8**, e236 (2019). https://doi.org/10.1002/sta4.236
22. R Core Team, R.: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2018). https://www.R-project.org/
23. Strobl, C., Wickelmaier, F., Zeileis, A.: Accounting for individual differences in Bradley-Terry models by means of recursive partitioning. J. Educ. Behav. Stat. **36**(2), 135–153 (2011)
24. Therneau, T., Clinic, M.: User written splitting functions for RPART (2015)
25. Therneau, T., Atkinson, B., Ripley, B.: rpart: recursive partitioning and regression trees. R package version 4.1–10 (2015)
26. Yilmaz, E., Aslam, J.A., Robertson, S.: A new rank correlation coefficient for information retrieval. In: Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 587–594 (2008)

European Funds and Regional Convergence: From the European Context to the Italian Scenario

Gennaro Punzo, Mariateresa Ciommi, and Gaetano Musella

Abstract The inclusive economic growth and the territorial cohesion represent the central points of the EU agenda. The regional funds are the main European Regional Policy aimed to increase the employment levels in the Union and to reduce the territorial divides between the backward and forward regions. This work aims to verify the effectiveness of 2007–2013 EU funding by means of the Difference-in-Differences regression on official data referring to European NUTS-2 regions. The aim is twofold. First, we verify whether the regional funds narrowed employment disparities in the European context. Second, we focus on Italy as one of the largest beneficiary countries of funds. The results suggest a general ineffectiveness of funds across the European countries and an even worst scenario in Italy. The quality of institutions, the fund management by the national and regional governments and the monitoring activities are the main causes of failure of regional policies in Italy.

Keywords EU funding · Territorial convergence · Employment · Diff-in-Diff · Italy

G. Punzo (✉)
Department of Economic and Legal Studies, University of Naples "Parthenope", via Generale Parisi, 13–80132 Napoli, Italy
e-mail: gennaro.punzo@uniparthenope.it

M. Ciommi
Department of Economic and Social Sciences, Università Politecnica delle Marche, Piazzale Martelli, 8–60121 Ancona, Italy
e-mail: m.ciommi@univpm.it

G. Musella
Department of Management and Quantitative Studies, University of Naples "Parthenope", via Generale Parisi, 13–80132 Napoli, Italy
e-mail: gaetano.musella@uiparthenope.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_16

1 Introduction

Since its foundation, the European Union (EU) has pursued the goal of reducing socio-economic disparities between its territories (Art. 158 of the founding Treaty, art. 130a of the Maastricht Treaty). Regional policy finds its origins in the Treaty of Rome founding the European Economic Community in 1957. To achieve the aim of territorial convergence, the EU Parliament promotes policies to support the poorest regions. In this context, the European Structural and Investment Funds (ESIF, hereafter) are the leading instruments to help the least developed regions.

To support economic development, the European Commission has set up five main Founds: (1) the European Regional Development Fund (ERDF); (2) the European Social Fund (ESF); (3) the Cohesion Fund (CF); (4) the European Agricultural Fund for Rural Development (EAFRD), and (5) the European Maritime and Fisheries Fund (EMFF).[1] During the last programming cycles (2000–2006, 2007–2013, 2014–2020), ERDF, ESF and CF funnelled most of the ESIFs, with the aim of leading the less well-off regions to a real convergence within countries (European [22] and making the EU's system of market integration viable [10]. While every EU region may benefit from the ERDF and ESF, the CF is addressed to the least developed regions.

The first budget allocated for the period 1988–1992 was of ECU 64 billion and it increased in each subsequent period: ECU 168 billion for 1994–1999; €235 billion (€213 for the 15 old member countries and the remaining for the new members) for 2000–2006; €347 billion for both 2007–2013 and 2014–2020.[2] Thus, regional policy contributes to succeed EU policy objectives and integrates them in the fields of education, employment, energy, environment, single market, research and innovation.

With particular reference to employment, territorial convergence consists in lowering disparities in employment rates between regions through a spiral of economic growth. In the current programming period (2014–2020), a new legislative framework for these funds has been brought forward and a new set of rules clearly links the ESIFs with the EU 2020 Strategy for smart, sustainable and inclusive growth in the European Union [22]. The EU 2020 Strategy has five targets: (1) to reach the employment of at least the 75% of the 20–64 year-olds EU-citizens (Employment),(2) to dedicate the 3% of the EU's GDP for R&D (Research & Development); (3) to reduce the greenhouse gas emissions, improve the energy from renewables and increase of energy efficiency (Climate change and energy sustainability); (4) to decrease the rates of early school leavers (Education); (5) to reduce drastically the number of people in or at risk of poverty and social exclusion (Fighting poverty and social exclusion).

At least three of the five headline goals of the EU 2020 Strategy relate directly to employment and productivity, with a focus on the target of 'new skills for new jobs'. In particular, within the Strategy's objective of sustainable growth, EU 2020

[1] See https://ec.europa.eu/regional_policy/en/funding/ for more details.

[2] See http://ec.europa.eu/regional_policy/index.cfm/en/ for more details.

fosters a high-employment economy that delivers social and territorial cohesion. In this respect, within the seven flagship initiatives, the Commission puts forward 'an agenda for new skills and jobs' [21] to empower people by developing skills throughout the lifecycle with a view to better match labour supply and demand.

The aim of this paper is to investigate the impact of the EU funding on employment levels at the regional level (NUTS-2)[3] within some EU countries, looking at the period 2007–2013. In particular, the paper first evaluates the overall contribution of regional funds in achieving the objective of territorial convergence in employment rates for ten European countries, paving the way for the first headline goal of the EU 2020 Strategy; Then, it focuses on Italy as a specific case-study by controlling for a set of socio-economic and institutional characteristics at the regional level with the aim of exploring the main reasons behind the controversial impact of EU funds on employment rates and their regional convergence.

For 2007–2013, the EU authorities set a threshold for regional GDP to allow NUTS-2 regions to apply for the 'convergence objective' fund, i.e., their per capita GDP should be less than 75% of the EU average GDP. Because of the enlargement to EU-25, a phasing-out support was assured to NUTS-2 regions that were below the threshold of 75% of GDP for the EU-15 but above the same threshold for the EU-25.[4]

Based on the Council Regulations (No 1080/2006 and 1083/2006), the convergence objective concerned 84 regions within 18 Member States of EU-27 with a total of 154 million inhabitants, and on the "phasing-out" basis, other 16 regions within 8 countries with a population of 16.4 million. The amount allocated for ESIFs was assigned to the convergence regions (70.47%) and "phasing-out" regions (4.95%), while the remaining part concerned the Cohesion Fund (24.58%).

After Poland and Spain, Italy is the third largest beneficiary of the EU Cohesion Policy, receiving almost €29 billion from the ERDFs and ESFs. In total, 66 projects have been approved for Italy: 19 programmes under the Convergence objective, 33 programmes under the Regional Competitiveness and Employment objective, and 14 programmes under the European Territorial Cooperation Objective. However, the economic crisis, which characterised the period 2007–2013, affected Italy more than other EU members, accentuating the weaknesses of the Country. The 2014 GDP was

[3] Acronym of Nomenclature of Territorial Units for Statistics, NUTS is a hierarchical classification proposed by Eurostat (Commission Regulation no. 31/2011) that divides up the EU economic territory for the purpose of harmonisation of regional statistics (see https://ec.europa.eu/eurostat/web/nuts/background). The NUTS-2 level is one of the most important territorial partition in that it represents the basic regions for the application of regional policies. In Italy, the NUTS-2 level corresponds to the 19 administrative regions plus the 2 autonomous provinces of Trento and Bolzano.

[4] According to the GPD criterion, the potential beneficiaries of the funds were some NUTS-2 regions belonging to 10 EU countries – namely Czech Republic, Germany, Greece, Spain, France, Hungary, Italy, Portugal, Slovakia, and the United Kingdom – and the whole territory of 8 EU countries, namely Bulgaria, Estonia, Latvia, Lithuania, Malta, Poland, Romania, and Slovenia. To those regions, additional 16 eligible regions belonging to Belgium, Germany, Greece, Spain, Italy, Austria, Portugal and United Kingdom were covered by the Convergence objective in the EU with 15 members, but not the EU with 25. See https://ec.europa.eu/regional_policy/en/policy/how/is-my-region-covered/2007-2013/ for the complete list of NUTS-2 regions.

9% lower than at the beginning of the programming period, while the unemployment rate increased by about 6 percentage points. If we focus only on ERDFs, in the above mentioned period, the subsidy was equivalent of around 0.2% of GDP and 4.4% of Government capital expenditure.

From the methodological perspective, the role of European funds in the regional convergence process is evaluated through the Difference-in-Differences (DiD) tools. The DiD method allows the estimation of the effect of a specific treatment by comparing the change in outcomes occurred over time between a population enrolled in a program, the so-called 'intervention group', and a population that is not, namely 'control group'. Labour economists were the first to apply DiD tools. For instance, Ashenfelter [6] investigated the effect of training programs on earnings. Afterwards, [13] analysed the impact of a New Jersey rise in the minimum wage on employment in fast-food restaurants.

The paper is structured as follows. Section 2 reviews the more relevant literature in this field. Section 3 introduces the DiD methodology and describes data used for the analysis. Section 4 discusses the results of the DiD null models for the EU countries. Section 5 estimates the DiD model for Italian regions by controlling for a set of covariates. Finally, Sect. 6 is devoted to the discussion and conclusions.

2 Related Works

Several works concerned about the analysis of the relationship between the structural funds of the previous programming cycles and the economic convergence at the national and regional levels in Europe. It is clear that the progress made in a period lays the basis for the time after. Studies usually showed conflicting results, which often also depend on the empirical strategy used [5]. Hagen and Mohl [27] reviewed the econometric methods for evaluating EU Cohesion Policy and discussed the reasons for the divergent results that emerged. Similarly, Dall'Erba and Fang [16] performed a meta-analysis on the econometric approaches used.

A critical review of the main results from the dominant literature on the impact of structural funds on economic regional growth has been proposed by Mohl and Hagen [35]. Dall'Erba [15] found a positive effect of structural funds on economic growth for 145 European regions over the long period 1989–1999 by analysing the relationship between the spatial distribution of per capita GDP and the regional funds.

By exploring the EU annual regional payments and commitments dataset for the period 2000–2006, Mohl and Hagen [35] proved that the effectiveness of structural funds in fostering economic growth depends on the type of objective. However, this is not the only factor influencing the incisiveness of structural funds. Percoco [42] found in the economic structure of the regions the main determinant of strength of the impact of Cohesion Policy. By comparing Romania and Bulgaria, Surubaru [49] attributed success in the use of funds to the administrative capacity (i.e., institutional capacity, bureaucratic capacity, human resources capacity, political stability, political support and political clientelism) of the regions. Analogously, Gagliardi and Percoco [24]

ascribed the different impact of Cohesion Policy to the geographical characteristics and structures of regions receiving the funds. Giordano [26] found similar results by considering three case studies (i.e., the province of Cuenca in the region of Castilla-La Mancha in Spain, the Swedish County of Norrbotten and the small Danish island of Bornholm), proving that geographical features not only do matter but also that ERDF policy responses are strongly conditioned by them. Bähr [7] already studied the effectiveness of the EU regional policy by evaluating the sub-national autonomy of EU countries. The results indicated that if the country had a higher degree of decentralisation the Structural Funds are more effective.

Maynou et al. [34] analysed the economic convergence of 174 regions belonging to 17 EU countries during the period 1990–2010, showing a reduction of disparities between countries directly attributable to Structural and Cohesion Funds. More in detail, a 1% increase in regional funds would lead to an average growth of 0.9% of per capita GDP in these countries, all other conditions being equal. Becker et al. [10] analysed the four programming periods (1989–1993, 1994–1999, 2000–2006, 2007–2013) and found that the economic crisis reduced the effect of transfers mainly on per capita income and this effect was greater in countries where the crisis hit the hardest.

According to Percoco [41], the economic literature analysing the role of Structural Funds accounts for two main branches. The first one focuses on the analysis of the impact of Structural Funds on the economic growth and, in particular, on its convergence. The second one aims at evaluating the impact of Structural Funds spending. Our paper refers to the first research line with the specific focus on convergence in employment levels in Italy. Pupo and Aiello [43] investigated the effectiveness of cohesion policies in Italy, focusing on the effects of EU spending on the convergence for the Italian regions from 1996 to 2007. The main conclusion they came to is that even if EU funds have been distributed to benefit the more backward regions, the way in which they have managed the resources vary across regions. Moreover, Structural Funds have had no impact on the dynamics of labour productivity. The same authors also examined the role of EU Cohesion Policy in Italy [2], proving that the impact of the funds was greater in the South compared to the Centre-North of Italy, although this has not contributed to reducing the differences in the productivity levels.

3 Method and Data

We adopt the DiD strategy [6] to evaluate the effect of the regional funds on local employment and their contribution to convergence across EU NUTS-2 regions. The DiD methodology allows evaluating the treatment effect by looking at the pre- and post-treatment differences in the outcome variable between two groups, one exposed to treatment and the other one used as a control group.

The DiD tools can be applied under the assumption that the trends in the treatment and control groups would have continued the same way as before in the absence of treatment. This is equivalent to affirm that the assumption of parallel paths is satisfied,

that is, in absence of treatment, the unobserved differences between treatment and control groups are the same over time. In other words, without treatment, the average change for the treated would have been equal to the observed average change for the controls (for details, see Mora and Reggio [36]).

The DiD tool has several strengths. First, it has an intuitive interpretation; second, it focuses on change rather than absolute levels. Third, it allows accounting for change due to factors other than intervention. On the other hand, the method does not come without limitations. It requires, for instance, baseline data as well as non-intervention group; thus, the use of the DiD technique requires the availability of data from pre- and post-treatment. In particular, the control group should be comparable to the treatment group and, more importantly, the two groups should have the same data availability. In its original formulation, the method cannot be applied if the groups in comparison have different outcome trends. Finally, the DiD tools require that pre/post change groups to be stable.

The DiD strategy allows us to compare the NUTS-2 regions that received the EU funding (treated group) between 2007 and 2013 with those NUTS-2 regions that did not (control group). By applying that tool, we can evaluate the differential effect of a treatment by comparing the change over time in the outcome for the treated group with the change that has occurred in the same timespan for the control group. Both groups of regions are analysed before (2000–2006) and after (2014–2016) the 2007–2013 funding period.

The DiD model is drawn as follows:

$$y = \alpha_0 + \alpha_1 S + \delta_0 T + \delta_1 T \bullet S + X\beta + u \tag{1}$$

where y is the *Nx1* vector of the outcome variable. S is a dummy variable that takes on value 1 for the recipient regions and 0 otherwise; α_1 accounts for the difference between the two groups prior to treatment; α_0 is the intercept. T is a dummy whose value is 1 for the period of treatment and 0 otherwise; δ_0 captures the time effects, and thus the changes in the outcome variable in the absence of the treatment. The interaction term ($T \bullet S$) is the same as a dummy equal to 1 for those treated units in the treatment period; δ_1 represents the treatment effect, this coefficient captures the difference in changes over time (diff-in-diff). X is the *NxK* matrix of covariates and β the vector of coefficients. Finally, $u \sim N(0, \sigma_u^2)$ is the *Nx1* vector of uncorrelated errors.

Formally, the coefficients are obtained by the following difference in differences:

$$\alpha_0 = E(y|S = 0, T = 0)$$

$$\alpha_1 = E(y|S = 1, T = 0) - E(y|S = 0, T = 0)$$

$$\delta_0 = E(y|S = 0, T = 1) - E(y|S = 0, T = 0)$$

$$\delta_1 = [E(y|S=0,T=0) - E(y|S=0,T=1)] + -[E(y|S=1,T=0) - E(y|S=1,T=1)] \quad (2)$$

where, $E(.|.)$ stands for conditional averages computed on the sample, as usually.

Let $\overline{y}_{S,T}$ be the average of the outcome variable in the group of regions $S = \{0,1\}$ at time $T = \{0,1\}$, the estimation of the treatment effect (diff-in-diff) is:

$$\hat{\delta}_1 = \left(\overline{y}_{0,0} - \overline{y}_{0,1}\right) - \left(\overline{y}_{1,0} - \overline{y}_{1,1}\right) = \widehat{\Delta}_0^{\overline{y}} + \widehat{\Delta}_1^{\overline{y}} \quad (3)$$

Thus, $\widehat{\Delta}_0^{\overline{y}}$ denotes the difference of the averages of the outcome variable *before* the EU funding between recipient and control regions, and $\widehat{\Delta}_1^{\overline{y}}$ is the difference of the averages *after* the EU funding between recipient and control regions.

We draw upon official data from Eurostat, Istat, and SIEPI (Italian Society of Economics and Industrial Policy). First, we perform the so-called DiD null model by country, using the Total Employment Rate (EMRT, hereafter, Tables 1 and 2). Second, regarding to Italy, we estimate DiD by controlling for a set of covariates, while keeping the EMRT as the outcome variable (Table 3).

In details, EMRT is the percentage of employed people aged 20–64 out of the working-age population. The EMRT data are related to NUTS-2 regions and cover a seventeen-year period, spanning from 2000 to 2016. In what follows, we describe the covariates. The 'Total Population by Educational Attainment Level' (PEALT)

Table 1 Diff-in-Diff estimates by country. 2000–2016

	Austria	Belgium	Czech republic	Germany	Spain
Before					
Control	70.8	66.29	77.26	69.8	67.38
Treated	70.84	57.6	70.04	66.57	58.58
Diff (T – C)	0.05 (1.12)	−8.69*** (1.79)	−7.17*** (1.18)	−3.24*** (0.59)	−8.81*** (0.87)
After					
Control	74.67	68.2	77.35	76.38	66.19
Treated	73.83	58.62	71.82	75.63	57.2
Diff (T – C)	−0.84 (0.94)	−9.58*** (1.5)	−5.53*** (0.99)	−0.73 (0.45)	−8.99*** (0.73)
Diff-in-Diff	−0.89 (1.46)	−0.89 (2.34)	1.65 (1.54)	2.51*** (0.74)	−0.18 (1.14)
Parallel Assumption	3.52	4.86	0.74	1.89	0.18
(*p-value*)	0.741	0.562	0.993	0.929	0.999

* *Significant at 10%; ** Significant at 5%; *** Significant at 1%; Standard errors in brackets*

Table 2 Diff-in-Diff estimates by country. 2000–2016

	Hungary	Italy	Portugal	Slovak republic	United Kingdom
Before					
Control	66.64	64.34	72.2	75.94	75.45
Treated	60.33	48.18	72.38	62.18	72.34
Diff (T – C)	−6.31*** (2.18)	−16.17*** (1.03)	0.18 (1.08)	−13.76*** (1.3)	−3.11*** (0.89)
After					
Control	68.33	66.71	68.58	76.1	75.74
Treated	61.87	47.1	69.37	65.01	74.31
Diff (T – C)	−6.46*** (1.82)	−19.61*** (0.85)	0.79 (0.9)	−11.09*** (1.08)	−1.42* (0.74)
Diff-in-Diff	−0.15 (2.84)	−3.45*** (1.33)	0.62 (1.4)	2.66 (1.69)	1.68 (1.16)
Parallel Assumption	1.51	0.41	0.5	6.94	3.21
(*p-value*)	0.959	0.998	0.984	0.325	0.781

* *Significant at 10%; ** Significant at 5%; *** Significant at 1%; Standard errors in brackets*

Table 3 Descriptive statistics for variables, Italy

	Min	Max	Mean	Standard deviation
EMRT	42.1	74.8	60.762	9.348
Per capita GDP	14,200	38,950	25,238.13	6,435.697
PEALT	6.7	23.3	13.651	3.348
ALMP	1,824.667	161,636.3	41,748.68	36,593.03
ELETT	6.7	34.3	18.24	5.779
PRPSE	0	471	32.614	44.024
Corruption	0.204	0.989	0.811	0.172
Government Effectiveness	0.064	0.721	0.362	0.147
Regulatory Quality	0.097	0.928	0.484	0.183
Rule of law	0.128	0.863	0.556	0.174
Voice and Accountability	0.148	0.721	0.433	0.144

represents the share of the total population with tertiary education (ISCED: 5–8).[5] The 'Active Labour Market Policies' (ALMP) depicts the average annual number

[5] The International Standard Classification of Education (ISCED) is the international reference proposed by UNESCO in the mid-1970s and then revised in 1997 and 2011 for a comparable classification of education programs and related qualifications by levels and fields. ISCED is based on education attainment, which refers to the highest level of formal education a person has successfully completed, from early childhood education (ISCED 0) to a short-cycle tertiary education, bachelor's, master's, doctoral or equivalent degree (ISCED: 5-8).

of beneficiaries of active policies. The 'per capita Gross Domestic Product' (per capita GDP) is the value of all goods and services produced in one year divided by the total population. The 'Total Early Leavers from Education and Training' (ELETT) captures the share of 18–24 year olds having attained ISCED 0–2 level and not receiving any formal or non-formal education or training. The variable 'People at-Risk-of-Poverty or Social Exclusion' (PRPSE) considers the percentage of the total population below either the poverty threshold (fixed as the 60% of the national median equivalised disposable income after social transfers) or severely materially deprived or living in a household with a very low work intensity.

The last variable collects five dimensions of the 'Institutional Quality Index' (IQI), namely, (*i*) Corruption, which measures the degree of corruption of those performing public functions and crimes against the public administration; (*ii*) Government effectiveness that evaluates the quality of public service and the policies of local governments; (*iii*) Regulatory quality, representing the ability of government to promote and formulate effective regulatory interventions; (*iv*) Rule of law that quantifies the crime levels, shadow economy, police force, and magistrate productivity; (*v*) Voice and accountability, assessing the degree of freedom of the press and association [37]. Each of the five dimensions of IQI ranges between 0 and 1, and the closer the indicator to 1, the higher the quality of local institution associated with that specific dimension of IQI. Table 3 shows basic descriptive statistics of the above-mentioned covariates for Italy.[6]

4 Results: The European Context

We test the effectiveness of the EU funding in fostering homogeneity in employment rates across 256 NUTS-2 regions within 10 EU countries. Once the treated group (84 regions exposed to convergence objective) and the control group (172 regions not included in the objective) are correctly identified, we perform the DiD null model by country, using the EMRT as the outcome variable.

The DiD null model exclusively evaluates the differential effects of EU funds on EMRTs between the treated and control regions. Subsequently, we compare the average employment rates across the two groups (treated and control) of regions before and after the provision of EU funds. Reasons of data availability on structural funds did not allow us to include France and Greece in the analysis.

The DiD methodology allows us to evaluate three main differences. First, the difference of the average EMRTs *before* the EU funding between treated and control regions; second, the difference of the average EMRTs *after* the EU funding between the same two groups of regions; third, the difference between the two previous

[6] Details on the variables used in this analysis can be found at the following sources: (*i*) ALMP variable, see ISTAT online dataset (http://dati.istat.it/); (*ii*) IQI variable, see SIEPI webpage (https://siepi.org/); (*iii*) other variables, see EUROSTAT online dataset (https://ec.europa.eu/eurostat/web/main/data/database). Moreover, the dataset used is available from the authors upon request.

differences (diff-in-diff, $\hat{\delta}_1$), which is expected to be significantly positive because only then it would represent a positive effect of EU funding on the convergence process for employment (Tables 1 and 2).

Before assessing the effect of EU funds on the convergence objective, the parallel-path assumption must be verified. The Mora & Reggio's approach [36] allowed us to test whether the average change in outcome for the treated in the absence of treatment equals the average change in outcome for the control. In our specific case, we verified that in the absence of the EU policy the averaged EMRTs of the two groups of regions follow the same trend.

The last row of Tables 1 and 2 shows that the parallel-trend assumption is met for each country covered. Thus, the group of control regions can be considered as a suitable counterfactual for the group of treated regions. This result is quite remarkable when one considers that the economic recession reached its highest intensity before the end of 2009 in most EU countries, that is, towards the middle of the 2007–2013 programming cycle, and that EU funds were not designed to offset the adverse effects of the crisis. Although some Member States, especially the countries of the Southern Europe, were particularly vulnerable, the parallel-trend test claims that the averages of employment rates of the treated regions would follow the same trend of those of the control regions even during the years of the crisis. Probably, the crisis had a quite pervasive impact on employment levels within each country without affecting the trends of employment rates in the less- and more-developed regions.

In general, the empirical evidence suggests that the EU funding did not influence significantly the convergence process for employment in most countries. In fact, except for Germany and Italy, the insignificance of $\hat{\delta}_1$ coefficients shows that, on average, the differences between the employment rates of the treated and control groups look much the same before and after the provision of EU funds or that any variation cannot be associated with the EU funds. However, some considerations can be made on specific groups of countries.

The first group of countries consists of Czech Republic, Slovak Republic and the United Kingdom, in which it could be assumed that the convergence process is still far from being realised both before and after the EU funding. In these countries, the employment gap between the two groups of regions becomes slightly lower after the treatment, but not enough low to make statistically significant the difference between the two differences after and before the provision of EU funds. In other words, the employment rates seem to improve in treated regions of these countries, narrowing the employment divide with the control regions but this reduction can hardly be associated with the EU funds.

The second group is composed of Belgium and Hungary where, between 2007 and 2013, the average employment levels even raised faster in the control regions than in the treated regions. This dynamic widened the gap between the two groups of regions even further to the detriment of the treated regions. However, also these differentials cannot be associated with the EU funds in that the diff-in-diff remains insignificant.

The third group is made up of Spain and Portugal, where the average employment levels have decreased for both groups of regions, widening in either case the territorial divergence within country. However, unlike Portugal, in Spain the employment divide between the treated and control regions retains its statistical significance both before and after the EU funding. In Austria, the employment divide between the treated and control regions was already not statistically significant before the treatment. Both groups of Austrian regions show an increase in employment rates, whose difference remains statistically insignificant even after the provision of EU funding.

Finally, Italy and Germany show opposite effects of the EU funding. In Germany, the employment divide between the treated and control regions is practically disappeared after the treatment, and the diff-in-diff is significantly positive, highlighting the effectiveness of EU funding in this country.

More importantly, the EU target of 75% of the population employed has been reached, on average, in both groups of German regions long before the date appointed by the EU 2020 Strategy. Instead, in Italy, the two groups of regions show an increased gap in average rates of employment such that the diff-in-diff estimate is even significantly negative. Italian regions exposed to convergence retain, on average, significant lower levels of employment than those of the control group: they still have a long way to go before converging in employment. This is one of the main reasons why we thought it appropriate to focus on Italy in order to investigate the potential reasons behind the failure of EU funds.

5 Results: The Italian Case

With the aim of exploring how much of the lack of regional convergence in Italy can be attributed to the incapacity of the Country and its recipient regions to generate economic growth and employment, we perform the DiD model for Italian regions (16 control and 5 treated) by controlling for a set of covariates. In fact, the impact of EU funds on economic development cannot avoid some institutional, political, and socio-economic factors. Precisely, four NUTS-2 regions, namely Calabria, Campania, Puglia and Sicilia, took advantage by the Convergence objective. Basilicata received the so-called "phasing out" assistance since it was covered by the Convergence objective in the EU with 15 members, but not the EU with 25.

As for the previous DiD null models, all data refer to the NUTS-2 regions and cover the period from 2000 to 2016. In addition, the EMRT is kept as the outcome variable, while several covariates are added in DiD models, as reported in Table 3.

The log–log specification of DiD model allows the coefficients (except for those of the IQI dimensions) to be interpreted as the percentage change in employment rates for a unit percentage change in a given covariate (i.e., elasticity of the dependent variable with respect to covariates).

In a preliminary step, we resorted to the instrumental variables (IV) method to overcome the endogenous relationships between the employment rates and per capita GDP and some specific dimensions of IQI. In detail, we used exogenous covariates

and their lagged versions as instruments. The Sargan-Hansen's J test [28, 45, 46] evaluated the validity of over-identifying restrictions, that is, the regression coefficients are exactly identified ($m = k$) or overidentified ($m > k$), where m and k denote the number of instruments and endogenous regressors, respectively. The Sargan-Hansen's J test lead to conclude with the exogeneity, and thus, the validity of these instruments.

Table 4 shows the estimation results of DiD model with covariates for Italy. First, the employment growth is positively related to the GDP level; hence, the relationship between GDP and employment appears not to have been affected during the period of global economic and financial crisis [13]. Per capita GDP can be interpreted as a proxy of the regional levels of economic prosperity, in that the differences in GDP help explain the differences in living standards between regions. The results suggest that economic growth provides an impetus to employment in Italy, in accordance with the literature that investigated this topic for most other developed countries (see, among others, Burggraeve et al. [15], Seyfried [48], Paladino and Vivarelli [40], Boltho and Glyn [12]). In other words, the results find employment sensitive to growth in current per capita GDP in line with the Okun's law (1962), which empirically asserted that GDP growth has an unemployment-reducing effect.

However, the southern Italian regions lag significantly behind the rest of the Country with regional per capita GDP equals to about half that of the remaining Italian regions [8]. This may help explain part of the observed employment divide between the treated and control regions, being the five treated regions all located in the South of Italy.

GDP grew somewhat uniformly along the national territory in more recent years, confirming in the South of Italy the turning point in the business cycle that occurred in 2015 after seven consecutive years of contraction [8]. However, as argued by Seyfried [48], although economic growth has a significantly positive impact on employment growth, some of these effects may take time before being fully felt. In other words, economic growth may have to occur for a period of time before it can significantly

Table 4 Diff-in-Diff estimates, Italy. Model with covariates, 2004–2012

	Coefficient	Standard error
Per capita GDP	0.280***	0.023
PEALT	−0.041	0.026
ALMP	−0.005*	0.003
ELETT	−0.036**	0.016
PRPSE	−0.016***	0.005
Corruption	0.135***	0.021
Government effectiveness	0.162***	0.026
Regulatory quality	0.115***	0.032
Rule of law	−0.012	0.016
Voice and accountability	−0.066*	0.037

* *Significant at 10%; ** Significant at 5%; *** Significant at 1%;*

affect the labour market, given the original disadvantage of the southern regions, the North–South territorial divide is unlike to disappear, inevitably reflecting on the regional employment dynamics. This would require taking further labour market measures in order to facilitate the adjustment of regional employment in response to changing economic settings [20, 23].

It is definitely clear the negative effect on regional employment performance of low human capital and poor/deprived/excluded people. While education is known to play a crucial role in society as a transformative agent [29, 39], the results show that the high level of education does not help explain the employment divide between the two groups of regions. In general, the Italian production structure is characterised by a lower use of skilled workforce compared to other European countries [8, 14, 25], and this aspect is particularly evident for the Southern Italian regions, which are even characterised by a weaker demand for high-educated and high-skilled workers. Although the observed employment gap cannot be ascribed to the different incidence of highly educated workforce between the two groups of regions, it is worth noting that low levels of education and poor standards of living conditions exacerbate each other's negative impact on employment [51, 52]. Indeed, adults with low qualification have a four times greater risk of being unemployed than their better educated counterpart [38]. Such a risk is also intensified by the difficulty of staying in the labour market once low educated persons managed to enter. Once in the workforce, low educated individuals are at increased risk for transition from paid employment into unemployment, disability pension, and early retirement [4, 47]. Therefore, this can be considered as a further backward factor for the southern regions, where low qualifications and jobs in low productivity sectors dominate [8], significantly affecting regional growth rates and, consequently, employment rates. Similarly, income poverty and material deprivation remain dramatic features of the least well-off regions of southern Italy, bringing many ill effects [11]. In these regions, socio-economic characteristics of poor people (i.e., unskilled manual workers, people in home duties, no or low education qualification) are more likely associated with underutilisation of labour, whether in the form of unemployment or massive underemployment [32].

The active labour market policies, which could be regarded as a proxy of the degree of unionisation of a country [31, 30], are government programs that would fight unemployment. Based on our results, in Italy, the package of active labour market policies seems not to help the unemployed in finding jobs. As discussed by OECD [38], these policies are usually based on under-funded projects, lacking of transparency and targeting, and rarely training provision addresses the lack of skills among adults. With an unemployment rate almost double the OECD average, in 2015, for example, Italy spent 0.51% of GDP on active labour market policies, which is just below the average spending among OECD countries (0.53% of GDP). It indicates that Italy invests relatively little in active measures per unemployed person, implying a low performance of the Italian system of employment services [38]. Moreover, spending on active labour market programs is not well directed in that half of the budget for active labour market policies is over-reliant on employment incentives. The frequent lack of a coherent national framework in allocating funding to regions and the still poor cooperation between regional institutions lead to unequal

access of services and to differences in quality and effectiveness of active labour market policies, usually at the expense of the less wealthy regions (see also Pupo and Aiello [43]). For example, the discrepancies in the type and quality of services to jobseekers—see, for example, the gap in the proportion of the unemployed who contact public employment offices in the North (38%) and the South (25%)—are mostly against the southern regions and the ability of their public employment offices to help individuals find a job, expanding the regional gaps in employment rates.

A significant share of the divide in employment rates between the two groups of regions is also attributable to differences in institutional quality, which may explain territorial imbalance as a result of differences in the size of the informal sector and shadow economy [1, 17 44]. In particular, three out of the five dimensions of institutional quality (i.e. corruption, government effectiveness and regulatory quality) suggest that a more efficient legal system and a lower propensity to corruption improve the employment performance of regions, in line with the strand of the economic literature that has proved the crucial role played by institutions on economic development [31, 49]. The highest values of IQI in the North and Centre give evidence about its potential relationship with employment rates, consistently with the literature recognising the direct association of a high quality of local institutions with a larger growth of the industrial sector, of production and hence of participation in the labour market [3, 33].

For example, in Italy, only about half of unemployed persons are registered with the public employment service. As discussed above, the great regional variability—from more than 50% in the autonomous provinces of Trento and Bolzano to less than 20% in Campania and Sicily [38]—is a clear sign of the low credibility of these agencies and institutions in general, mainly by citizens from the southern regions. A low quality of institutions may also lead labour market stakeholders to be reluctant to use the public employment service, as the access to active measures is partial and often perceived ineffective. Public participation would imply a broader consensus, greatly enhancing political interaction between citizens, job-brokers and government. Policies will be more effective if a broad coalition supports the proposal and works together to deliver it.

6 Discussion and Conclusions

The decennial European strategy ‘Europe 2020’ started in early 2010s with the aim to aid the Member States to overcome the economic crisis and to promote a smart, sustainable, and inclusive economic growth. This paper focused on the inclusive growth, according to which EU intend to promote an economy with a high employment rate that favours social and territorial cohesion. In particular, the EU set the objective of 75% employment rate for people aged 20–64 to reach by the end of the decade. To reach this objective, the EU resorted to a system of ‘indirect funding’ to which belong the regional funds such as the EFRD, ESF, and CF. The main objective of these funds is to reduce economic, social, and territorial disparities between the

various European regions. In other words, the EU goal is to trigger a convergence process between the forward and backward regions of Member States. The analysis we carried out highlighted the negligible effect of the EU indirect funding in narrowing the employment gap between these two groups of regions. In fact, the empirical analysis first highlighted that the funds have not produced effect across Member States (except for Germany) and then, by through a focus on Italian context, it has sketched an even worse scenario. In Italy, the gap between the convergence regions (i.e., those in southern area of the country) and the other ones widened after the EU funding programs 2007–2013. Some reasons could be responsible of this failure in Italy: (*i*) the financing system adopted by EU and the funds management; (*ii*) the monitoring activities on the progress and type of founded projects; (*iii*) heterogeneity of institutional quality between Italian regions.

The EU funding were structured as a co-financed system in which EU provided a half of the costs of the funding while the other half was paid through the national contributions (i.e., the Italian government). The co-financing system is not a bad idea by itself because in this way the interests of the provider (EU) and the recipient (national government) of the funds would coincide. However, a compensatory effect has taken place. In other words, the national government has strategically reduced the national spending targeted to backward regions replacing the national funds with the European ones. In this way, the net amount of funding has remained almost the same because of the EU financing has act more like a substitute of the national funds rather than an addition to them. Another critical point is the management of funds. The final beneficiaries of funds are the local administrations (e.g., the regions) but the financial planning is extremely complex involving the European, national, and local levels in a 7-year long strategic plan. Since the local administration represents the last level considered in this scheme, the most of plan is managed at centralised level (i.e., European, and national level). However, the centralised system tends to be less efficient than the decentralised one because the local administrations are more aware of the needs of their territories. In other words, local governments typically feature an advantage in terms of information and may propose prompter responses to bottom-up regional demands while central governments cannot.

A relevant problem is the effectiveness of monitoring actions on the right progression of projects funded. While the managers may not have the skills required to handle the funding properly, the projects heterogeneity represents something to take really into account. The project heterogeneity allows to implement proper active labour market policies and missing to reach variety could lead to the so-called *locking-in effect* [50]. According to this effect, unemployed people entered in a job seeking/training program tend to become locked-in to a temporary job reducing their search for a regular job. This is a risk that cohesion policy should try to avoid, since a great number of projects financed focused on human capital formation. The most of these programmes are payment-based since the participants have a subside to follow them. When the formation ends, whether the unemployed people do not find a job, they try to receive another subside by following another programme. This represent a trap and unemployed people are caught in the training loop without entering in a stable employment condition. In other words, people cannot use the

skills gained to find a job because of the employment demand does not grow and they could push out of labour market. To avoid this, the funded projects should focus on the creation of new enterprise and working opportunities. This is particularly true in the most economically depressed areas where the main problem is the lack in demand of workers rather than their competences. Our empirical results confirm this risk because they highlight a high difficulty of less educated people in finding a job and a substantial lack in employment advantages for more educated ones. In other words, while the less educated people could be chained in the lock-in effect because involved in a cycle of human capital development programmes, the more educated people are in an environment unable to valorise their competences. The aim of cohesion policy should be to stimulate the entrepreneurial initiatives (particularly by young people) by funding the creation of new small and medium-sized enterprises and creating in this way new employment opportunities.

The economic performances and the efficiency of public investments are strongly related to the quality of institutions, and of governments. This relationship is also valid in the case of the returns of EU funds. The same EU pointed out how low institutions quality may undermine efforts to achieve greater economic and social cohesion. In the Fifth Cohesion Report, the EU stated as follow: '*poor institutions can, in particular, hinder the effectiveness of regional development strategies*' ([19] p. 65). In other words, the weaker the institutional quality, the greater the difficulties in transforming EU funds into growth and development. Italy is historically characterised by high heterogeneity in institutional quality. The quality of local governments in southern regions (most of them are recipient of the funds) is far below that of the northern regions. This has contributed to create the socio-economic disparities that have characterised the nation so far [37]. In this work, we have controlled for five dimensions of institutional quality highlighting how these factors play a crucial role in explaining the failure of cohesion funds in Italy. In particular, our results showed the presence of complex relationships between the effectiveness of EU cohesion funds and the dimensions of governments effectiveness, regulatory quality, rule of law, and corruption. The reinforcement of these institutions' characteristics is essential for managing the projects funded and for controlling the quality of labour market policies. In a nutshell, independently of the policies adopted to improve the effectiveness of EU regional actions passes unquestionably through definition of clear rules that project managers have to respect, lower level of corruption in local administrations, and higher capability of local governments in defining labour market policies able to create a good environment for the creation of new business star-ups.

References

1. Acemoglu, D., Robinson, J.: The Role of Institutions in Growth and Development. World Bank, Washington DC (2008)
2. Aiello, F., Pupo, V.: Structural funds and the economic divide in Italy. J. Policy Model. **34**(3), 403–418 (2012)

3. Agovino, M., Garofalo, A., Cerciello, M.: Do local institutions affect labour market participation? The Italian case. BE J. Econ. Anal. Policy **19**(2) (2019)
4. Alavinia, S.M., Burdorf, A.: Unemployment and retirement and ill-health: a cross-sectional analysis across European countries. Int. Arch. Occup. Environ. Health **82**(1), 39–45 (2008)
5. Arbia, G., Le Gallo, J., Piras, G.: Does evidence on regional economic convergence depend on the estimation strategy? Outcomes from analysis of a set of NUTS2 EU regions. Spat. Econ. Anal. **3**(2), 209–224 (2008)
6. Ashenfelter, O.: Estimating the effect of training programs on earnings. Rev. Econ. Stat. **60**, 47–57 (1978)
7. Bähr, C.: How does sub-national autonomy affect the effectiveness of structural funds? Kyklos **61**(1), 3–18 (2008)
8. Bank of Italy.: The economy of the Italian regions—short-term dynamics and structural features, vol. 23 (2017)
9. Barham, C., Walling, A., Clancy, G., Hicks, S., Conn, S.: Young people and the labour market. Econ. Labour Mark. Rev. **3**(4), 17–29 (2009)
10. Becker, S.O., Egger, P.H., Von Ehrlich, M.: Effects of EU regional policy: 1989–2013. Reg. Sci. Urban Econ. **69**, 143–152 (2018)
11. Berry, A.: Inequality, poverty, and employment: What we know. Rev. Interv. économiques. Pap. Polit. Econ. **47** (2013)
12. Boltho, A., Glyn, A.: Can macroeconomic policies raise employment. Int. Labour Rev. **134**, 451–470 (1995)
13. Burggraeve, K., de Walque, G. and Zimmer, H.: The relationship between economic growth and employment. Econ. Rev. 32–52 (1994); Card, D., Krueger, A.B.: Wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania. Am. Econ. Rev. **84**, 772–793 (2013)
14. Castellano, R., Musella, G., Punzo, G.: Structure of the labour market and wage inequality: evidence from European countries. Qual. Quant. **51**(5), 2191–2218 (2017)
15. Dall'Erba, S.: Distribution of regional income and regional funds in Europe 1989–1999: an exploratory spatial data analysis. Ann. Reg. Sci. **39**(1), 121–148 (2005)
16. Dall'Erba, S., Fang, F.: Meta-analysis of the impact of European union structural funds on regional growth. Reg. Stud. **51**(6), 822–832 (2017)
17. Di Liberto, A., Sideri, M.: Past dominations, current institutions and the Italian regional economic performance. Eur. J. Polit. Econ. **38**, 12–41 (2015)
18. Esping-Andersen, G.: The Three Worlds of Welfare Capitalism. Cambridge, Polity (1990)
19. EU—EUROPEAN UNION.: Investing in Europe's Future. In: Fifth Report on Economic, Social and Territorial Cohesion. Publications Office of the European Union, Luxembourg (2010)
20. European Central Bank.: The employment-GDP relationship since the crisis. Econ. Bull. **1**(6), 53–71 (2016). Box 3, "Recent employment dynamics and structural reforms"
21. European Commission.: An agenda for new skills and jobs: a European contribution towards full employment (2010)
22. Commission, E.: European Structural and Investments Funds 2014–2020: Official text and commentaries. Publication Office of the European Union, Luxembourg (2015)
23. Farsio, F., Quade, S.: An empirical analysis of the relationship between GDP and unemployment. Humanomics **19**(3), 1–6 (2003)
24. Gagliardi, L., Percoco, M.: The impact of European cohesion policy in Urban and rural regions. Reg. Stud. **51**(6), 857–868 (2017)
25. Garofalo A., Castellano R., Punzo G., Musella G.: Skills and labour incomes: how unequal is Italy as part of the Southern European countries? Qual. Quant. 1–30 (2017)
26. Giordano, B.: Exploring the role of the ERDF in regions with specific geographical features: Islands, mountainous and sparsely populated areas. Reg. Stud. **51**(6), 869–879 (2017)
27. Hagen, T., Mohl, P.: 16 Econometric evaluation of EU cohesion policy: a survey1. In: International Handbook on the Economics of Integration: Factor Mobility, Agriculture, Environment and Quantitative Studies, vol. 3, p. 343 (2011)

28. Hansen, L.P.: Large sample properties of generalized method of moments estimators. Econ.: J. Econ. Soc. 1029–1054 (1982)
29. Harvey, L.: New realities: the relationship between higher education and employment. Tert. Educ. Manag. **6**(1), 3–17 (2000)
30. Huo, J., Nelson, M., Stephens, J.: Decommodification and activation in social democratic policy: resolving the paradox. J. Eur. Soc. Policy **18**, 5–20 (2008)
31. Knack, S., Keefer, P.: Institutions and economic performance: cross-country tests using alternative institutional measures. Econ. Politics **7**(3), 207–227 (1995)
32. Kuiper, M., Ree, K.V.D.: Growing out of poverty: how employment promotion improves the lives of the urban poor (No. 993817723402676). International Labour Organization (2005)
33. Lasagni, A., Nifo, A., Vecchione, G.: Firm productivity and institutional quality: evidence from Italian industry. J. Reg. Sci. **55**(5), 774–800 (2015)
34. Maynou, L., Saez, M., Kyriacou, A., Bacaria, J.: The impact of structural and cohesion funds on Eurozone convergence, 1990–2010. Reg. Stud. **50**(7), 1127–1139 (2016)
35. Mohl, P., Hagen, T.: Do EU structural funds promote regional growth? new evidence from various panel data approaches. Reg. Sci. Urban Econ. **40**(5), 353–365 (2010)
36. Mora, R., Reggio, I.: didq: a command for treatment-effect estimation under alternative assumptions. Stata J. **15**(3), 796–808 (2015)
37. Nifo, A., Vecchione, G.: Do institutions play a role in skilled migration? The case of Italy. Reg. Stud. **48**(10), 1628–1649 (2014)
38. OECD.: Strengthening Active Labour Market Policies in Italy. Connecting People with Jobs, OECD Publishing, Paris (2019)
39. OECD.: Education at a Glance Interim Report: Update of Employment and Educational Attainment Indicators. OECD, Paris, (2015)
40. Padalino, S., Vivarelli, M.: The employment intensity of economic growth in the G-7 countries. Int. Labour Rev. **136**, 191–213 (1997)
41. Percoco, M.: The impact of structural funds on the Italian Mezzogiorno. Région de Développement **21**, 141–153 (2005)
42. Percoco, M.: Impact of European cohesion policy on regional growth: does local economic structure matter? Reg. Stud. **51**(6), 833–843 (2017)
43. Pupo, V., Aiello, F.: L'Impatto della politica regionale dell'Unione Europea. Uno studio sulle regioni italiane. *Rivista italiana degli economisti*, *14*(3), 421–454 (2009)
44. Rodríguez-Pose, A.: Do institutions matter for regional development? Reg. Stud. **47**(7), 1034–1047 (2013)
45. Sargan, J.D.: The estimation of economic relationships using instrumental variables. Econ.: J. Econ. Soc. 393–415 (1958)
46. Sargan, J.D.: Testing for misspecification after estimating using instrumental variables. In: Contributions to Econometrics: John Denis Sargan, vol. 1, pp. 213-235 (1988)
47. Schuring, M., Robroek, S.J., Otten, F.W., Arts, C.H., Burdorf, A.: The effect of ill health and socioeconomic status on labor force exit and re-employment: a prospective study with ten years follow-up in the Netherlands. Scand. J. Work., Environ. Health 134–143 (2013)
48. Seyfried, W.: Examining the relationship between employment and economic growth in the ten largest states. Southwest. Econ. Rev. **32**, 13–24 (2011)
49. Surubaru, N.C.: Administrative capacity or quality of political governance? EU cohesion policy in the new Europe, 2007–13. Reg. Stud. **51**(6), 844–856 (2017)
50. Van Ours, J.C.: The locking-in effect of subsidized jobs. J. Comp. Econ. **32**(1), 37–55 (2004)
51. Van Zon, S.K., Reijneveld, S.A., de Leon, C.F.M., Bültmann, U.: The impact of low education and poor health on unemployment varies by work life stage. Int. J. Public Health **62**(9), 997–1006 (2017)

A BoD Composite Indicator to Measure the Italian "Sole 24 Ore" Quality of Life

Mariantonietta Ruggieri, Gianna Agrò, and Erasmo Vassallo

Abstract The measure of Quality of Life (QoL) is still a topic widely discussed in literature. In Italy, the newspaper "Il Sole 24 Ore" publishes a famous ranking that highlights strong disparities among provinces. In this paper, "Il Sole 24 Ore" and BoD-DEA methods are compared in order to show how different types of normalization and aggregation significantly influence the results making these rankings very fragile and questionable.

Keywords Quality of life · Composite indicators · BoD approach · Non-compensatory aggregation

1 Introduction

The "Il Sole 24 Ore" represents the most influential Italian economic and financial newspaper. It has been publishing for about 30 years a ranking of Quality of Life (QoL) for Italian provinces (areas at NUTS3 level) (Il Sole 24 Ore, 2020). The analysis, published annually in December, produces a great media resonance and great attention from politics, that uses this ranking to evaluate the good or bad government of the territories and to emphasize the persistence of strong economic and social disparities among provinces of northern and southern Italy. There are other similar investigations in Italy, for example the report of the newspaper "Italia Oggi" [8], but the report of the "Il Sole 24 Ore" is the best known, quoted and used one [9], for this reason, here, we refer to this specific study. Any ranking is strongly influenced by the choice of the basic dimensions, the kind of data and indicators, the standardization and the aggregation techniques, used to obtain a single overall score or position, so many criticisms can be addressed to investigations of this type, and even the purpose of these studies may appear questionable [11]. Actually, it is very difficult to provide a unique definition of QoL, as this concept invests personal and community aspects, depending on subjective and objective well-being and happiness [2]. Certainly, the

M. Ruggieri · G. Agrò · E. Vassallo (✉)
Department of Economics, Business and Statistics (SEAS), University of Palermo, Palermo, Italy
e-mail: erasmo.vassallo@unipa.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_17

concept of QoL is much more that "standard of living", basically connected to income levels [7]. The absence of a clear and univocal definition generates a partly indeterminate concept more difficult to translate into widely accepted data, indicators and measures, about that, there would be a lot to write. Anyway, the aim of this paper is not to discuss about the different definitions of QoL (and close concepts) used in the theoretical and applied literature. The target of this paper is to focus on an element often underestimated in these rankings, that is the statistical techniques used to standardize and synthesize the data: in fact, choosing one or another method can greatly influence the final ranking of the units, and this important limitation makes these studies much more fragile than they appear. Therefore, we accept the choice of "Il Sole 24 Ore" about the selection of data, indicators and dimensions used in its QoL dossier, although we have many perplexities on the redundancy of some data and their ability to correctly represent some pillars, but our interest is to focus on the methods and to compare results without being influenced by a different selection of data and variables. In particular, we compare the "Il Sole 24 Ore" procedure with a BoD-DEA method that, unlike other and simpler techniques, aggregates the basic indicators using non-constant weights, endogenously determined without external intervention of the researcher, in this way, the procedure automatically adapts the weights to the specific case, limiting personal and often arbitrary choices, that can significantly affect the final ranking [3, 4, 10]. The peculiarity of the BoD approach presented in this paper is the use of a model based on a non-compensatory multiplicative aggregation, with weights defined by a Benefit-of-Doubt procedure, where the specific restrictions are based on minimum and maximum limits of the impact of the component indicators. Moreover, in this way, using logarithms, it becomes simple to calculate in additive terms the contribution of the single indicators on the composite indicator. Section 2 compares the two procedures "Il Sole 24 Ore" and this specific BoD-DEA approach. Results and comments are reported in Sect. 3, while Sect. 4 concludes.

2 "Il Sole 24 Ore" QoL Approach and the BOD Method

We refer to the 2019 QoL dossier of the "Il Sole 24 Ore", the last available at the time of the first draft of this paper. The 2019 report uses 90 indicators divided into 6 dimensions (15 indicators for each): (1) income and consumption, (2) environment and services, (3) justice and safety, (4) business and labor, (5) demography and society, (6) culture and leisure, measured in all the 107 Italian provinces (according to the official NUTS3 nomenclature) [5]. For example, the first dimension "income and consumption" includes among others: value added per capita, bank deposits per capita, average selling price of houses, whereas the last dimension "culture and leisure" includes cinemas per capita, libraries per capita, and so on. All information related to data and the detailed list of indicators used by the newspaper, as well as in this work, can be found easily on the "Il Sole 24 Ore" website or in "Il Sole 24 Ore" [5, 6] and is not here reported for reason of space. We discuss the procedure

used by the newspaper to aggregate the data, while the choice of indicators and dimensions needs a separate discussion that is beyond the scope of this paper. We note that although name and meaning of the 6 dimensions have been similar since the 1990 report, variables and weighting technique have changed a lot over time; for example, the number of indicators were 42 in 2018 but 90 in 2019. For this reason, we will focus only on the 2019 report, without proposing any historical comparison. Moreover, the 2019 report could represent a turning point because, for the first time, a wide panel of external experts was recruited to select the best indicators available according to a holistic view, assuming to make this choice stable and durable for the following years. Finally, although "Il Sole 24 Ore" has recently published the 2020 report, unfortunately, during the year 2020, the SARS-CoV-2 epidemic has impacted on many economic and social aspects, altering the territorial representation of some indicators and their adequate expression of the level of quality of life. The 2019 "Il Sole 24 Ore" dossier attributes, for each indicator, 1000 points to the province with the best value and zero to that one with the worst value; the score for the other units is distributed in function of the distance from the two extremes. For each of the 6 dimensions, the ranking of each province is determined by the simple mean of the points of the basic indicators then, the final ranking, is obtained as simple arithmetic mean of the six dimensions; in summary, all the indicators have the same importance. This method has the advantage of being very simple and intuitive, but it ignores some statistical features of data, for example the different variability exhibited by different indicators. Besides, the simple arithmetic mean is a fully compensatory aggregation procedure; this implies that indicators and dimensions have the same importance and can also be considered perfectly complementary, even if this is not generally true. Unlike the "Il Sole 24 Ore", we first normalize the indicators with the well-known min–max procedure with translation on scale 1–10 [12]: $\left(x_{ij} - \min_j\left\{x_{ij}\right\}\right) \Big/ \left(\max_j\left\{x_{ij}\right\} - \min_j\left\{x_{ij}\right\}\right) \cdot 9 + 1$ for indicators with positive polarity (higher values correspond to better conditions, for example the level of GDP per capita), and $\left(\max_j\left\{x_{ij}\right\} - x_{ij}\right) \Big/ \left(\max_j\left\{x_{ij}\right\} - \min_j\left\{x_{ij}\right\}\right) \cdot 9 + 1$ for indicators with negative polarity (higher values correspond to worse conditions, for example the number of crimes per inhabitant). In such a way, we have values between 1 (worst value) and 10 (best value), with the advantage to obtain indicators with unique orientation, that is higher values always correspond to better conditions, and to take into account the intrinsic variability of each indicator. About the question of the compensatory nature, we use a geometric mean of indicators that, as known, compared to the simpler arithmetic mean, has the advantage of not being excessively affected by a few extremely small or large values. Moreover, we apply non-constant weights, that can vary both for each indicator and for each province, these weights are determined, according to the "Benefit of Doubt" (BoD) logic with an "automatic" procedure independent of the researcher's opinions. This last aspect is not without weaknesses: why an expert's opinion is less relevant than an automatic procedure? On the other hand, in practice, the alternative to automatic procedures is very often the simple identical weighting. In this context, automatic methods based on a BoD logic have received many consents [1, 12, 13, 16]. These methods exploit the Data Envelopment Analysis (DEA), a frontier technique that has been usually used for measuring efficiency in

production. Following Giambona and Vassallo [3, 4], a generic composite indicator (CI) is

$$CI_i = \prod_{j=1}^{n} I_{ij}^{w_{ij}} \tag{1}$$

where I_{ij} is the jth normalized indicator ($j = 1, \ldots, n$) for the ith province ($i = 1, \ldots, m$) with weight w_{ij}. The weights are determined endogenously by an algorithm based on a multiplicative optimization model, that solves the following problem

$$CI_i = \max_{w} \prod_{j=1}^{n} I_{ij}^{w_{ij}} \tag{2}$$

with constraints

$$\prod_{j=1}^{n} I_{ij}^{w_{ij}} \leq e \; and \; w_{ij} \geq 0 \tag{3}$$

where e is the Napier's constant [17]. The CI is obtained by multiplying the component indicators with weights calculated in the best possible conditions, i.e. increasing as much as possible the composite score for a given province. In short, a low value of the CI for the ith province is due to low values of the indicators that compose it and not to specific weights, calculated to obtain the best result, i.e. the maximum possible value for the ith unit compared to the benchmark province. At the end, we obtain scores between 1 and 2.71 (the Napier's constant), attributable to the most intuitive interval between 0 and 1 by applying the antilogarithm. However, the optimization problem could attribute weight zero to some indicators and too much weight to others, and this is not desirable if all the dimensions are relevant. At this aim, we add specific constraints on the weights, in particular, we add proportion constraints to the model (2)–(3)

$$\left(\prod_{j=1}^{n} I_{ij}^{w_{ij}} \right)^{L} \leq I_{ij}^{w_{ij}} \leq \left(\prod_{j=1}^{n} I_{ij}^{w_{ij}} \right)^{U} \tag{4}$$

where U and L ranging between 0 and 1 to represent the lower and the upper bound (in percentage terms) for the contribution of the jth indicator. In fact, without constraints on the weights, the model could ignore the contribution of the under-performing indicators or dimensions to obtain the best solution, and this is not admissible. For our application on the "Il Sole 24 Ore" data, we use the BoD procedure in two steps separating the indicators by the six dimensions. In this way, in a first step, we have six CIs, each calculated on the 15 indicators of each dimension and, in a second step, we obtain a new CI, an overall CI, where the component indicators are the six CIs obtained previously in the first step. This operational choice is dictated by the need not to disperse the characteristics of each dimension and not to risk making

insignificant the contribution of some basic indicator. This two-step procedure allows a greater detail of the analysis and it allows the minimum limits on the weights to be adequately varied. In our case, we consider $L = 5\%$ for the CIs of the six dimensions and $L = 14\%$ for the overall CI, while U is defined accordingly. We note that bounds slightly lower or slightly higher lead to superimposable results, so an intermediate value has been chosen among the possible alternatives, taking into account that L (to guarantee sufficient flexibility to the method) must be relatively low: in fact, this is the only (weak) discretionary choice of the researcher within the proposed procedure. How do the results of "Il Sole 24 Ore" change by applying this BoD method instead of using simple arithmetic means? We answer to the question in the next section.

3 Comparing the Two Procedures: Results and Comments

In this section we compare the original 2019 ranking of the "Il Sole 24 Ore", based on simple means with identical weights among indicators and for all the 107 provinces, and our multiplicative BoD ranking based on min–max normalizations and geometric means with weights endogenously defined and specific for each indicator and for each province (each indicator and each province could have a specific and different weight). We compute six BoD composite indicators for each of the six dimensions indicated by the "Il Sole 24 Ore" newspaper and an overall BoD composite indicator which combines the previous six composite indicators of the dimensions; in the first case, the lower limit of the weights is set at 5% and, in the second case, at 14%. It is interesting to note that the two rankings appear consistent between the two methods, but with many significant differences.

Figure 1 shows a correlation plot among the "Il Sole 24 Ore" ("sole") and the BoD ("bod") composite indicators for the overall case and for each of the six dimensions (0 represents the overall score, 1–6 refers to the six dimensions). Figure 2 reports the kernel density estimates for each of the six CIs and for the overall CI; for each pair of indicators "sole" and "bod", the average bandwidth is used [15, 14]. To allow a direct comparison, the composite indicators were rescaled in the range 0–100, and this naturally does not affect the position and ranking of the 107 provinces. Some "twin peaks" can be noted in the distributions reported in Fig. 2, they are clearly evident for the overall indicators, substantially in correspondence with the provinces of the southern and northern Italy. Nevertheless, the interpretative framework is much more fragmented than the simple division between areas in the North with higher quality of life and areas in the South with lower quality of life. Rather, a more complex classification appears which, although consistent between the two techniques, shows the effect of a different method of aggregation and weighting.

Unlike the simple arithmetic mean, the BoD method tends to reward the best results of the provinces and it puts each unit in the best possible conditions in absence of widely accepted information on the role of the basic indicators. In this regard, it is interesting to note a strong change of position of some provinces compared to the "Il Sole 24 Ore" original classification.

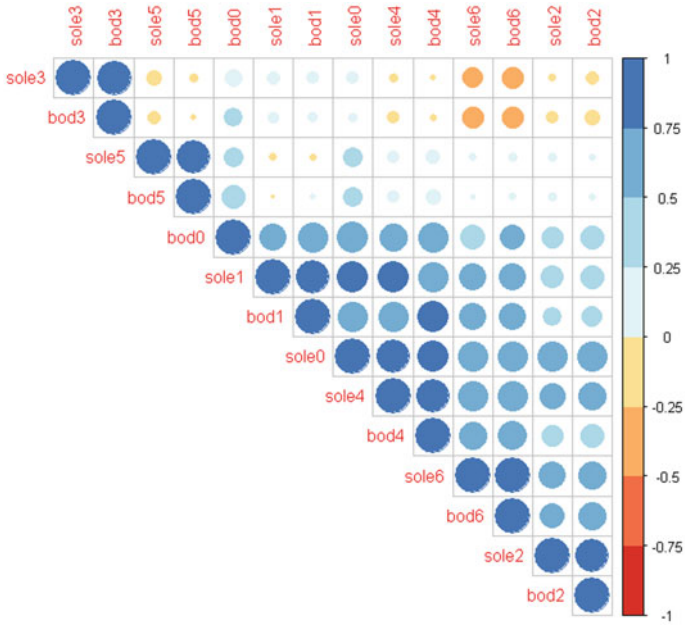
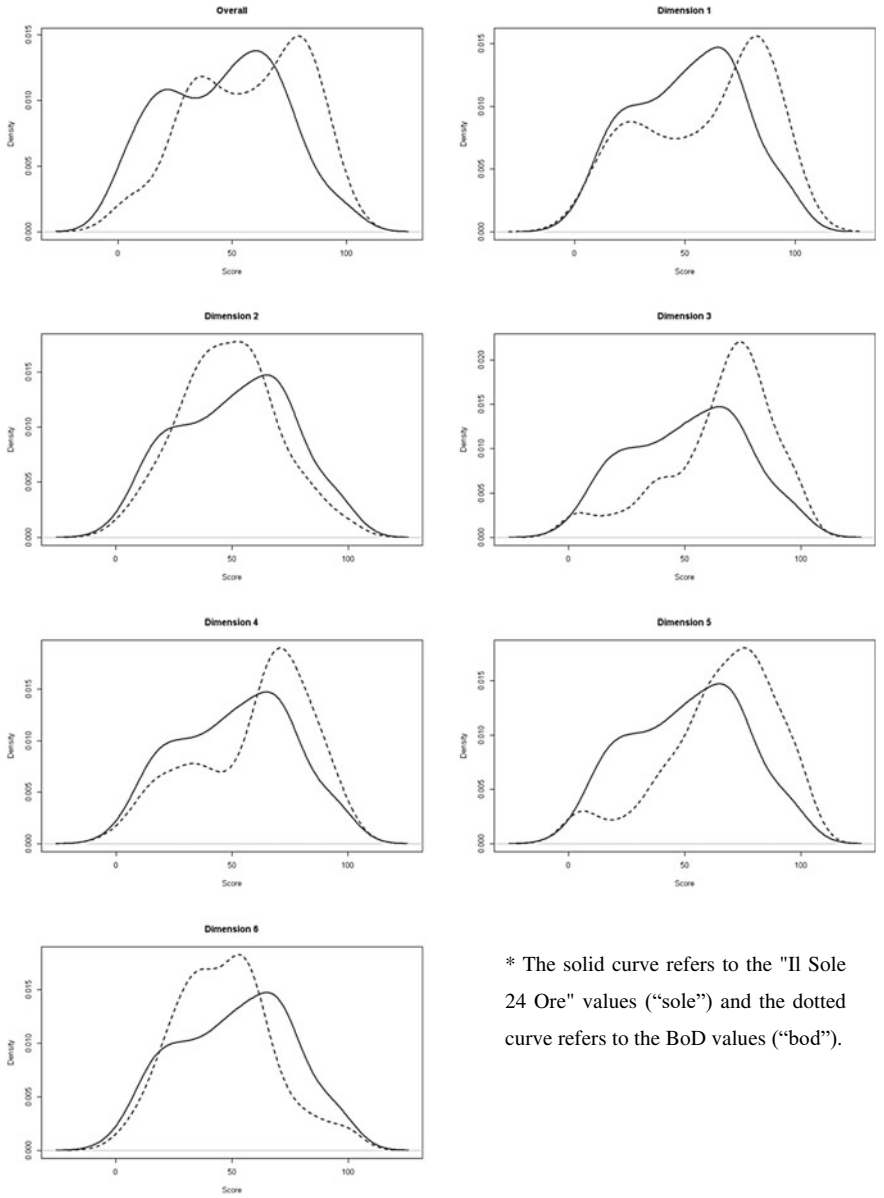


Fig. 1 Correlation plot among the "sole" and "bod" CI values

In particular, for the overall indicator in Fig. 3, the provinces no.38 (Genova), 49 (Lodi), 82 (Roma), 16 (Bolzano), 34 (Firenze), 56 (Milano) and 98 (Trieste) show a BoD score lower than "Il Sole 24 Ore" score (panel *a* in Fig. 3) and significant changes of position appear for these and other provinces (panel *b* in Fig. 3).

Figure 4 shows the maps of the 107 provinces with CI above third quartile, between the second and the third quartile, between the first and the second quartile and under the first quartile for the "sole" (panel *a*) and "bod" (panel *b*) scores; darker colours correspond to higher QoL values: it is immediate to note that the "bod" distribution is less polarized between the northern and southern provinces than the "sole" distribution, with a Moran's I of spatial association lower for "bod" (0.469, p-value 0.001) than for "sole" (0.645, p-value 0.001). In detail, just to give some examples, Milano loses 97 positions in BoD (from 1st to 98th), Trieste loses 96 positions (from 5th to 101st), Firenze 78 positions (from 15th to 93rd), whereas Messina gains 25 positions (from 100 to 75th), Salerno 26 (from 86 to 60th), Lucca 29 (from 54 to 25th). Some changes are therefore very significant: Milano loses (and quite a lot) its primacy in favour of Venice (ninth for the "Il Sole 24 Ore"). This happens because the "BoD" technique does not make a simple mean of the basic indicators, then, in our opinion, it generates a more representative ranking of the reality (which obviously depends on the very questionable basic indicators). Moreover, still referring to Milano, it is useful to note that "Il Sole 24 Ore" attributes rank 2, 5, 107, 1, 9 and 2 for the six dimensions, respectively: evidently, the same general classification of the newspaper does not adequately evaluate the bad performance in the 3rd dimension of Milano.



* The solid curve refers to the "Il Sole 24 Ore" values (“sole”) and the dotted curve refers to the BoD values (“bod”).

Fig. 2 Kernel density estimates of the “sole” and “bod” overall CI values

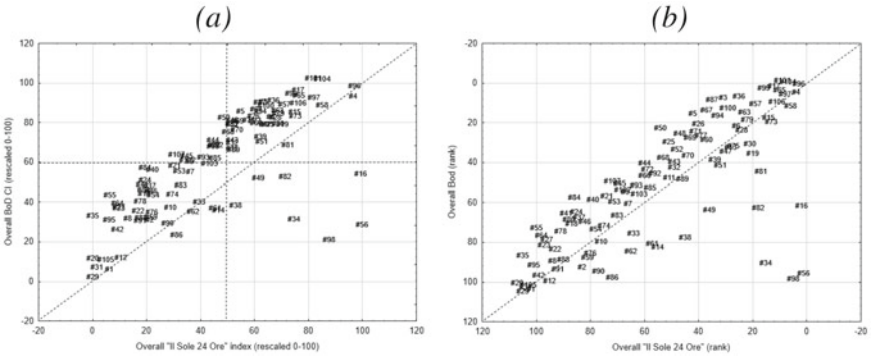


Fig. 3 Score (panel *a*) and ranking (panel *b*) of the pair "sole"-"bod" for the overall CIs

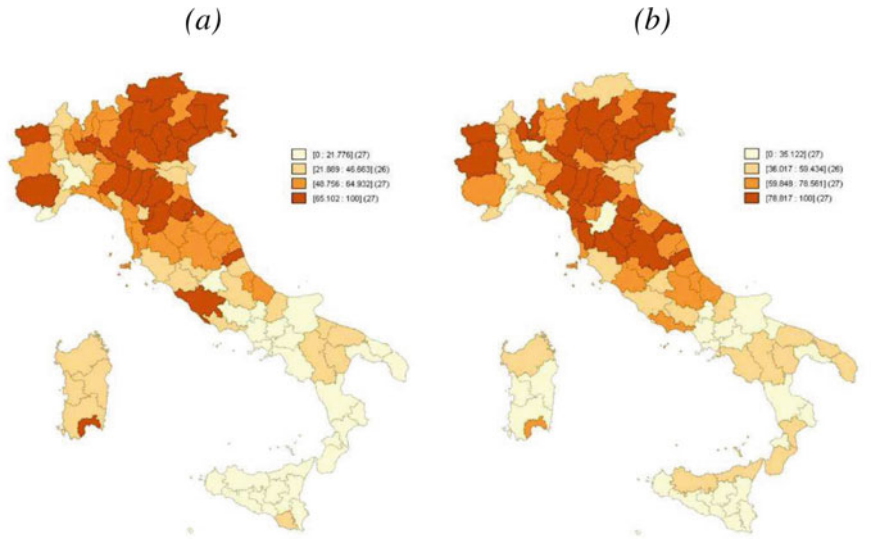


Fig. 4 Quantile maps of the "sole" (panel *a*) and "bod" (panel *b*) overall CI scores

The corresponding rank obtained by BoD for Milano are: 80, 13, 106, 1, 11 and 1, with a low performance in dimension 1 and 3, that is taken into account in the general classification justifying the strong change of position.

4 Conclusions

In this paper, a composite indicator is proposed via multiplicative BoD-DEA procedure in order to obtain a general score of quality of life (QoL) for the 107 Italian

provinces. We have used the same data and indicators considered in the well-known 2019 QoL dossier published by the "Il Sole 24 Ore" Italian newspaper, that represents the most detailed and famous analysis on QoL in Italy. This analysis has been repeated annually for 30 years and, for the year 2019, it is based on 90 indicators distributed over 6 dimensions. The choice of data and indicators is not without criticism, but here we do not discuss this point. Rather, we focus on aggregation and weighting of the data and on the consequent risk of excessive weakness of these rankings. The original "Il Sole 24 Ore" procedure uses a simple standardization and a fully compensatory arithmetic mean, whereas the multiplicative BoD method applies different weights to dimensions and provinces according to specific characteristics of the data; besides, it uses a geometric mean to avoid the problem of full compensation among indicators or dimensions. From a technical point of view, the BoD is a more specific and advanced method and addresses some critical aspects compared to simpler methods; nevertheless, it is not possible to assert that an endogenous choice of the weights is always better than an exogenous choice or than the use of identical weights for units and dimensions. We cannot say in absolute terms which is the best possible choice: in fact, there could be substantial and qualitative reasons in favour of an expert judgment. Certainly, the techniques of standardization and aggregation strongly influence the rankings, and this generates many doubts about the utility of these classifications if they are not accompanied by due caution. In our case, although in a substantially coherent general context, we note significant changes in ranking and position for some provinces: for example, Milano has a very low score in BoD and this totally reverses the ranking of the "Il Sole 24 Ore". In short, the provinces having good results in all dimensions and indicators are rewarded with BoD, despite of the provinces that have a strong polarization in some indicators or dimensions.

References

1. Aparicio, J., Kapelko, M., Monge, J.F.: A well-defined composite indicator: an application to corporate social responsibility. J. Optim. Theory Appl. **186**, 299–323 (2020)
2. Cummins, R.A.: Quality of life definition and terminology: a discussion. In: Document from the International Society for Quality-of-Life Studies. International Society for Quality-of-Life Studies. *(mimeo)* (1998)
3. Giambona, F., Vassallo, E.: Composite indicator of social inclusion for European countries. Soc. Indic. Res. **116**(1), 269–293 (2014)
4. Giambona, F., Vassallo, E.: Composite indicator of social inclusion for the EU countries. In: Alleva, G., Giommi, A. (eds.) Topics in theoretical and applied statistics. Springer, Berlin (2016)
5. Il Sole 24 Ore.: Dossier Qualità di vita 2019, Il Sole 24 Ore, Milano (2019)
6. Il Sole 24 Ore.: Dossier Qualità di vita 2020, Il Sole 24 Ore, Milano (2020)
7. ISTAT.: Rapporto BES 2019, Istat, Roma (2019)
8. Italia Oggi.: Report sulla qualità della vita 2020, Italia Oggi, Milano (2020)
9. Lun, G., Holzer, D., Tappeiner, G., Tappeiner, U.: The stability of rankings derived from composite indicators: analysis of the '"Il Sole 24 Ore"' quality of life report. Soc. Indic. Res. **77**, 307–331 (2006)

10. Munda, G., Nardo, M.: Noncompensatory/nonlinear composite indicators for ranking countries: a defensible setting. Appl. Econ. **41**(12), 1513–1523 (2009)
11. Nussbaum, M., Sen, A.: The Quality of Life. Oxford Press, Oxford (2003)
12. OECD: Handbook on Constructing Composite Indicators: Methodology and User Guide. Oecd, Paris (2008)
13. Rogge, N., Van Nijverseel, I.: Quality of life in the european union: a multidimensional analysis. Soc. Indic. Res. (2018). https://doi.org/10.1007/s11205-018-1854-y
14. Scott, D.: Multivariate Density Estimation: Theory, Practice, and Visualization. Wiley, New York (1992)
15. Silverman, B.W.: Density Estimation for Statistics and Data Analysis. Chapman $ Hall, London (1986)
16. Szuwarzyński, A.: Benefit of the doubt approach to assessing the research performance of Australian universities. High. Educ. Q. **73**(2), 235–250 (2018)
17. Zhou, P., Ang, B.W., Poh, K.L.: A mathematical programming approach to constructing composite indicators. Ecol. Econ. **62**, 291–297 (2007)

Trends and Random Walks in Mortality Series

Giambattista Salinari and Gustavo De Santis

Abstract The notion that time series cannot be properly dealt with until their nature has been established is nowadays largely accepted among economists, less so among demographers. In this paper, based on theoretical considerations and empirical data, we prove that mortality evolves over time following a geometric random walk with drift. If this is true, other series too must follow a non-stationary path, for instance person-years and survivors in mortality tables, and survivors in actual populations. In the empirical part of the paper, we carry out 160 tests on age-specific log-mortality rates in France and England-Wales (at ages 0–79) over the years 1850–2016. The DS (difference stationary), not TS (trend stationary) nature of the series emerges clearly, probably with just one unit root.

Keywords Mortality dynamics · Difference stationary · Trend stationary · Unit roots

1 Introduction

Modern time series analysis distinguishes between trend stationary (TS) and difference stationary (DS) phenomena, depending on the transitory (TS) or permanent (DS) effects of innovations. In an attempt to explain the sustained economic growth of some European countries in the nineteenth century, scholars originally assumed a TS process: a long-term trend, determined by structural factors (e.g. capital accumulation and population growth), with minor and ultimately irrelevant short-term fluctuations. In 1982, however, Nelson and Plosser proposed the DS hypothesis: economic growth may be a random walk with drift, where innovations, for example in technology, end up by permanently diverting the system from its previous trend.

G. Salinari (✉)
Department of Economics and Business, University of Sassari, Sassari, Italy
e-mail: gsalinari@uniss.it

G. De Santis
DiSIA, University of Florence, Florence, Italy
e-mail: gustavo.desantis@unifi.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_18

This debate has been largely ignored in demography, where, however, the distinction between TS and DS time series may be equally important. Consider, for example the interpretations of the mortality decline since the end of the nineteenth century: some attribute it to nutritional improvements (and therefore to economic growth; [6], some instead, to public policies, especially health care [3], but neither hypothesis can be tested until the TS or DS nature of the process has been ascertained.

In the rest of this paper, we proceed in two steps. First, we argue that mortality rates follow a DS-type process, and we prove that if the dynamics of mortality rates is DS, so is the dynamics of other related demographic variables. Then, we try to support our hypothesis empirically, using "unit root" tests on mortality series.

2 On the Dynamics of Mortality Rates

Within a cohort life table, mortality rates are defined as the ratio between deaths at age x and the person-years at age x

$$m_{t,x} = \frac{d_{t,x}}{L_{t,x}} = \frac{l_{t+1,x+1} - l_{t,x}}{L_{t,x}} \tag{1}$$

where

$$L_{t,x} = \frac{l_{t,x} + l_{t+1,x+1}}{2} \tag{2}$$

and where $l_{t,x}$ represents those who celebrate their x-th birthday in year t (Fig. 1).

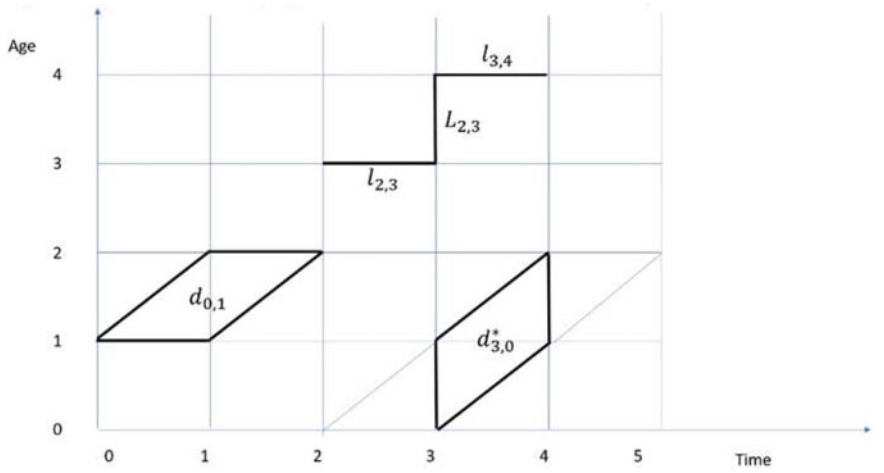


Fig. 1 Selected demographic life table variables (Lexis diagram)

The instantaneous rate of mortality, called force of mortality $\mu_t(x)$, can be written as

$$\mu_t(x) = \lim_{\Delta x \to 0} \frac{l_t(x) - l_t(x + \Delta x)}{\Delta x l_t(x)} \tag{3}$$

Over time t, for each age x, the evolution of $\mu_t(x)$ is subject to two different forms of "innovations", or shocks. On the one hand, those that produce permanent effects on the series such as vaccines, penicillin and antibiotics (or, with opposite effects, the formation of antibiotic-resistant bacteria, the spread of smoking and alcohol drinking, etc.). On the other hand, there are innovations that induce only transitory effects. For example, the worldwide rise in mortality caused by Covid-19 in 2020 is unlikely to last. The same happened with the heat wave in Europe in the summer of 2003 or the tsunami on the Indonesian coast in 2004.

The effect of the first type of innovations can be captured by the following equation:

$$\ln[\mu_t(x)] = \ln(\delta) + \ln[\mu_{t-1}(x)] + \ln(w_t) \tag{4}$$

where the log-force of mortality in year t depends on that of the previous year, with ln (w_t) representing a white-noise process and ln(δ) the drift, arguably with $\delta < 1$ (decline). In this model, "random" innovations w_t affect simultaneously all age classes (proportionality assumption; [10]).

Equation (4) can be applied recursively, from a conventional starting year 0, which results in

$$\ln[\mu_t(x)] = \ln(\delta)t + \ln[\mu_0(x)] + \sum_{j=1}^{t} \ln(w_j) \tag{5}$$

showing that innovations w_j produce permanent effects in this model, by modifying all subsequent values of $\mu_t(x)$.

Innovations with transitory effects can be introduced as follows

$$\ln[\mu_t(x)] = \ln(\delta)t + \ln[\mu_0(x)] + \sum_{j=1}^{t} \ln(w_j) + e_t \tag{6}$$

where e_t indicates an ARMA-type (autoregressive moving average) stationary process.

3 The Evolution of Survivors at Age *x*

As mentioned, the purpose of this paper is double:

(1) to show that mortality evolves as described in Eq. (6), and,
(2) to show that, if this is true, the dynamic of a long series of other demographic variables (not solely those that can be found within a mortality table) will also be non-stationary.

Let us start with the latter demonstration. Within a life table with radix $l_0 = 1$, survivors at age x can be written as [9]

$$l_t(x) = e^{-M_t(x)} \tag{7}$$

where $M_t(x)$ is the cumulated force of mortality

$$M_t(x) = \int_0^x \mu_t(a)da \tag{8}$$

Ignoring the transitory component for a moment, we can use Eq. (4) to rewrite $M_t(x)$ as

$$M_t(x) = \int_0^x \mu_t(a)da = \delta w_t \int_0^x \mu_{t-1}(a)da = M_{t-1}(x)\delta w_t \tag{9}$$

Substituting (9) into (7) and taking the logarithm we obtain

$$\ln[l_t(x)] = \ln[l_{t-1}(x)]\delta w_t \tag{10}$$

which suggests that the time series of the log-survivors at age x evolves as a geometric random walk.

If, instead, the transient component of the model is not negligible, survivors evolve as a geometric random walk with noise

$$\ln[l_t(x)] = \ln[l_0(x)]e^{e_t}\delta^t \prod_{j=1}^{t} w_t. \tag{11}$$

With a few adjustments, these conclusions apply also to the years of life lived, or person years (Appendix A), so

$$\ln[L_t(x)] = \ln[L_{t-1}(x)]\delta W_t \tag{12}$$

where $W_t = \frac{1}{2}(w_t + w_{t-1})$.

4 The Dynamics of Population Aged *x*

The probability $p_{t,x}$ of surviving until the exact age $x + 1$ for those who have reached age x is related to the mortality rate $m_{t,x}$ by the following equation [9]

$$\ln(p_{t,x}) = -m_{x,t} \tag{13}$$

As $\mu_t(x) \approx m_{t,x}$, the log-probabilities of survival follow a geometric random walk

$$\ln(p_{t,x}) = \delta \ln(p_{t-1,x}) w_t \tag{14}$$

and this can be used to derive the dynamics of the population surviving at exact age x in year t, $N_{t,x}$

$$\ln(N_{t,x}) = \ln(p_{t-1,x-1}) + \ln(N_{t-1,x-1}) \tag{15}$$

This is a non-stationary process around a trend, because every innovation w_t (in Eq. 14) has a permanent effect on the ln ($N_{t,x}$) series.

The conclusion is that if the dynamics of the log mortality rates follow a DS process, then several other series will do the same, among which $p_{t,x}$, $l_{t,x}$, $L_{t,x}$ and $N_{t,x}$.

5 Unit Root Tests

Let us now take a step backwards. Equation (4) assumes that the evolution of the log-force of mortality follows a random walk with drift. Exploiting the approximation $\mu_t(x) \approx m_{t,x}$, we can use the mortality rate series of the Human Mortality Database (HMD) to carry out a series of unit root tests, to assess the DS or TS character of the series.

The most widely known of these is probably the Dickey-Fuller (DF) [2] test. However, we will not use it here because it is negatively affected by autocorrelation and heteroskedasticity of residuals, and at least the latter is likely to be particularly problematic for our analyses. Assuming, with Brillinger [1] that the deaths of the population $D_{t,x}$ distribute as a Poisson

$$D_{t,x} \sim Poisson[\mu_t(x) E_t(x)], \tag{16}$$

where $E_t(x)$ represents those who are exposed to the risk of dying at age x in the population, the variance of the mortality rates $m_{t,x}$ equals $\frac{1}{D_{t,x}}$ [4]. As both the force of mortality $\mu_t(x)$ and the population at risk $E_t(x)$ vary over time (the former generally declining, the latter frequently increasing), heteroskedasticity is probable.

This is why instead of the DF test we will use the Phillips and Perron (PP) test, similar to the former, but capable of monitoring also the presence of heteroskedasticity and serial autocorrelation in the residuals [7].

In both the DF and the PP tests, the null hypothesis assumes the series to be non-stationary (i.e., with a unit root). The non-rejection of the null hypothesis is therefore a necessary first step, but it may not suffice, because non-rejection could be due to the limited power of the test. Additional support may come from the [5] test (KPSS), where the null hypothesis is that the process is stationary around a trend.

As the PP and the KPSS unit root tests are based on markedly different assumptions, they are customarily used simultaneously to determine the DS or TS character of a series. In our analyses, we will carry out both tests twice: first on the original series $\ln(m_{x,t})$ and then on their first differences $\Delta\ln(m_{x,t})$. With the former, we test the hypothesis of non-stationarity (at least one unit root). If these tests are passed, there is still the possibility of multiple unit roots. This is why we need the second series of tests, carried out on the first differences: if the process becomes stationary after differentiation there is only one unit root, and the evolution of mortality over time is consistent with a geometric random walk as in Eqs. (4) and (6).

6 Our Data

So far we have relatively disregarded the transitory component e_t of Eq. (6), which, however, may play an important role. The transitory component e_t includes two different phenomena. First, it captures all transient variations in mortality due to climate, epidemics, wars, natural disasters, etc. The importance of this component was probably greater in the past, before the nineteenth century, than it is today, when, after the epidemiological transition, mortality is relatively "under control" [8].

Secondly, e_t captures the sample variability of the log-mortality rates. As their variance is $\frac{1}{D_{t,x}}$, other things being equal, the smaller the population (or the age group), the greater the noise. This makes estimation difficult for the age groups where mortality is low (e.g., between 7 and 35 years) and for those with few survivors (e.g., above 80 years).

To circumvent these difficulties, two strategies are possible:

(1) Select countries with long time series, or
(2) Select large populations, to reduce the sample variability of the log-mortality rates.

Among the possibilities offered by the HMD, two populations stand out as the best candidates for our analyses: France and England-Wales. In both of them, the Human Mortality Database (HMD) series began even before mid-nineteenth century (but our analyses begin in 1850, because data quality is lower before that date) and extend over a period of 167 years, up to 2016. Both countries have a medium-sized population, currently of about 60 million individuals (about 56 million in England-Wales and 67 million in France). Finally, in both cases it is possible to distinguish

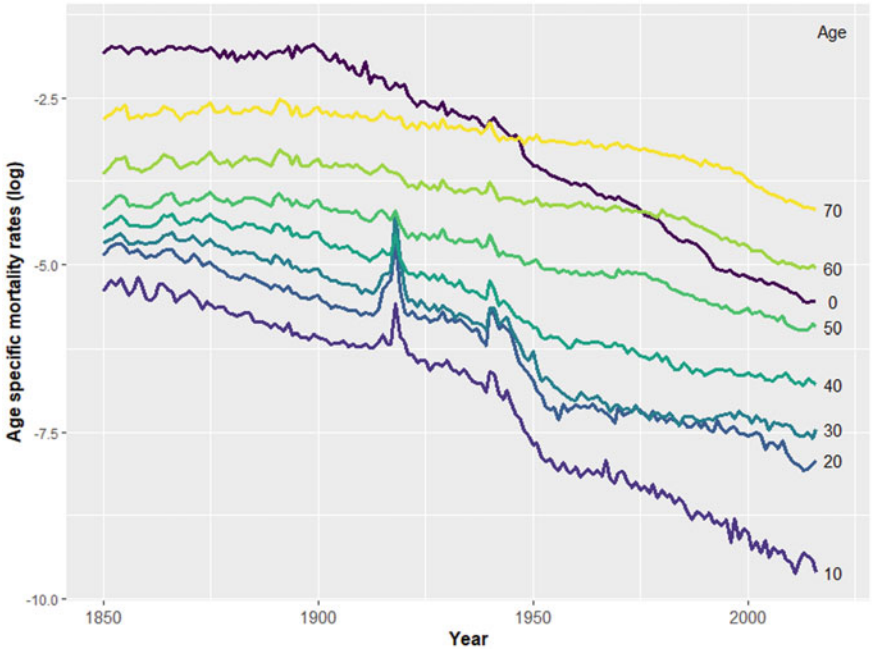


Fig. 2 Evolution of age specific mortality rates by selected ages. England-Wales 1850–2016 (in logarithms) *Source* Own calculations on HMD data (https://www.mortality.org)

the civilian from the military population, which allows us to reduce the (transitory) mortality effects of the two world wars.

Figures 2 and 3 show a few series of age specific log-rates of mortality for these two countries.

Most of these series tend to move in parallel, consistently with the hypothesis of proportionality, excluding however the young, up to 10 years, and the very young (age 0). Mortality at young and very young ages seems to have declined faster, especially after the end of the nineteenth century. This means that the proportionality hypothesis—implicit in the drift and the innovation parameters of Eqs. (4) and (6), i.e. $\ln(\delta)$ and $\ln(w_t)$ respectively—may hold only starting from a certain age, for instance 10 years.

In both the French and the English series, the mortality shocks of the twentieth century (two World Wars and the Spanish flu epidemic) are clearly visible and affect mainly the population below age 30. In the French case, two further mortality crises can be detected in the second half of the nineteenth century, probably due to cholera (once again, hitting harder on the young). We mention these crises both because they are interesting per se, and because they can make it more difficult for us to identify the DS or TS character of our (age-specific) series.

In short: we used the yearly series of age-specific log mortality rates for France and England-Wales between 1850 and 2016. The tests were repeated for each age

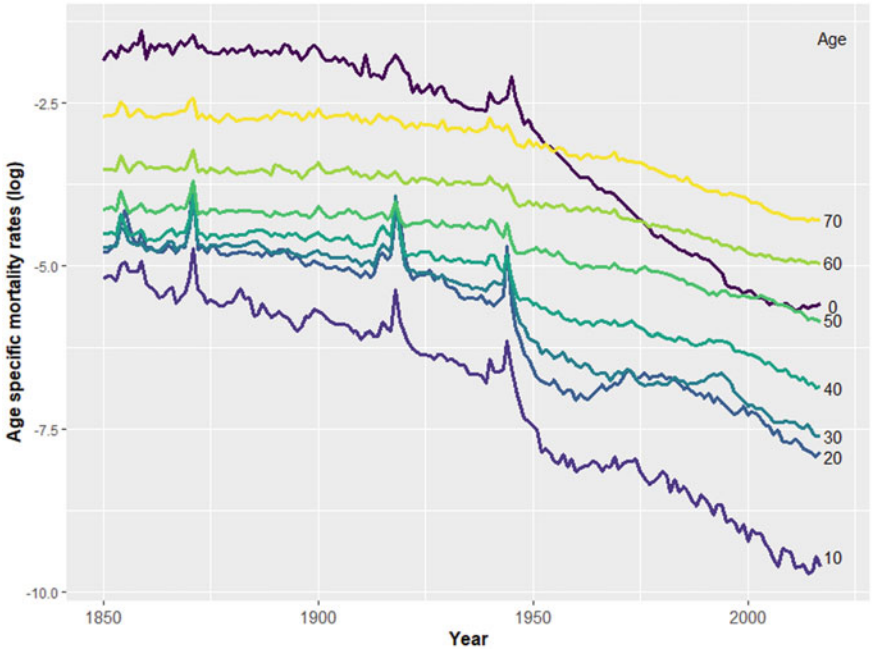


Fig. 3 Evolution of age specific mortality rates by selected ages. France 1850–2016 (in logarithms). *Source* Own calculations on HMD data (https://www.mortality.org)

in the 0–79 years range. We stopped before age 80, in part because the indication of age at death becomes less reliable at older ages and in part because observations become too few and results erratic.

7 Unit Root Tests: Results

In Fig. 4 (England-Wales) and Fig. 5 (France), each bar represents the *p* value associated with a unit root test carried out on the log-mortality rates at a that specific age over 152 observations (years), and the horizontal dotted line marks the 5% significance level. In general, our results appear to be consistent: the KPSS tests tend to reject the null hypothesis that series are stationary around a trend, and, consistently, the PP tests tend *not* to reject the null hypothesis that the evolution of the series is *non*-stationary. A partial exception are the ages around 20 years, where the tests indicates trend-stationarity in England-Wales and something not well identified in France. However, the idea that the path of mortality at age 20 or so evolves over time differently from what it does at other ages does not seem very convincing. We submit that the exception at this age may be a consequence of the great mortality crises of

the twentieth century (two World Wars and the Spanish epidemic; see again Figs. 2 and 3): the "noise" is too strong for our tests to produce clear-cut results.

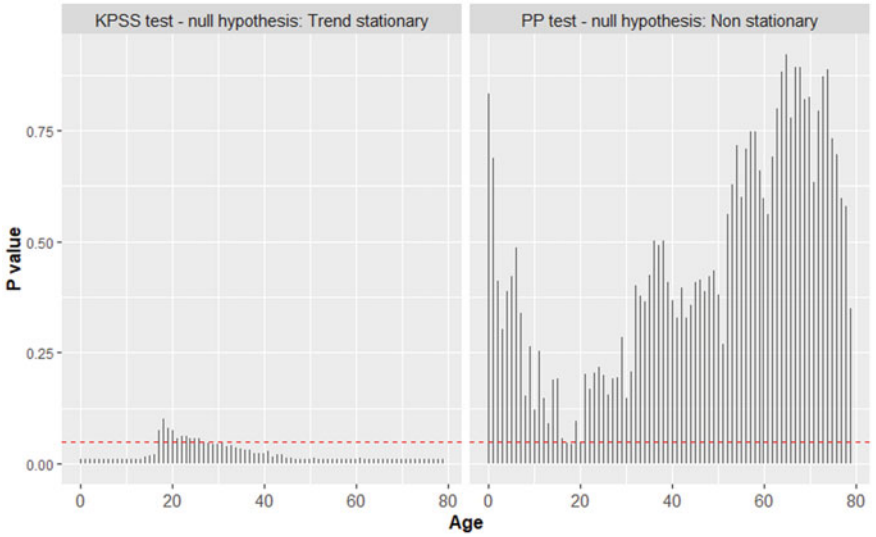


Fig. 4 Results of the unit root tests on the log-mortality rates in England-Wales, 1850–2016 (age-specific *p* values) *Source* Own calculations on HMD data (https://www.mortality.org)

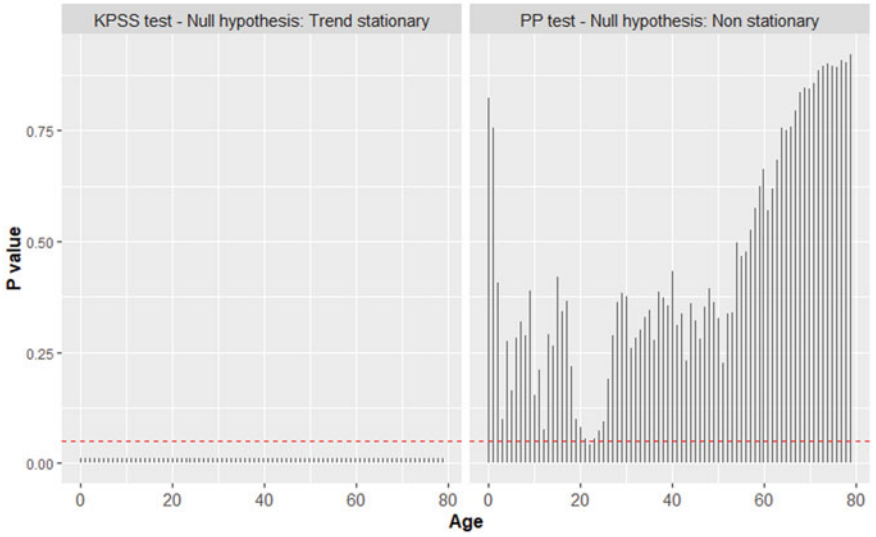


Fig. 5 Results of the unit root tests on the log-mortality rates in France, 1850–2016 (age-specific *p* values) *Source* Own calculations on HMD data (https://www.mortality.org)

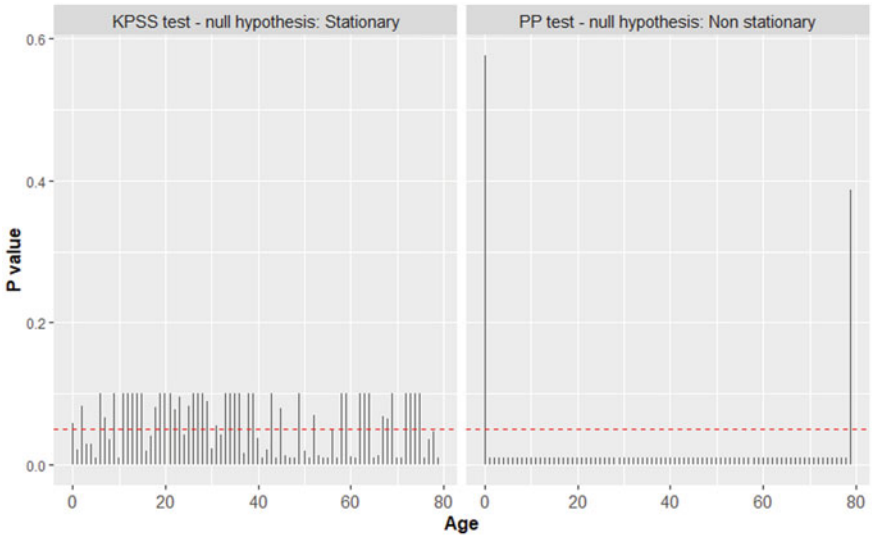


Fig. 6 Results of the unit root tests on the first differences of the log-mortality rates in England-Wales, 1850–2016 (age-specific *p* values) *Source* Own calculations on HMD data (https://www.mortality.org)

What we can conclude at this stage is that the hypothesis that the dynamics of the mortality rates in England-Wales and France over the period 1850–2016 are non-stationary holds (i.e., cannot be rejected). The question remains as to how many unit roots these series may have. To verify this, we repeated the unit root tests on the first differences of the log-rates of mortality (Figs. 6 and 7).

In both countries, the PP tests reject the null hypothesis of non-stationarity, meaning that the series are apparently consistent with the dynamics described by Eqs. (4) or (6). Conversely, the KPSS tests yield ambiguous results (Table 1): about 55% of them do not reject the null hypothesis of stationarity, but the remaining 45% do, either because there are multiple unit roots, or because the test cannot produce reliable results (e.g., because the series are too noisy).

In conclusion, the unit root tests performed on levels clearly indicate that the log-rates of mortality are non-stationary. However, the unit root tests performed on the first differences offer only partial support to the hypothesis of Eqs. (4) and (6): the possibility exists that the dynamics of mortality is more complex than these equations assume.

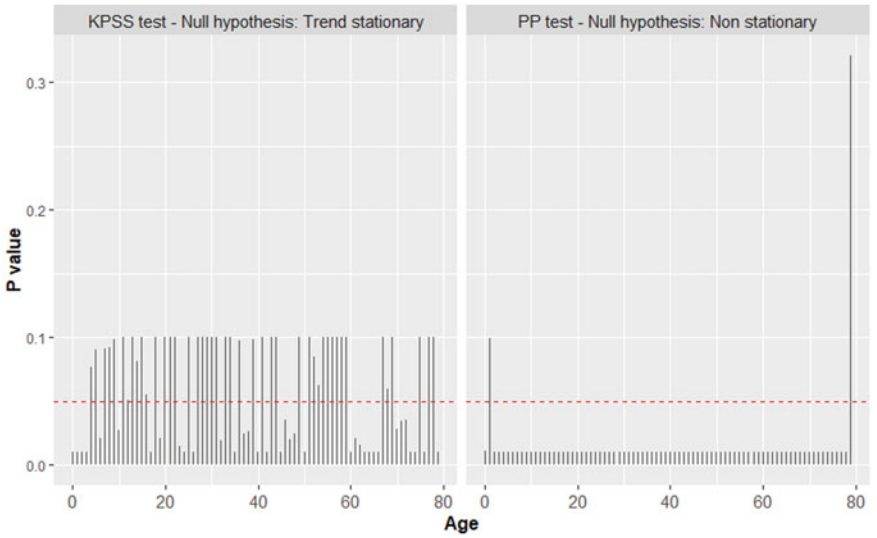


Fig. 7 Results of the unit root tests on the first differences of the log-mortality rates in France, 1850–2016 (age-specific *p* values) *Source* Own calculations on HMD data (https://www.mortality.org)

Table 1 Summary of the results of the unit root tests

Test	Series	Null hypothesis	No of series	England-wales: rejections	France: rejections
PP	Levels	Non stat	80	3	1
KPSS	Levels	Stat	80	70	80
PP	Diff	Non stat	80	78	78
KPSS	Diff	Stat	80	36	36

Source Own calculations on HMD data (https://www.mortality.org)

8 Summary and Conclusions

The main purpose of this paper is to prompt a debate on the TS or DS nature of demographic series. We showed, both theoretically and empirically, that the force of mortality probably evolves over time as a geometric random walk with drift (so, a DS process). If this is true, other series too must follow a non-stationary path, for instance person-years and survivors at age x in mortality tables, and survivors at age x in actual populations. For the empirical part, we carried out 160 tests on the series of the log-rates of mortality in France and in England-Wales (ages 0–79). These tests largely confirmed the DS nature of the series. To determine the order of integration of the process (number of unit roots) we repeated the tests on the differentiated series. In this case, however, we obtained ambiguous results: the PP tests support

our conjecture (the first differences of the log-mortality rates appear to be stationary, which is consistent with the hypothesis that the dynamics of the force of mortality follows a geometric random walk with drift), but this is less true with the KPSS tests (supportive only in 55% of the cases, i.e., ages).

Ultimately, the hypothesis that mortality follows an integrated process of order 1 remains the most plausible, but the possibility that the degree of integration is greater than 1 cannot be completely excluded, as yet.

Appendix

Let us derive Eq. (12), introduced, but not proved, in the main text.

Let us first introduce the concept of prospective force of mortality $\mu_t^*(x)$, which is analogous to its "traditional" version of Eq. (3) except that it refers to person years

$$\mu_t^*(x) = \lim_{\Delta x \to 0} \frac{L_t(x) - L_t(x + \Delta x)}{\Delta x L x} \tag{A.1}$$

and that deaths $d_{t,x}^*$ are defined accordingly (see Fig. 1 in the text).

The dynamic of the prospective force of mortality $\mu_t^*(x)$ is approximately equal to the geometric average of two adjacent "ordinary" forces of mortality $\mu_t(x)$:

$$\mu_t^*(x) \cong [\mu_t(x+1)\mu_{t-1}(x)]^{\frac{1}{2}} \tag{A.2}$$

Taking the log of Eq. (A.2) and applying Eq. (4) from the main text we get:

$$\ln[\mu_t^*(x)] = \ln(\delta) + \ln[\mu_{t-1}^*(x)] + W_t \tag{A.3}$$

where $W_t = \frac{1}{2}w_t + \frac{1}{2}w_{t-1}$.

Analogously to what we did in Eq. (7) for $l_t(x)$, we can now compute $L_t(x)$ from the prospective force of mortality:

$$L_t = e^{-M_t^*(x)} \tag{A.4}$$

where $M_t^*(x)$ is the cumulated prospective force of mortality. Substituting Eq. (A.3) into Eq. (A.4) and reorganizing (as in Eq. 10), we get Eq. (12):

$$\ln[L_t(x)] = \ln[L_{t-1}(x)]\delta W_t \tag{A.5}$$

References

1. Brillinger, D.R.: The natural variability of vital rates and associated statistics. Biometrics **42**, 693–734 (1986)
2. Dickey, D.A., Fuller, W.A.: Distribution of the estimators for autoregressive time series with a unit root. Econometrica **49**, 1057–1072 (1979)
3. Easterlin, R.: How beneficent is the market? a look at the modern history of mortality. Eur. Rev. Econ. Hist. **3**(3), 257–294 (1999)
4. Horiuchi, S., Wilmoth, J.R.: Deceleration in the age pattern of mortality at older ages. Demography **35**(4), 391–412 (1998)
5. Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y.: Testing the Null Hypothesis of stationarity against the alternative of a unit root. J. Econ. **54**, 159–178 (1992)
6. McKeown, T.: The Origins of Human Disease. Basil Blackwell/Oxford (1988)
7. Newey, W.K., West, K.: A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica **55**, 703–708 (1987)
8. Omran, A.R.: The epidemiologic transition: a theory of the epidemiology of population change. Milbank Mem.L Fund Q. **49**(4), 509–538 (1971)
9. Preston, S.H., Heuveline, P., Guillot, M.: Demography: Measuring and Modeling Population Processes. Blackwell/Oxford (2001)
10. Tuljapurkar, S., Li, N., Boe, C.: A universal pattern of mortality decline in the G7 countries. Nature **405**, 789–792 (2000)

Fuzzy and Model Based Clustering Methods: Can We Fruitfully Compare Them?

Alessio Serafini, Luca Scrucca, Marco Alfò, Paolo Giordani, and Maria Brigida Ferraro

Abstract During the last years, fuzzy and model-based approaches to clustering have received a great deal of attention and have been increasingly used in several empirical contexts. Even if they are very different from a theoretical point of view, they are similar in practice. In fact, model-based clustering gives posterior probabilities of component, treated as cluster, membership. Fuzzy clustering assigns observations to clusters through fuzzy membership degrees, while no probabilistic assumption is made to represent the clusters. The aim of this work is to compare the performance of some clustering methods belonging to the two approaches, in terms of recovering the true clusters, in a large scale simulation study.

Keywords Cluster analysis · Model-based clustering · Heuristic clustering · Mixture models · Fuzzy algorithms

1 Introduction

Unsupervised clustering techniques represent a set of methods that aim at discovering groups of similar observations, separated from other groups of observations without any a priori information about the underlying clustering structure. Several clustering

A. Serafini · L. Scrucca
Dipartimento di Economia, Università degli Studi di Perugia, Perugia, Italy
e-mail: alessio.serafini@unipg.it

L. Scrucca
e-mail: luca.scrucca@unipg.it

M. Alfò · P. Giordani (✉) · M. B. Ferraro
Dipartimento di Scienze Statistiche, Sapienza Università di Roma, Rome, Italy
e-mail: paolo.giordani@uniroma1.it

M. Alfò
e-mail: marco.alfo@uniroma1.it

M. B. Ferraro
e-mail: mariabrigida.ferraro@uniroma1.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_19

techniques have been proposed during the past decades, with a growing interest in many application areas. Those techniques that use probabilistic assumptions about the data generating mechanism are referred to as *model-based* clustering methods, while methods that make no explicitly probabilistic assumptions are called *heuristic* clustering algorithms. In model-based clustering, data are assumed to be generated by a mixture of probability distributions, where each component in the mixture describes a cluster [20]. Once the model is fitted, each unit may be allocated to the component/cluster associated to the maximum posterior probability of component membership. In heuristic clustering methods, groups are no longer represented in terms of parametric distributions but rather by centres, and observations are assigned to a cluster by some specific procedure or criterion. These methods include the k-Means algorithm [30] and the so-called *fuzzy* algorithms [41]. In the latter family of algorithms, observations are assigned to one of the clusters according to the maximum membership degree, defined as a function of the distance from the specific centre, attaining values in the range [0, 1]. Each unit may belong to several clusters, and, in this sense, such algorithms are often referred to as *soft* clustering methods. The softness of the partition is achieved by means of the concept of fuzziness [41].

Although these two classes of methods share the same aim, i.e. clustering observations in homogeneous groups, they are conceptually different. Both posterior probabilities and membership degrees take values in [0, 1], but the membership degree cannot be interpreted as a measure of probability because no random generation process is assumed and, conversely, the posterior probability is not related to the fuzziness of the partition. Nevertheless, as shown in [25], there exists a strong connection between these two classes of methods. In fact, the EM algorithm can be defined as a particular coordinate descent algorithm minimising a loss function equal to the negative sum of the entropies of the posterior probabilities plus a weighted sum of probabilistic distances between observations and mixture components. The corresponding heuristic clustering method [29] can be easily derived by replacing the probabilistic distances with non-probabilistic distances, such as the Euclidean one.

Our aim is to compare these methods and provide some guidelines for model/method selection by a simulation study. To our knowledge, such a comparison between model-based and fuzzy clustering methods has never been carried out, but for a few limited cases, see [7, 14]. The simulation study considers finite mixture models with Gaussian [20], t [36] and generalised hyperbolic component-specific densities [8]. These mixture models are recalled in Sect. 2. Moreover, we consider the Fuzzy k-Means algorithm [6] and the Gustafson-Kessel variant of the Fuzzy k-Means [4, 24]. These algorithms are described in Sect. 3. Section 4 contains the simulation study comparing the selected model-based and fuzzy clustering methods. A final discussion can be found in Sect. 5.

2 Model-Based Clustering

In model-based clustering, observed data are assumed to be generated by a statistical model. Let $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) \in \mathbb{R}^p$ be a sample of n independent observations, where p denotes the number of variables. Each vector $\boldsymbol{x}_i$ is drawn from a density that describes the population of interest; the density is defined by a finite mixture of probability density functions:

$$f(\boldsymbol{x}_i; \boldsymbol{\Theta}) = \sum_{k=1}^{g} \pi_k f_k(\boldsymbol{x}_i|\boldsymbol{\Psi}_k), \tag{1}$$

where π_k are the mixing proportions, $k = 1, \ldots, g$, such that $\pi_k > 0$, and $\sum_{k=1}^{g} \pi_k = 1$, g is the number of mixture components, $f_k(\boldsymbol{x}|\boldsymbol{\Psi}_k)$ is the density of the kth mixture component, and $\boldsymbol{\Theta} = \{\pi_1, \pi_2, \ldots, \pi_{g-1}, \boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \ldots, \boldsymbol{\Psi}_g\}$ is the global set of model parameters [31]. Each component-specific density may belong to a different parametric class and represents a cluster. In the following we will proceed by assuming $f_k(\boldsymbol{x}|\boldsymbol{\Psi}_k) = f(\boldsymbol{x}|\boldsymbol{\Psi}_k), \quad \forall k = 1, \ldots, g$, which is the somewhat standard way to proceed in these contexts.

The log-likelihood for a mixture model is

$$\ell(\boldsymbol{\Theta}) = \log \prod_{i=1}^{n} \sum_{k=1}^{g} \pi_k f(\boldsymbol{x}_i|\boldsymbol{\Psi}_k) = \sum_{i=1}^{n} \log \sum_{k=1}^{g} \pi_k f(\boldsymbol{x}_i|\boldsymbol{\Psi}_k).$$

Maximum likelihood estimation of model parameters is usually pursued by applying the well-known Expectation-Maximisation (EM) algorithm [15]. This is an iterative procedure for maximising the expected value of the complete data log-likelihood:

$$\ell_c(\boldsymbol{\Theta}) = \log \prod_{i=1}^{n} f(\boldsymbol{x}_i, \boldsymbol{z}_i|\boldsymbol{\Theta}) = \sum_{i=1}^{n} \{\log p(\boldsymbol{z}_i|\boldsymbol{\Theta}) + \log f(\boldsymbol{x}_i|\boldsymbol{z}_i, \boldsymbol{\Theta})\},$$

where $\boldsymbol{z}_i = (z_{i1}, \ldots, z_{ig})$ is a latent variable with $z_{ik} = 1$ if the ith observation $\boldsymbol{x}_i$ belongs to component k and $z_{ik} = 0$ otherwise. The EM algorithm is briefly described in Algorithm 1.

Algorithm 1 EM algorithm for fitting finite mixture models

initialisation: set starting values for the unknown parameters $\boldsymbol{\Theta}$
repeat
  E-step: estimate $\mathbb{E}(z_i)$ for each observation given $\boldsymbol{\Theta}$
  M-step: estimate $\boldsymbol{\Theta}$ by maximising $\ell_c(\boldsymbol{\Theta})$
until convergence

2.1 Finite Mixtures of Gaussian Densities

Due to its flexibility and mathematical tractability, the most popular model-based clustering approach postulates that, within cluster k, the data follow a Gaussian distribution (GMM), i.e. $f(\boldsymbol{x}_i, |\boldsymbol{\Psi}_k) \equiv \phi(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ where $\phi(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is the density of a multivariate Gaussian distribution with mean vector $\boldsymbol{\mu}_k$ and covariance matrix $\boldsymbol{\Sigma}_k$ [20]. The density of the GMM is then given by

$$f(\boldsymbol{x}_i; \boldsymbol{\Theta}) = \sum_{k=1}^{g} \pi_k \phi(\boldsymbol{x}_i | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \tag{2}$$

where $\boldsymbol{\Theta} = \{\pi_1, \pi_2, \ldots, \pi_{g-1}, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_g\}$ is the global set of parameters for the model in (2), and $\phi(\boldsymbol{x}_i | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is the underlying component-specific density function. Thus, the model in (2) generates ellipsoidal clusters centred at the mean $\boldsymbol{\mu}_k$, with $\boldsymbol{\Sigma}_k$ controlling the shape and orientation of each cluster. Parsimonious parameterisations of the cluster covariance matrices can be obtained through the eigen-decomposition [5, 10]

$$\boldsymbol{\Sigma}_k = \lambda_k \boldsymbol{D}_k \boldsymbol{A}_k \boldsymbol{D}_k, \tag{3}$$

where λ_k is a scalar controlling the *global* volume of the ellipsoid, $\boldsymbol{A}_k$ is a diagonal matrix controlling its shape, $\boldsymbol{D}_k$ is a columnwise orthogonal matrix controlling the orientation of the ellipsoid. By suitable constraining the quantities in (3), a class of 14 flexible models with different geometrical properties can be derived. Estimation for any of these models is implemented in the R package **mclust** [38]. The number of clusters and the parameterisation of the covariance matrices can be selected using model selection criteria, such as the Bayesian information criterion (BIC) [19, 37].

2.2 Finite Mixtures of t Densities

A heavy-tailed alternative for the component specific density in (2) involves the use of the multivariate t distribution [36]. The density of the resulting t mixture model (tMM) has the following form:

$$f(\boldsymbol{x}_i; \boldsymbol{\Theta}) = \sum_{k=1}^{g} \pi_k f_t(\boldsymbol{x}_i | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \nu_k),$$

where $\boldsymbol{\Theta} = \{\pi_1, \pi_2, \ldots, \pi_{g-1}, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_g, \nu_1, \ldots, \nu_g\}$ is the set of parameters and $f_t(\boldsymbol{x}_i | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \nu_k)$ denotes the density of the multivariate t distribution with mean $\boldsymbol{\mu}_k$, scale matrix $\boldsymbol{\Sigma}_k$, and ν_k degrees of freedom. Based on the eigen-decomposition for the Gaussian case in (3), and by constraining the degrees of

freedom to be equal/unequal among groups, a parsimonious class of mixture models with t-distributed components, called tEIGEN, can be defined [1]. This family of models is implemented in the R package **teigen** [3]. A variant of the classic EM algorithm for mixture models proposed by [33] called ECM is employed for maximum likelihood estimation of tMM [2]. Such algorithm replaces the M-step with a sequence of conditional maximisation substeps. As for the GMMs case, BIC is employed to select the number of components, the parameterisation of the scale matrices, and the degrees of freedom [1].

2.3 Finite Mixtures of Generalised Hyperbolic Densities

Generalised hyperbolic mixture models (GHMMs) has been introduced in [8] who propose finite mixture models with non-elliptical component-specific densities as an alternative to GMMs and tMMs. The GHMM density is formalised as

$$f(\boldsymbol{x}_i; \boldsymbol{\Theta}) = \sum_{k=1}^{g} \pi_k f_{\mathcal{G}}(\boldsymbol{x}_i | \lambda_k, \chi_k, \psi_k, \boldsymbol{\mu}_k, \boldsymbol{\Delta}_k, \alpha_k),$$

where $f_{\mathcal{G}}(\boldsymbol{x}_i | \lambda_k, \chi_k, \psi_k, \boldsymbol{\mu}_k, \boldsymbol{\Delta}_k, \alpha_k)$ is the density of a generalised hyperbolic distribution with location vector $\boldsymbol{\mu}_k$, scale matrix $\boldsymbol{\Delta}_k$, skewness parameter α_k, index parameter λ_k and concentration parameters χ_k and ψ_k. The EM algorithm can again be employed for parameters estimation and BIC is used for parameter estimation. GHMMs are implemented in the R package **MixGHD** [39].

3 Fuzzy Clustering

As opposed to model-based clustering methods, heuristic algorithms for clustering do not require any probabilistic assumption. One of the most popular heuristic clustering method is probably the k-Means (kM) algorithm [30]. This algorithm partitions the n observations into g groups by solving the following optimisation problem:

$$\begin{aligned} \min_{\boldsymbol{U}, \boldsymbol{H}} J_{kM} &= \sum_{i=1}^{n} \sum_{k=1}^{g} u_{ik} d^2(\boldsymbol{x}_i, \boldsymbol{h}_k), \\ \text{s.t.} \quad & u_{ik} \in \{0, 1\}, \quad i = 1, \ldots, n, \quad k = 1, \ldots, g, \\ & \sum_{k=1}^{g} u_{ik} = 1, \quad i = 1, \ldots, n, \end{aligned} \tag{4}$$

where the generic element of the matrix $\boldsymbol{U}$ of size $(n \times g)$, u_{ik}, denotes the membership of observation i to cluster k, i.e. $u_{ik} = 1$ if the observation i belongs to cluster k

and $u_{ik} = 0$ otherwise. Moreover, $\boldsymbol{h}_k = [h_{k1}, \ldots, h_{kp}]$, the kth row of the prototype matrix $\boldsymbol{H}$ of size $(g \times p)$, is the prototype for cluster k $(k = 1, \ldots, g)$, and $d^2(\boldsymbol{x}_i, \boldsymbol{h}_k)$ is the squared Euclidean distance between observation i and the kth prototype. Thus, the loss function to be minimised represents the within-cluster sum of squares. In this clustering algorithm, each observation is assigned to the closest prototype, guaranteeing that each observation belongs to one and only one cluster. Coordinate descent algorithms can be adopted for finding the optimal partition. The number of clusters can be selected by several methods. A popular choice is the Silhouette (S) index [28].

Fuzzy clustering methods have been introduced to handle situations where the membership of an observation to a cluster is not clearcut [41]. The units are assigned to clusters according to a so-called fuzzy membership degree, taking values in [0, 1]. This is inversely related to the dissimilarities between the units and the cluster prototypes. A membership degree close to 1 implies that a data point is close to the corresponding prototype, and therefore it can be assigned to the corresponding cluster. In the literature, several fuzzy clustering methods exist. Among them, the most common is the Fuzzy k-Means (FkM) algorithm. The following subsections briefly review the FkM algorithm and the closely related Gustafson-Kessel variant (GK-FkM). These algorithms are implemented in the R package **fclust** [16, 17].

3.1 Fuzzy k-Means

The FkM clustering algorithm [6] aims at grouping n units into g clusters by solving the following constrained optimisation problem:

$$\begin{aligned} \min_{\boldsymbol{U},\boldsymbol{H}} J_{FkM} &= \sum_{i=1}^{n}\sum_{k=1}^{g} u_{ik}^m d^2(\boldsymbol{x}_i, \boldsymbol{h}_k), \\ \text{s.t.} \quad & u_{ik} \in [0, 1], \quad i = 1, \ldots, n, \quad k = 1, \ldots, g, \\ & \sum_{k=1}^{g} u_{ik} = 1, \quad i = 1, \ldots, n. \end{aligned} \tag{5}$$

In this formulation, the matrices $\boldsymbol{H}$ and $\boldsymbol{U}$ have the same meaning as in (4), except for the different constraints on the elements in $\boldsymbol{U}$. In this case, the loss function measures the within-cluster sum of squares in a fuzzy setting, taking into account the membership degree information. The fuzziness parameter $m > 1$ tunes the level of fuzziness of the resulting partition. The higher the value of m, the fuzzier the partition with membership degrees tending to $1/g$ as $m \to \infty$. When m is close to 1, the FkM solution approaches that of kM in (5). The standard choice is $m = 2$ [35]. The solution of (5) is obtained through an iterative optimisation algorithm briefly described in Algorithm 2.

Algorithm 2 FKM algorithm

initialisation: set starting values for the membership degree matrix $\boldsymbol{U}$
repeat
  update of $\boldsymbol{H}$: solve the problem in (5) with respect to $\boldsymbol{H}$
  update of $\boldsymbol{U}$: solve the problem in (5) with respect to $\boldsymbol{U}$
until convergence

To select the optimal number of clusters, several cluster validity indexes can be adopted. Among them, we mention the Fuzzy Silhouette (FS) index [9] extending the S index to the fuzzy setting by considering the membership degree information.

3.2 The Gustafson-Kessel Variant of FkM

The FkM algorithm yields spherical clusters and it may be inadequate when clusters have different geometrical shapes. The Gustafson-Kessel variant of the FkM (GK-FkM) [24] is designed to solve this issue. GK-FkM uses a Mahalanobis-type dissimilarity:

$$d_M^2(\boldsymbol{x}_i, \boldsymbol{h}_k) = (\boldsymbol{x}_i - \boldsymbol{h}_k)^\top \boldsymbol{M}_k (\boldsymbol{x}_i - \boldsymbol{h}_k), \tag{6}$$

where $\boldsymbol{M}_k$ is a symmetric and positive-definite matrix. The GK-FkM algorithm can then be formulated as

$$\begin{aligned} \min_{\boldsymbol{U},\boldsymbol{H},\boldsymbol{M}_1,\ldots,\boldsymbol{M}_g} J_{GK-FkM} &= \sum_{i=1}^{n}\sum_{k=1}^{g} u_{ik}^m d_M^2(\boldsymbol{x}_i, \boldsymbol{h}_k), \\ \text{s.t.} \quad & u_{ik} \in [0,1], \quad i = 1,\ldots,n, \quad k = 1,\ldots,g, \\ & \sum_{k=1}^{g} u_{ik} = 1, \quad i = 1,\ldots,n, \\ & |\boldsymbol{M}_k| = \rho_k > 0, \quad k = 1,\ldots,g, \end{aligned} \tag{7}$$

where the last set of constraints avoids the trivial solution with $\boldsymbol{M}_k = \boldsymbol{0}, k = 1,\ldots,g$. Note that the most common choice is $\rho_k = 1$ for all $k = 1,\ldots,g$.

An iterative solution of the optimisation problem in (7) can be found by an iterative optimisation algorithm. To avoid numerical problems, a computational improvement has been proposed in [4] where the condition number of $\boldsymbol{M}_k$ is constrained to be higher than a pre-specified threshold. Note that this condition is similar to that imposed to covariance matrices in finite mixture models with either Gaussian or t-components [23, 27].

4 Simulation Study

We investigate the finite sample performance of the previously described clustering methods by means of a large simulation study, trying to understand in which conditions a method tends to perform reasonably well and when it fails to recover the clustering structure. In the following, we first describe the set-up of the simulation study, and then we report the obtained results.

Clustering methods are compared using the Adjusted Rand Index (ARI) [26], a popular measure of agreement between two partitions, one estimated by a statistical procedure independent of the labelling of the groups, and one being the true classification. The ARI has zero expected value in the case of a random partition, and it achieves the value of 1 in the case of perfect agreement between two partitions. An extension of the ARI has been recently proposed by [18] to incorporate soft clustering information, as produced by both model-based clustering and fuzzy clustering. Finally, we mention that several other criteria for clustering comparison are available; for a review see [32].

4.1 Simulation Set-Up

Synthetic data are generated from two different mixture model scenarios, one with a Gaussian (MVN), and one with t_{10} component-specific density (MVt10). Varying number of components, ($g = 3$ and $g = 4$), sample sizes, ($n = 200$ and $n = 1000$) and $p = 4$ variables are considered. The aim is to evaluate model/method performance in recovering the true cluster structure when component-specific densities are elliptical, with potentially heavier tails than a MVN density and with a possibly reduced available information (small sample size). This last argument plays a potentially fundamental role as we know that ML estimates for the covariance matrices may be unstable in this case.

To evaluate whether a model is sensitive to the size of each cluster, the mixing probabilities are chosen to have either a balanced

$$\boldsymbol{\pi} = (1/3, 1/3, 1/3), \qquad \text{and} \qquad \boldsymbol{\pi} = (1/4, 1/4, 1/4, 1/4),$$

or an unbalanced structure

$$\boldsymbol{\pi} = (0.5, 0.4, 0.1), \qquad \text{and} \qquad \boldsymbol{\pi} = (0.4, 0.3, 0.2, 0.1),$$

for $g = 3$ and $g = 4$, respectively. The clustering algorithms discussed in Sects. 2 and 3 are based on specific, different, shapes for the clusters. The FkM, as the classical kM, adopts spherical cluster shapes, with the same orientation and volume, whereas GK-FkM adopts a more general elliptical shapes with common volume. On the other hand, different parameterisations of the covariance matrices lead to clusters

with different shapes, volumes and orientations. Since these geometric characteristics may influence the final classification, we start by considering spherical clusters with diagonal covariance structures ($\boldsymbol{\Sigma} = \mathbf{I}$). Then, we consider ellipsoidal shapes defined by the eigen-decomposition in (3). For $g = 3$ and $p = 4$ we use:

$$\lambda_1 = 0.2, \quad \boldsymbol{A}_1 = \begin{pmatrix} 0.5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{D}_1 = \begin{pmatrix} 1.5 & 1 & 1 & 1 \\ 1 & 1.5 & 1 & 1 \\ 1 & 1 & 1.5 & 1 \\ 1 & 1 & 1 & 1.5 \end{pmatrix};$$

$$\lambda_2 = 0.05, \quad \boldsymbol{A}_2 = \begin{pmatrix} 1.5 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 4.5 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix} \quad \text{and} \quad \boldsymbol{D}_2 = \begin{pmatrix} 1.6 & -1 & 1 & -1 \\ -1 & 1.6 & -1 & -1 \\ 1.2 & -1 & 1.6 & -1 \\ -1 & -1 & -1 & 1.6 \end{pmatrix};$$

$$\lambda_3 = 1, \quad \boldsymbol{A}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \boldsymbol{D}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

while for $g = 4$ and $p = 4$ we consider the following:

$$\lambda_1 = 0.2, \quad \boldsymbol{A}_1 = \begin{pmatrix} 0.5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{D}_1 = \begin{pmatrix} 1.5 & 1 & 1 & 1 \\ 1 & 1.5 & 1 & 1 \\ 1 & 1 & 1.5 & 1 \\ 1 & 1 & 1 & 1.5 \end{pmatrix};$$

$$\lambda_2 = 0.05, \quad \boldsymbol{A}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \quad \text{and} \quad \boldsymbol{D}_2 = \begin{pmatrix} 1.6 & -1 & 1 & -1 \\ -1 & 1.6 & -1 & -1 \\ 1.2 & -1 & 1.6 & -1 \\ -1 & -1 & -1 & 1.6 \end{pmatrix};$$

$$\lambda_3 = 1, \quad \boldsymbol{A}_3 = \begin{pmatrix} 0.2 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.7 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \boldsymbol{D}_3 = \begin{pmatrix} 0.9 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.9 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.9 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.9 \end{pmatrix};$$

$$\lambda_4 = 1, \quad \boldsymbol{A}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \boldsymbol{D}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Given the component-covariance structures, the mean vectors can be used to arrange the clusters in the feature space, allowing for different degrees of separation between clusters. We design clusters lying along diagonal and scattered regions in the feature space, well separated clusters and partially overlapping clusters. The

mean vectors are presented as a $p \times g$ matrix, with the means for each cluster along the columns. For clusters aligned along the diagonal, the means for well separated clusters are, when $g = 3$ and $g = 4$, respectively:

$$\boldsymbol{\mu} = \begin{pmatrix} 0 & -7 & 7 \\ 0 & -7 & 7 \\ 0 & -7 & 7 \\ 0 & -7 & 7 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\mu} = \begin{pmatrix} 0 & -7 & 7 & 7 \\ 0 & -7 & 7 & -7 \\ 0 & -7 & 7 & 7 \\ 0 & -7 & 7 & -7 \end{pmatrix}.$$

For partially overlapping clusters, the means are:

$$\boldsymbol{\mu} = \begin{pmatrix} -3.5 & 0 & 3.5 \\ -3.5 & 0 & 3.5 \\ -3.5 & 0 & 3.5 \\ -3.5 & 0 & 3.5 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\mu} = \begin{pmatrix} -3.5 & 0 & 3.5 & 3 \\ -3.5 & 0 & 3.5 & -3 \\ -3.5 & 0 & 3.5 & 3 \\ -3.5 & 0 & 3.5 & -3 \end{pmatrix},$$

For scattered clusters we consider four different configurations. If clusters are well separated, we have

$$\boldsymbol{\mu} = \begin{pmatrix} 0 & 7 & 4 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\mu} = \begin{pmatrix} 0 & 0 & 7 & 7 \\ 0 & 7 & 7 & -2 \\ 0 & 7 & 7 & 3 \\ 0 & 7 & 7 & -3 \end{pmatrix},$$

in the case of diagonal covariance matrices and

$$\boldsymbol{\mu} = \begin{pmatrix} 0 & 7 & 7 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\mu} = \begin{pmatrix} 0 & 0 & 7 & 7 \\ 0 & 7 & 7 & -2 \\ 0 & 7 & 7 & 3 \\ 0 & 7 & 7 & -3 \end{pmatrix},$$

in the case of full covariance matrices. For partially overlapping clusters the mean vectors are defined as

$$\boldsymbol{\mu} = \begin{pmatrix} 0 & 3.5 & 3.5 \\ 0 & 0 & 3.5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\mu} = \begin{pmatrix} 0 & -3 & 3 & 3 \\ 0 & -3 & 3 & -3 \\ 0 & -3 & 3 & 3 \\ 0 & -3 & 3 & -3 \end{pmatrix}.$$

for diagonal, while for full covariance matrices we set:

$$\boldsymbol{\mu} = \begin{pmatrix} 0 & 3.5 & 3.5 \\ 0 & 0 & 3.5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\mu} = \begin{pmatrix} 0 & -3.5 & 3.5 & 3 \\ 0 & -3.5 & 3.5 & -3 \\ 0 & -3.5 & 3.5 & 3 \\ 0 & -3.5 & 3.5 & -3 \end{pmatrix}.$$

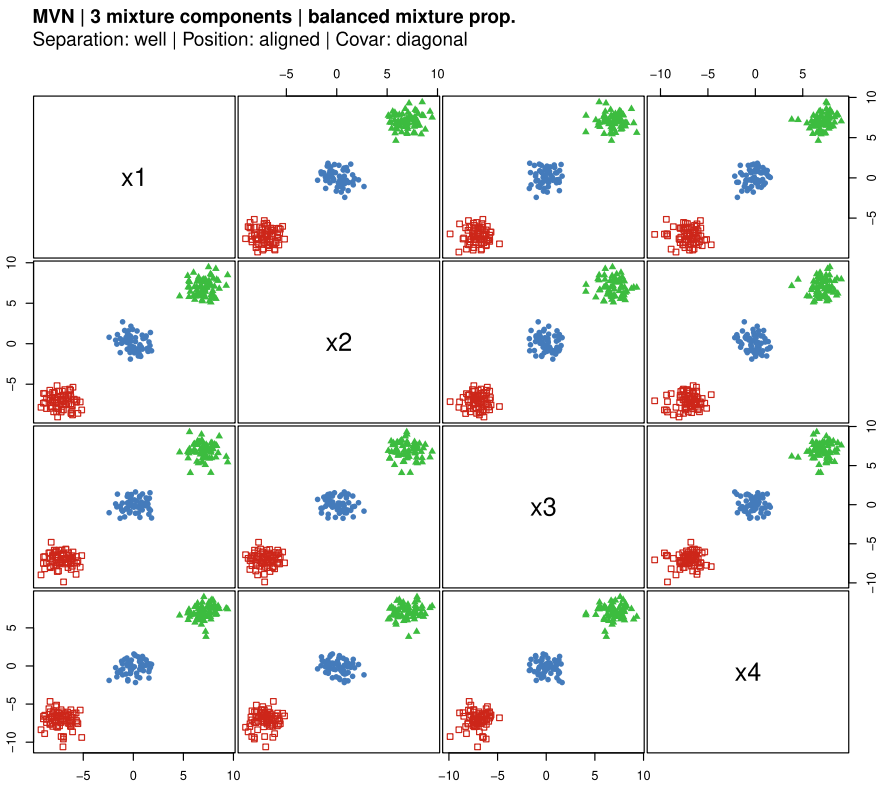


Fig. 1 A dataset used in the simulation study. Characteristics of the synthetic data are shown in the plot title

The `sim()` function, available in the **mclust** package, is used to generate the datasets following all the possible combinations of parameters presented above. For each of these cases, we generate 100 datasets. Examples of two generated datasets (case $g = 3$) are reported in Figs. 1 and 2. By looking at the figures, we may notice that the different scenarios present different levels of complexity moving from simpler cases, such as the one in Fig. 1, to more complex ones, such as the one in Fig. 2.

4.2 Results

Results from the simulation study are summarised in Figs. 3, 4, 5, 6, 7, 8, 9, 10, where each graph shows the boxplots of the Adjusted Rand Index (ARI) for every data generating mechanism, number of mixture components and type of prior mixture probabilities, as discussed in Sect. 4.1.

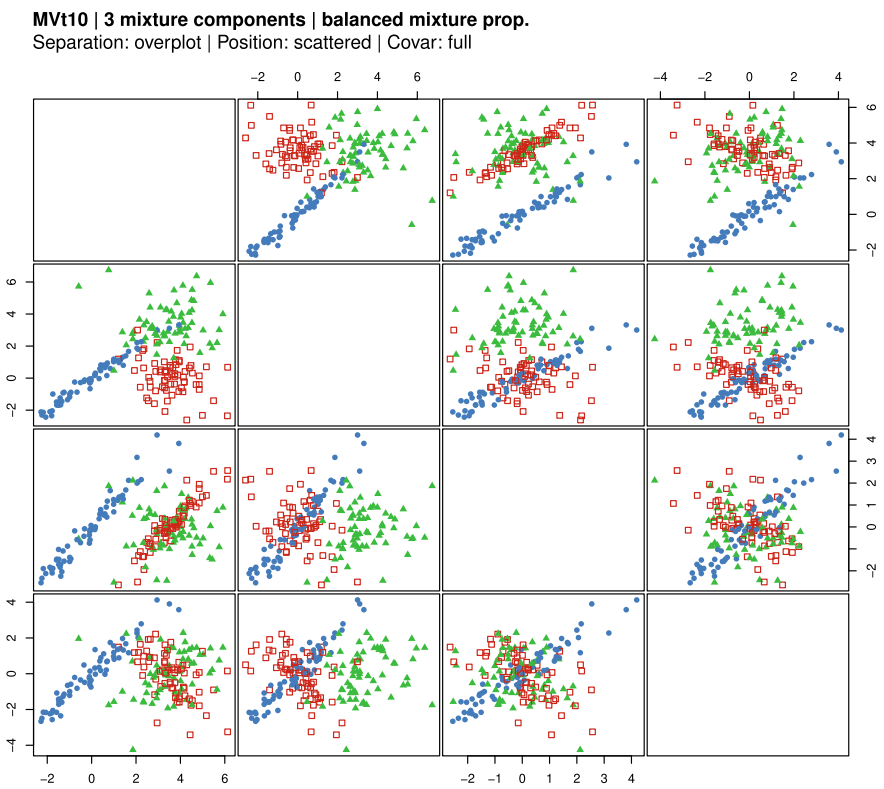


Fig. 2 A dataset used in the simulation study. Characteristics of the synthetic data are shown in the plot title

The clustering methods are applied by either using the true value for g or selecting the optimal number of clusters with model selection criteria. Specifically, FS and BIC have been adopted for, respectively, the fuzzy and the model-based clustering algorithms. Moreover, S is used to select the optimal number of clusters for the kM algorithm. Note that, when the true number of clusters is set, BIC is used to select the parameterisation of the within-cluster covariance matrix for GMM and tMM.

In general, we can observe that mixture models with Gaussian and t-distributed components tend to show a good classification performance. This is likely related to the fact that the adoption of parsimonious covariance structures may generate models that fit the shape of clusters. In fact, as we shall see, the observed differences in performance are essentially related to some specific scenarios. This especially holds for the FkM and kM methods that show approximately the same classification performance, whilst, for GK-FkM, the results appear very poor in terms of clustering accuracy. Although GK-FkM adopts a more general clustering structure, the need to estimate more parameters leads to poor performance in the simplest cases. This appears to be the case also for GHMM.

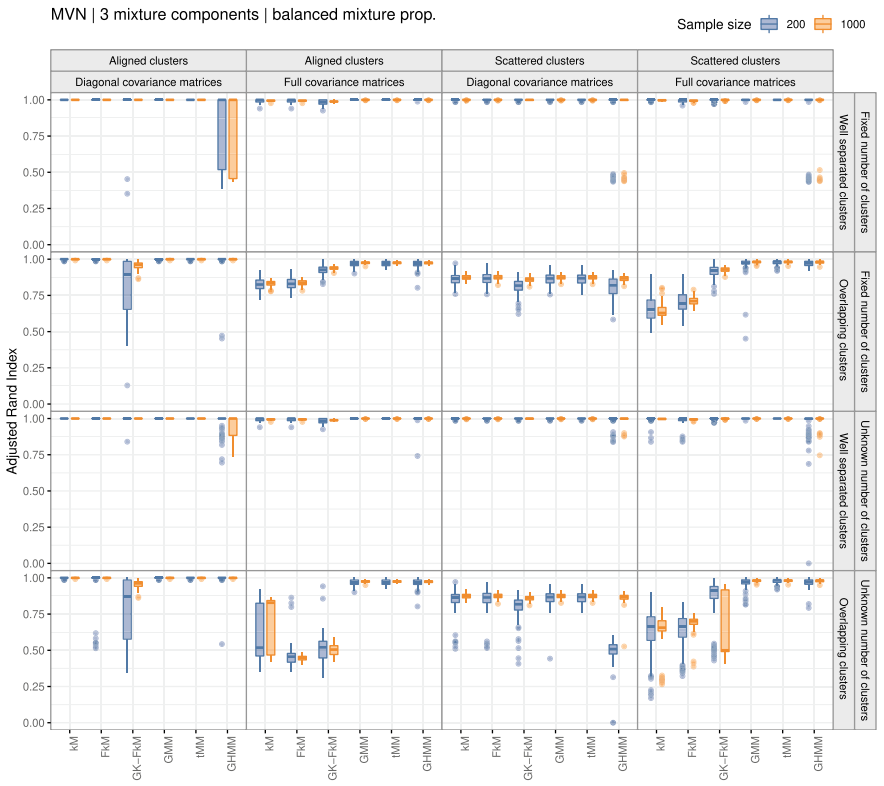


Fig. 3 Boxplots of ARI for each pair of configurations (GMM case, $g = 3$, balanced clusters)

Results for $g = 3$ and $g = 4$ look quite similar, indicating that, for each configuration, the number of clusters is not a relevant issue. The most relevant differences between model-based, in particular GMM and tMM, and heuristic methods, in particular FkM and kM, are observed in the cases where the clusters are partially overlapped with full covariance matrices. This occurs because the covariance matrix information plays a relevant role in distinguishing the clusters. In such scenarios, the GK-FkM method works reasonably well, but fails in the cases where clusters have diagonal covariance matrices.

GMM and tMM present more or less the same classification performance, even when the *true* data generating process (in terms of the component-specific density) is different from the one adopted by each mixture model. Variability of the final classification seems to be the main difference when the MVt10 model is used for simulating the data. For the GHMM we can observe a worse overall performance compared to both GMM and tMM models. The results showed that the GHMM approach is able to achieve a good classification performance when the data generation process fol-

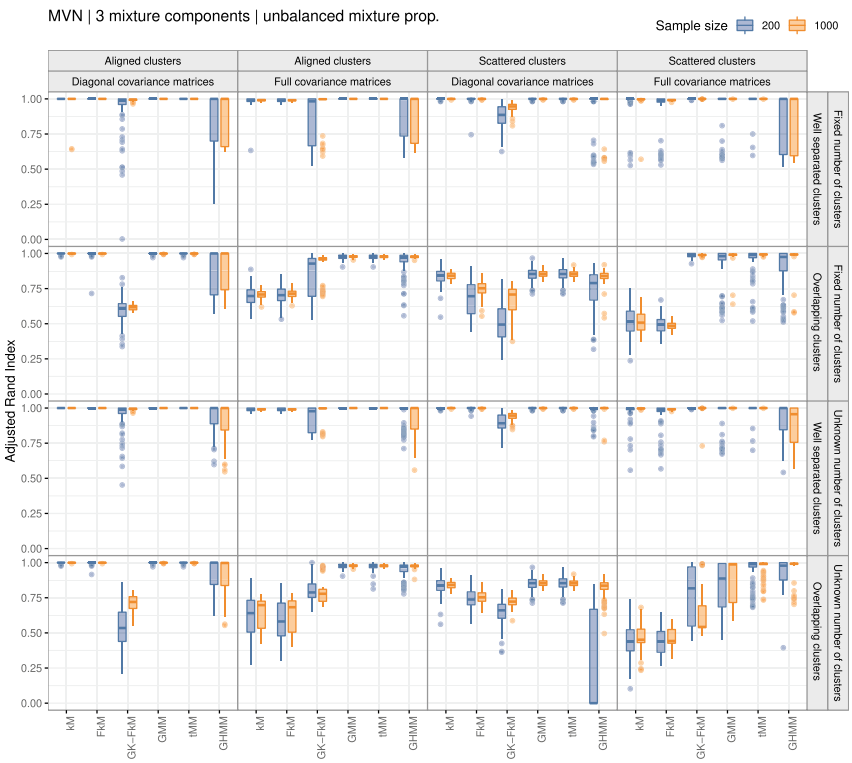


Fig. 4 Boxplots of ARI for each pair of configurations (GMM case, $g = 3$, unbalanced clusters)

lows an asymmetric distribution, but it is unsatisfactory otherwise. The performance of the heuristic methods are rather stable with respect to the data generating process showing that there is no need to determine the appropriate probability model to use.

5 Final Discussion

A growing use of clustering methods, in particular model-based and fuzzy clustering methods, has suggested the need for a thorough investigation of the differences and similarities between these two approaches in terms of identifying the *true* clustering structure. This has been done by an extensive simulation study. We have observed that mixture models, in particular those with Gaussian and t-distributed components, tend to provide uniformly good clustering performance. In contrast, the most common heuristic clustering methods, i.e., FkM and kM, work in a satisfactory way, but tend to fail when clusters are jointly overlapped and non-spherical.

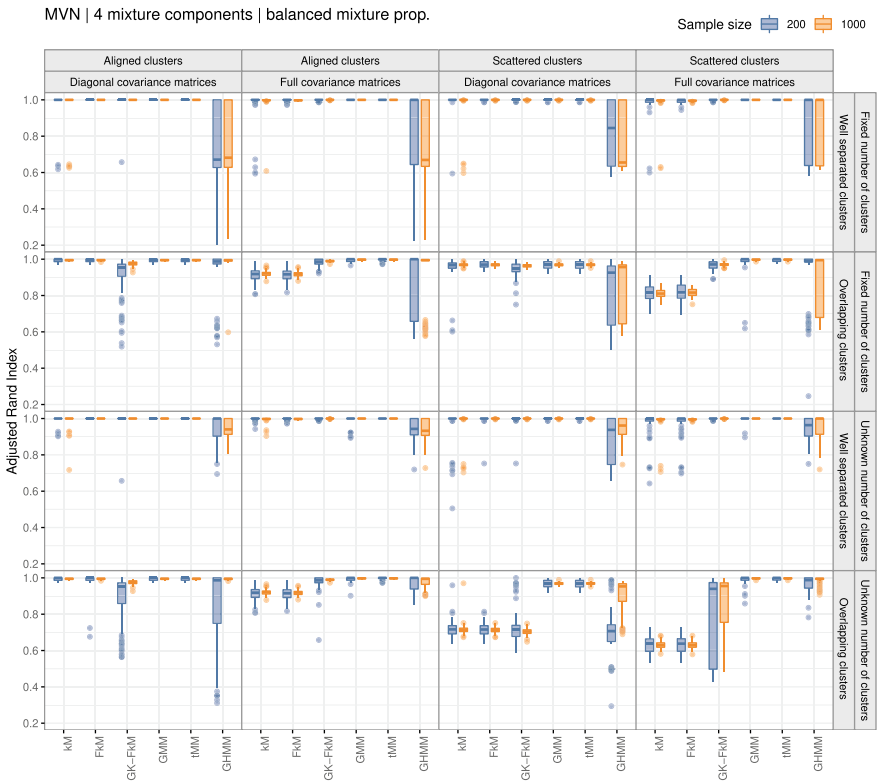


Fig. 5 Boxplots of ARI for each pair of configurations (GMM case, $g = 4$, balanced clusters)

When the underlying data generating process produces non-elliptical clusters, mixture models that assume symmetrical clusters (e.g. Gaussian and t-distributed components) may show poor classification performance. In this case, heuristic methods, such as k-Means and Fuzzy k-Means, appear to be more robust because they are less dependent on distributional assumptions. Conversely, when clusters are overlapped and essentially elliptical, the flexibility of the parsimonious decomposition of the covariance matrices for GMMs and tMMs allows to better fit the different cluster shapes.

The flexibility of GMMs and tMMs is achieved at the cost of a higher number of parameters with respect to FkM and kM. As we saw, these algorithms assign the observations to the clusters according to their distances with respect to the cluster means. No information provided by the cluster covariance matrices is used. Therefore, on the one hand, the number of free parameters is extremely low implying that FkM and kM are less affected from computational problems when handling large-sized data. On the other hand, when cluster memberships are essentially related to the different configurations of the cluster covariance matrices, in particular when

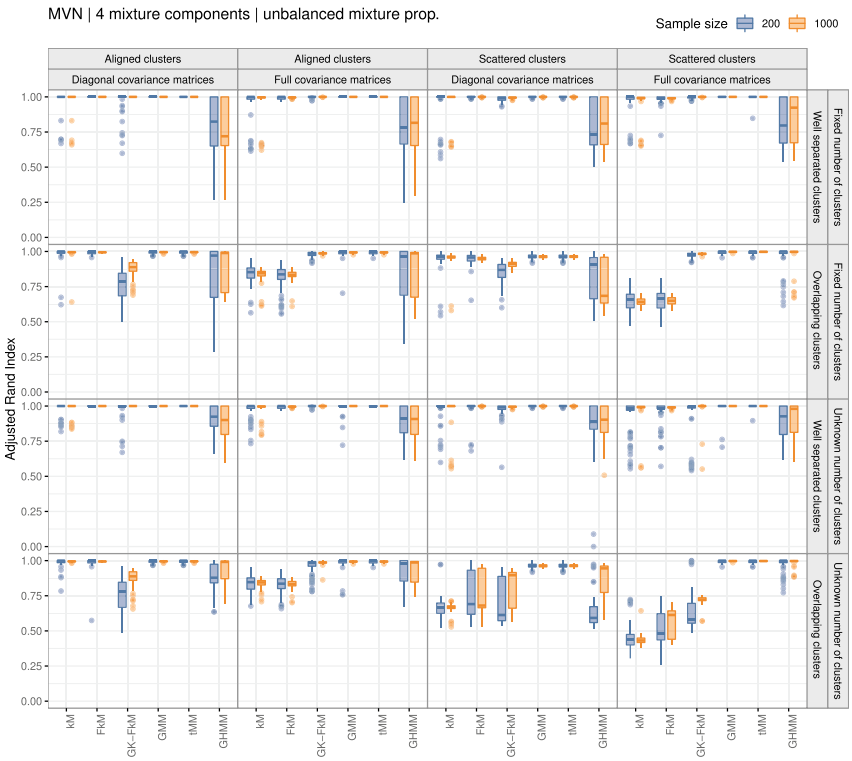


Fig. 6 Boxplots of ARI for each pair of configurations (GMM case, $g = 4$, unbalanced clusters)

clusters are overlapped, the performance of F*k*M and *k*M get worse because they tend to fail in identifying the clusters.

The results we have found should be seen as the starting point of a wider comparison between model-based and fuzzy clustering methods. Such a wider comparison can be done along various directions. The first involves the selection of the scenarios in the current simulation study. Datasets have been generated according to mixture models. It is trivial that model-based clustering methods assuming the *true* distribution work in a satisfactory way. Nevertheless, our interest was to assess its performance in comparison with those of the other (over- and under-parameterised) clustering methods. It will be interesting to see what happens in a new simulation study where datasets are randomly generated in a *non-parametric* way assessing whether and how the performance of model-based and fuzzy clustering methods differ. For instance, the set-up of the new simulation study might take inspiration from the seminal work by Milligan and Cooper [34].

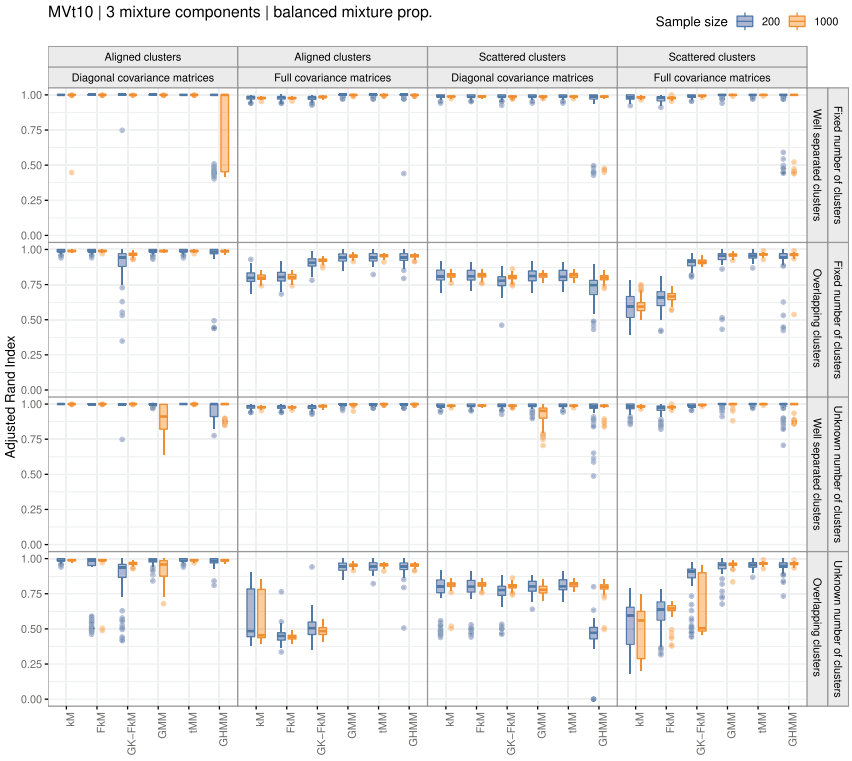


Fig. 7 Boxplots of ARI for each pair of configurations (tMM case, $g = 3$, balanced clusters)

Another line of research may refer to the clustering methods considered in the comparison. In the present paper, the most common model-based and fuzzy clustering methods have been applied. A new analysis including other more refined methods could however be useful. For instance, starting from the FkM algorithm, several fuzzy algorithms have been proposed in literature to solve specific problems. These can be the entropic k-Means algorithm [29] and many other, which try to mix the two approaches. These methods integrate parametric models into fuzzy clustering algorithms according to what we may refer to as *fuzzy model-based approach*. Examples can be found in [11, 40].

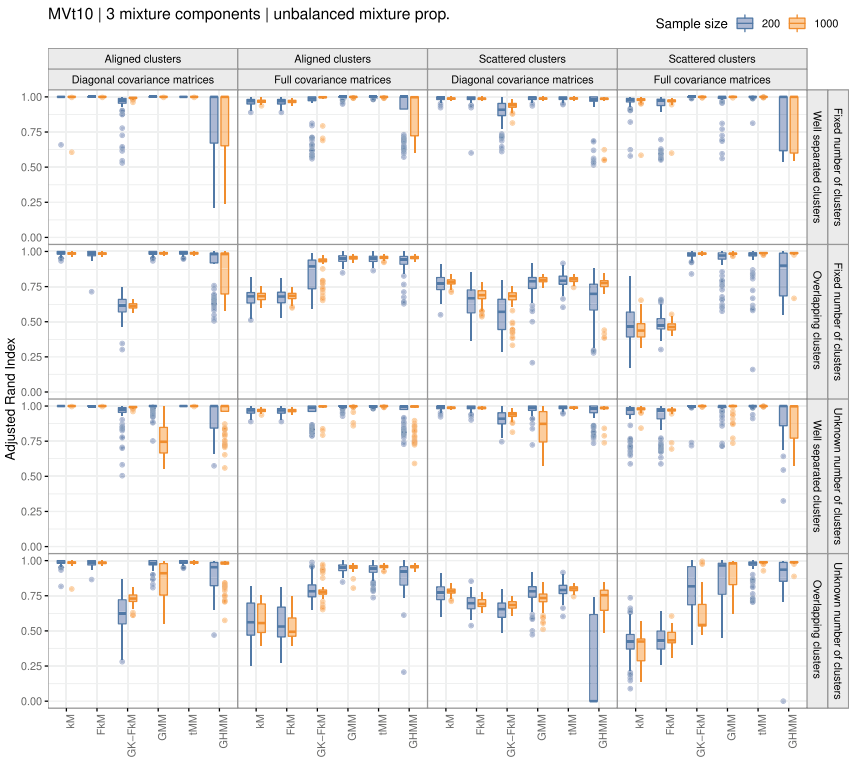


Fig. 8 Boxplots of ARI for each pair of configurations (tMM case, $g = 3$, unbalanced clusters)

Finally, since the presence of noise and outliers tend to degrade the estimation for both model-based and fuzzy clustering methods, the robustness of the discussed methods should also be investigated. For instance, this can be solved by adding an improper uniform component [12], in the model-based framework, or a noise cluster [13] in the fuzzy framework, or by trimming [21]. In the fuzzy model-based approach we mention [22]. This further comparison could represent a useful extension to the present work in order to provide a more in-depth view of the similarities and differences among the different approaches and methods.

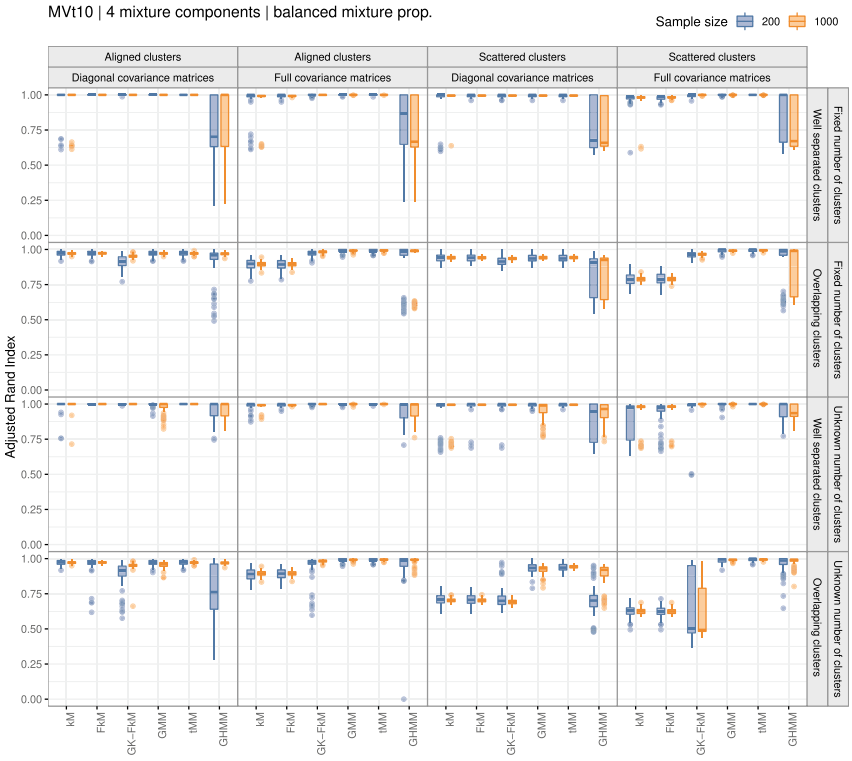


Fig. 9 Boxplots of ARI for each pair of configurations (tMM case, $g = 4$, balanced clusters)

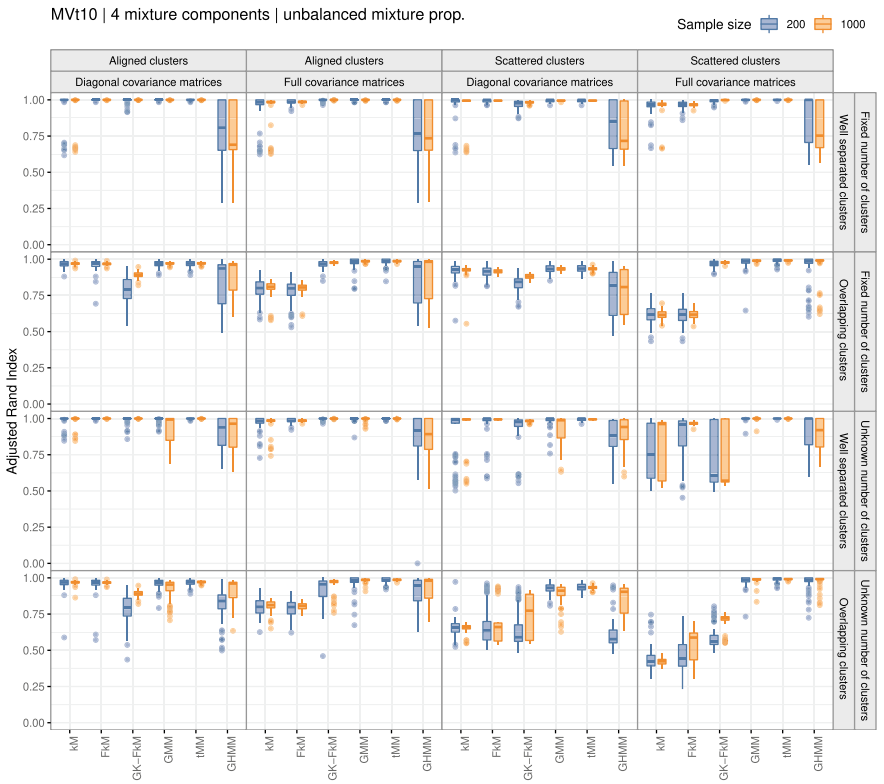


Fig. 10 Boxplots of ARI for each pair of configurations (tMM case, $g = 4$, unbalanced clusters)

References

1. Andrews, J.L., McNicholas, P.D.: Model-based clustering, classification, and discriminant analysis via mixtures of multivariate t-distributions. Stat. Comp. **22**, 1021–1029 (2012)
2. Andrews, J.L., McNicholas, P.D., Subedi, S.: Model-based classification via mixtures of multivariate t-distributions. Comput. Stat. Data Anal. **55**, 520–529 (2011)
3. Andrews, J.L., Wickins, J.R., Boers, N.M., McNicholas, P.D.: Teigen: an R package for model-based clustering and classification via the multivariate t distribution. J. Stat. Soft. **83**, 1–32 (2018)
4. Babuska, R., Van der Veen, P.J., Kaymak, U.: Improved covariance estimation for Gustafson-Kessel clustering. In: Proceedings of the 2002 IEEE International Conference on Fuzzy Systems, vol. 2, pp. 1081–1085 (2002)
5. Banfield, J., Raftery, A.E.: Model-based gaussian and non-gaussian clustering. Biometrics **49**, 803–821 (1993)
6. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Kluwer Academic Publishers, New York (1981)
7. Bezdek, J.C., Hathaway, R.J., Huggins, V.J.: Parametric estimation for normal mixtures. Pattern Recognit. Lett. **3**, 79–84 (1985)
8. Browne, R.P., McNicholas, P.D.: A mixture of generalized hyperbolic distributions. Can. J. Stat. **43**, 176–198 (2015)

9. Campello, R.J.G.B., Hruschka, E.R.: A fuzzy extension of the silhouette width criterion for cluster analysis. Fuzzy Sets Syst. **157**, 2858–2875 (2006)
10. Celeux, G., Govaert, G.: Gaussian parsimonious clustering models. Pattern Recognit. **28**, 781–793 (1995)
11. Choy, S.K., Lam, S.Y., Yu, K.W., Lee, W.Y., Leung, K.T.: Fuzzy model-based clustering and its application in image segmentation. Pattern Recognit. **68**, 141–157 (2017)
12. Coretto, P., Hennig, C.: Maximum likelihood estimation of heterogeneous mixtures of gaussian and uniform distributions. J. Stat. Plan. Infer. **141**, 462–473 (2011)
13. Davé, R.N.: Characterization and detection of noise in clustering. Pattern Recognit. Lett. **12**, 657–664 (1991)
14. Davenport, J.W., Bezdek, J.C., Hathaway, R.J.: Parameter estimation for finite mixture distributions. Comput. Math. Appl. **15**, 819–828 (1988)
15. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm (with discussion). J. R. Stat. Soc. B **39**, 1–38 (1977)
16. Ferraro, M.B., Giordani, P.: A toolbox for fuzzy clustering using the R programming language. Fuzzy Sets Syst. **279**, 1–16 (2015)
17. Ferraro, M.B., Giordani, P., Serafini, A.: fclust: an R package for fuzzy clustering. R J. **11**(1), 198–210 (2019)
18. Flynt, A., Dean, N., Nugent, R.: sARI: a soft agreement measure for class partitions incorporating assignment probabilities. Adv. Data Anal. Class. **13**, 303–323 (2019)
19. Fraley, C., Raftery, A.E.: How many clusters? which clustering method? answers via model-based cluster analysis. Comp. J. **41**, 578–588 (1998)
20. Fraley, C., Raftery, A.E.: Model-based clustering, discriminant analysis, and density estimation. J. Am. Stat. Assoc. **97**, 611–631 (2002)
21. García-Escudero, L.A., Gordaliza, A., Matrán, C., Mayo-Iscar, A.: A general trimming approach to robust cluster analysis. Ann. Stat. **36**, 1324–1345 (2008)
22. García-Escudero, L.A., Greselin, F., Mayo-Iscar, A.: Robust, fuzzy, and parsimonious clustering, based on mixtures of factor analyzers. Int. J. Approx. Reason. **94**, 60–75 (2018)
23. Greselin, F., Ingrassia, S.: Constrained monotone EM algorithms for mixtures of multivariate t distributions. Stat. Comput. **20**, 9–22 (2010)
24. Gustafson, D.E., Kessel, W.C.: Fuzzy clustering with a fuzzy covariance matrix. In: Proceedings of the IEEE Conference on Decision and Control, pp. 761–766 (1979)
25. Hathaway, R.J.: Another interpretation of the EM algorithm for mixture distributions. Stat. Prob. Lett. **4**, 53–56 (1986)
26. Hubert, L., Arabie, P.: Comparing partitions. J. Class. **2**, 193–218 (1985)
27. Ingrassia, S., Rocci, R.: Constrained monotone EM algorithms for finite mixture of multivariate Gaussians. Comput. Stat. Data Anal. **51**, 5339–5351 (2007)
28. Kaufman, L., Rousseeuw, P.: Finding Groups in Data: An Introduction to Cluster Analysis. Wiley, Hoboken (1990)
29. Li, R.P., Mukaidono, M.: Gaussian clustering method based on maximum-fuzzy-entropy interpretation. Fuzzy Sets Syst. **102**, 370–379 (1999)
30. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 281–297 (1967)
31. McLachlan, G.J., Peel, D.: Finite Mixture Models. Wiley, New York (2000)
32. Meila, M.: Criteria for Comparing Clusterings. In: Hennig, C., Meila, M., Murtagh, F., Rocci, R. (eds.) Handbook of Cluster Analysis, pp. 619–635. Chapman and Hall/CRC, New York (2016)
33. Meng, X.-L., Rubin, D.B.: Maximum likelihood estimation via the ECM algorithm: a general framework. Biometrika **80**, 267–278 (1993)
34. Milligan, G.W., Cooper, M.C.: An examination of procedures for determining the number of clusters in a data set. Psychometrika **50**, 159–179 (1985)
35. Pal, N.R., Bezdek, J.C.: Cluster validity for the fuzzy c-means model. IEEE T. Fuzzy Syst. **3**, 370–379 (1995)

36. Peel, D., McLachlan, G.J.: Robust mixture modelling using the t distribution. Stat. Comput. **10**, 339–348 (2000)
37. Schwartz, G.: Estimating the dimension of a model. Ann. Stat. **6**, 31–38 (1978)
38. Scrucca, L., Fop, M., Murphy, T.B., Raftery, A.E.: mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. R J. **8**(1), 205–233 (2016)
39. Tortora, C., ElSherbiny, A., Browne, R.P., Franczak, B.C., McNicholas, P.D., Donald D.A.: MixGHD: model based clustering, classification and discriminant analysis using the mixture of generalized hyperbolic distributions. R package version 2.3.3 (2019). https://CRAN.R-project.org/package=MixGHD
40. Yang, M.-S., Su, C.-F.: On parameter estimation for normal mixtures based on fuzzy clustering algorithms. Fuzzy Sets Syst. **68**, 13–28 (1994)
41. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)

An Analysis of Misclassification Rates in Rater Agreement Studies

Amalia Vanacore and Maria Sole Pellegrino

Abstract This study aims at investigating, via a Monte Carlo simulation, the performance of two non-parametric benchmarking procedures for characterizing the extent of rater agreement in non asymptotic conditions. The performance of each procedure has been evaluated by computing an overall weighted misclassification rate; moreover, in order to investigate whether the procedures overestimate or underestimate the level of agreement, misclassification frequencies have been computed for each agreement category.

Keywords κ-type agreement coefficients · Non-parametric benchmarking procedures · Monte Carlo simulation · Misclassification rate

1 Introduction

In several fields, ranging from manufacturing to health-care and risk management, agreement coefficients are widely adopted for assessing the precision of subjective evaluations provided by human raters to support strategic and operational decisions. κ-type agreement coefficients are commonly adopted for measuring the degree of agreement between the evaluations provided on the same sample of items by two or more raters (i.e. inter-rater agreement) or by the same rater in two or more occasions (i.e. intra-rater agreement).

In order to qualify the extent of rater agreement as good or poor the computed coefficient is compared against a benchmark scale. However, as the agreement coefficient is computed on a sample of items, its magnitude may strongly depend on some experimental factors such as rating scale dimension, number of rated items, trait prevalence and marginal probabilities [12, 18]. It is evident that the interpretation

A. Vanacore (✉) · M. S. Pellegrino
Department of Industrial Engineering, University of Naples "Federico II"
p.le Tecchio 80, 80125, Naples, Italy
e-mail: amalia.vanacore@unina.it

M. S. Pellegrino
e-mail: mariasole.pellegrino@unina.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8_20

based on the straightforward benchmarking should be treated with caution especially for comparison across studies when the experimental conditions are not the same.

A proper characterization of the extent of rater agreement should rely upon a benchmarking procedure that allows to identify a suitable neighborhood of the true value of rater agreement by taking into account sampling uncertainty. The most simple and intuitive way to accomplish this task is by comparing the lower confidence bound of the agreement coefficient against a benchmark scale. The widespread confidence interval is the parametric one, whose accuracy depends on the asymptotic normality of the coefficients and on the asymptotic solution for their variance [12]; both assumptions are questionable for small sample sizes, for which confidence intervals based on bootstrap resampling are the methods of choice [16, 17, 19].

Two non-parametric inferential benchmarking procedures based respectively on percentile and Bias-Corrected and Accelerated bootstrap confidence intervals will be in the following discussed. Their performance will be evaluated and compared via a Monte Carlo simulation study firstly in terms of weighted misclassification rate, that is a synthetic measure of the procedure performance, and then looking into each specific agreement category for investigating whether the procedures overestimate or underestimate the level of agreement. Several scenarios have been investigated, differing for sample size and rating scale dimension.

The remainder of the paper is organized as follows: in Sect. 2 two well-known paradox-resistant κ-type agreement coefficients are discussed; the commonly adopted benchmark scale and the investigated characterization procedures based on non-parametric confidence intervals are presented and discussed in Sect. 3; in Sect. 4 the simulation design is described and the main results are fully discussed; finally, conclusions are summarized in Sect. 5.

2 Definition of κ-type agreement coefficients

The κ-type agreement coefficients are relative measures of agreement obtained by rescaling the proportion of observed agreement with the proportion of agreement expected by chance alone. Several κ-type coefficients have been proposed in the literature differing from each other in the definition, and thus formulation, of agreement expected by chance alone.

Among them, the most common is Cohen kappa [6]. Despite its popularity, Cohen kappa is criticized because it is affected by two paradoxes [5, 9]: the degree to which raters disagree (bias problem) and the marginal distribution of the evaluations independently provided by each rater (prevalence problem). Specifically, "unbalanced marginal totals produce higher values of Cohen kappa than more balanced marginal totals"; whereas the chance agreement value produced by symmetric marginal distribution is higher than that obtained with asymmetric distribution, making the κ value considerably decrease.

A solution to face the above paradoxes firstly suggested by Bennett et al. [1] for the simplest case of 2 rating categories is to adopt the uniform distribution for chance

Table 1 $k \times k$ contingency table

		2^{nd} rater					
	Category	1	...	j	...	k	**Total**
1^{st} rater	1	n_{11}	...	n_{1j}	...	n_{1k}	$n_{1\cdot}$
	⋮	⋮	...	⋮	...	⋮	⋮
	i	n_{i1}	...	n_{ij}	...	n_{ik}	$n_{i\cdot}$
	⋮	⋮	...	⋮	...	⋮	⋮
	k	n_{k1}	...	n_{kj}	...	n_{kk}	$n_{k\cdot}$
	Total	$n_{\cdot 1}$	...	$n_{\cdot j}$	...	$n_{\cdot k}$	n

measurements, which – given a certain rating scale – can be defended as representing the maximally non-informative measurement system [7].

Specifically, let n be the number of items rated by two raters on an ordinal k-point rating scale, the data can be arranged in a $k \times k$ contingency table $(n_{ij})_{k\times k}$ (Table 1) where the generic (i, j) cell contains the joint frequency n_{ij} that counts the number of items classified into i^{th} category by the first rater and into j^{th} category by the second rater.

The weighted κ-type coefficients allow to consider that, on ordinal rating scales, disagreement on two distant categories is more serious than disagreement on neighboring categories.

The weighted version of the uniform kappa [2, 10, 13, 14], often referred to as Brennan-Prediger coefficient (hereafter $\widehat{BP}_w$), is formulated as:

$$\widehat{BP}_w = \frac{p_{a_w} - p_{a|c}^{BP_w}}{1 - p_{a|c}^{BP_w}}. \tag{1}$$

Being w_{ij} the symmetrical weight corresponding to (i, j) cell, the proportion of weighted observed agreement p_{a_w} is given by:

$$p_{a_w} = \sum_{i=1}^{k}\sum_{j=1}^{k} w_{ij}\frac{n_{ij}}{n}; \tag{2}$$

whereas the proportion of weighted agreement expected under the assumption of uniform chance measurements $p_{a|c}^{BP_w}$ is formulated as follows:

$$p_{a|c}^{BP_w} = \frac{T_w}{k^2} \tag{3}$$

where T_w is the sum over all weight values w_{ij}.

Another well-known paradox-resistant κ-type agreement coefficient is the AC_1 proposed by Gwet [11], whose weighted version AC_2 [12] is formulated as:

$$\widehat{AC}_2 = \frac{p_{a_w} - p_{a|c}^{AC_2}}{1 - p_{a|c}^{AC_2}} \tag{4}$$

The proportion of weighted observed agreement p_{a_w} is still formulated as in Eq. 2, whereas the probability of chance agreement $p_{a|c}^{AC_2}$ is given by:

$$p_{a|c}^{AC_2} = \frac{T_w}{k(k-1)} \cdot \sum_{i=1}^{k} p_i(1 - p_i) \tag{5}$$

Specifically, $p_{a|c}^{AC_2}$ is defined as the probability of the simultaneous occurrence of two events, one rater provides random rating (R) and the two raters agree (G):

$$p_{a|c}^{AC_2} = P(G \cap R) = P(G|R) \cdot P(R) \tag{6}$$

where $P(G|R) = T_w/k^2$ and $P(R)$ is approximated with a normalized measure of randomness defined by the ratio of the observed variance to the variance expected under the assumption of totally random ratings:

$$P(R) = \frac{\sum_{i=1}^{k} p_i(1 - p_i)}{(k-1)/k} \tag{7}$$

with p_i denoting the propensity that a rater assigns score i to an item which is estimated by $p_i = (n_{i\cdot} + n_{\cdot i})/2n$ being $n_{i\cdot}$ (resp. $n_{\cdot i}$) the total number of items classified into i^{th} category by the first (resp. second) rater.

3 Non-parametric inferential benchmarking procedures

After computing an agreement coefficient, a common issue is how to characterize the extent of agreement. For this purpose, several benchmark scales have been proposed mainly in social and medical sciences over the years aiming at providing an aid to qualify the magnitude of κ-type coefficients. The most widely adopted benchmark scale is by far the scale proposed by Landis and Koch [15] and reported in Table 2, consisting of six ranges of coefficient values corresponding to as many categories of agreement.

Despite its popularity, the straightforward benchmarking can be misleading because it does not associate the interpretation of the extent of agreement with a degree of uncertainty and it does not allow to compare the extent of agreement across

Table 2 Landis and Koch benchmark scale for κ-type coefficients

κ values	Agreement category
≤ 0.00	Poor
0.01–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

different studies, unless they are carried out under the same experimental conditions. In order to have a fair characterization of the extent of rater agreement, it is necessary to associate a degree of uncertainty to the interpretation of the coefficient.

Under non-asymptotic conditions, the magnitude of the κ-type coefficient could be related to the notion of extent of agreement by benchmarking the lower bound of $(1 - 2\alpha)\%$ non-parametric confidence interval based on bootstrap resampling that, being free from distributional assumptions, fits also the cases of small sample sizes. Among the available methods to build bootstrap confidence intervals, the simplest is the percentile bootstrap which takes the α and $1 - \alpha$ percentiles of the cumulative distribution function of the bootstrap replications of the κ coefficient as lower and upper confidence bounds. For severely skewed distribution, instead, the Bias-Corrected and Accelerated bootstrap confidence interval is recommended [3, 8], since it adjusts for any bias and lack of symmetry of the bootstrap distribution through the acceleration parameter and the bias correction parameter.

4 Monte Carlo Simulation

The benchmarking procedures based on percentile and Bias-Corrected and Accelerated bootstrap confidence intervals have been applied to characterize the extent of both BP_w and AC_2 under several conditions. Their statistical properties have been investigated via a Monte Carlo simulation study developed considering the classifications provided by the same rater in two occasions (or simultaneously by two raters) about the same set of n items into one of the k ordinal rating categories.

The simulation study has been designed as multi-factor experimental design with three multi-level factors: number of items n, rating scale dimension k and level of agreement. The factor n has 4 levels: $n = 5, 10, 15, 30$ items; the factor k has 4 levels: $k = 2, 3, 5, 7$ ordinal classification categories; whereas the factor level of agreement has 7 levels: 0, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9.

For each combination of n, k and level of agreement, $r = 2000$ Monte Carlo data sets have been generated; for each Monte Carlo data set, the two κ-type coefficients under study (BP_w and AC_2) have been assessed and for both of them percentile and

Bias-Corrected and Accelerated bootstrap confidence intervals have been built, for a total of 224 investigated scenarios.

The data have been simulated by sampling each Monte Carlo data set from a multinomial distribution with parameters n and $\mathbf{p} = (\pi_{11}, \ldots, \pi_{ij}, \ldots, \pi_{ik})$; the π_{ij} values have been set so as to obtain the seven true population values of agreement level (viz. 0, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9), assuming a linear weighting scheme [4].

The performance of the benchmarking procedures based on lower bound of percentile and Bias-Corrected and Accelerated bootstrap confidence interval (i.e. LB_p and LB_{BCa}, respectively) has been evaluated in terms of weighted misclassification rate (hereafter, $\mathbf{M}_w$). Specifically, let $\{X_h; r\}_\omega$ be a Monte Carlo data set containing r benchmarks X_h obtained for a population value taken as reference for a specific agreement category ω. $\mathbf{M}_w$ is evaluated as the weighted proportion of misclassified X_h:

$$\mathbf{M}_w = \frac{1}{r} \sum_{\omega=1,\Omega} w_{\omega\omega'} \cdot I\left[X_{h|\omega} \in \omega'\right]; \quad \omega' \neq \omega \tag{8}$$

where $I[\cdot]$ is an indicator function taking value 1 if the argument is true and 0 otherwise and $w_{\omega\omega'}$ is a linear misclassification weight adopted to account that on an ordinal benchmarking scale some misclassifications are more serious than others. The best and worst values of $\mathbf{M}_w$ obtained across the analysed benchmarking procedures for BP_w and AC_2 are reported in Table 3 for each combination of n and k values.

Simulation results reveal that $\mathbf{M}_w$ slightly differs across benchmarking procedures and agreement coefficients and it is affected by sample size and rating scale

Table 3 Best and worst $\mathbf{M}_w$ between benchmarking procedures (yellow: LB_p; orange: LB_{BCa}) for BP_w and AC_2

(a) Best $\mathbf{M}_w$ for BP_w

	$n=5$	$n=10$	$n=15$	$n=30$
$k=2$	0.372	0.238	0.168	0.117
$k=3$	0.230	0.175	0.140	0.090
$k=5$	0.188	0.158	0.132	0.086
$k=7$	0.182	0.089	0.087	0.072

(b) Worst $\mathbf{M}_w$ for BP_w

	$n=5$	$n=10$	$n=15$	$n=30$
$k=2$	0.372	0.277	0.220	0.137
$k=3$	0.281	0.177	0.156	0.099
$k=5$	0.214	0.177	0.137	0.094
$k=7$	0.192	0.097	0.088	0.080

(c) Best $\mathbf{M}_w$ for AC_2

	$n=5$	$n=10$	$n=15$	$n=30$
$k=2$	0.344	0.236	0.167	0.107
$k=3$	0.262	0.161	0.124	0.073
$k=5$	0.184	0.139	0.113	0.075
$k=7$	0.153	0.105	0.099	0.061

(d) Worst $\mathbf{M}_w$ for AC_2

	$n=5$	$n=10$	$n=15$	$n=30$
$k=2$	0.369	0.279	0.215	0.131
$k=3$	0.303	0.245	0.175	0.101
$k=5$	0.198	0.202	0.161	0.098
$k=7$	0.158	0.138	0.135	0.081

dimension; its value decreases as the number of both rated items and rating categories increases.

Specifically, the worst performance for both coefficients has been obtained for $n = 5$ where $\mathbf{M}_{\mathrm{w}}$ always exceeds 10%, ranging from 15% to 37%. The performance of non-parametric benchmarking procedures becomes satisfactory for samples of $n = 30$ items since $\mathbf{M}_{\mathrm{w}} \approx 10\%$, being quite indistinguishable across procedures and coefficients; the differences in $\mathbf{M}_{\mathrm{w}}$ across benchmarking procedures and agreement coefficients get smaller as n increases because of the decreasing skewness in the distributions of the coefficients: if the distribution is symmetric, $\mathrm{LB_p}$ and $\mathrm{LB_{BCa}}$ agree. In all other scenarios, the procedures have good performance only when $k = 7$.

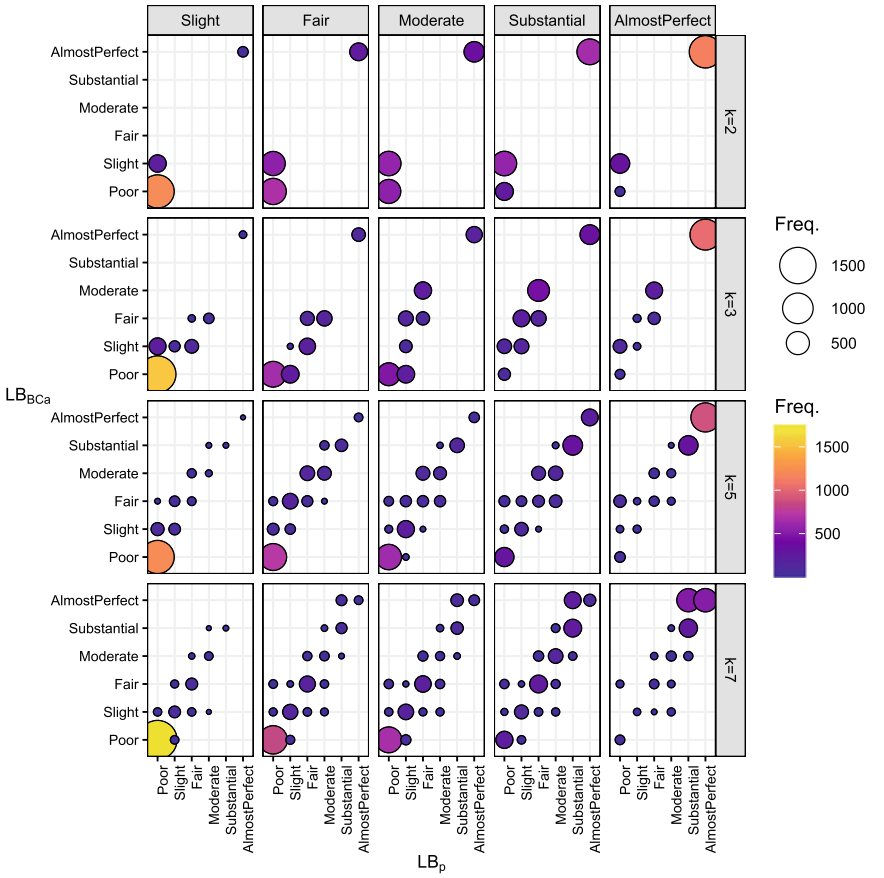


Fig. 1 Bubble charts of agreement characterizations provided by $\mathrm{LB_{BCa}}$ (left) and $\mathrm{LB_p}$ (bottom) for BP_w, stratified by true agreement levels (top) and number of rating categories k (right) with $n = 5$

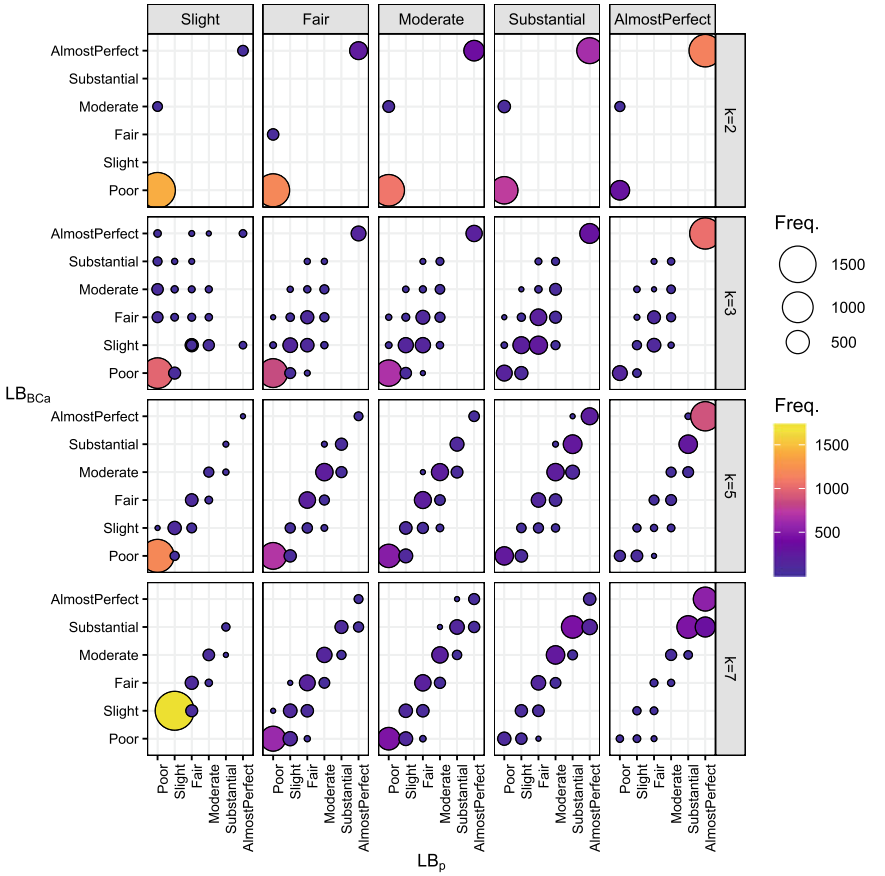


Fig. 2 Bubble charts of agreement characterizations provided by LB_{BCa} (left) and LB_p (bottom) for AC_2, stratified by true agreement levels (top) and number of rating categories k (right) with $n = 5$

The benchmarking procedure with a generally better performance in terms of weighted misclassification rate is LB_{BCa} for BP_w; whereas LB_p is the best benchmarking procedure for AC_2.

For scenarios of $n = 5$ and $n = 30$ items, it is further investigated whether benchmarking procedures underestimate or overestimate the level of agreement. In this respect, the simulated Monte Carlo data sets have been stratified according to their true level of agreement into five strata corresponding to the Landis and Koch agreement categories (see Table 2). For each stratum the number of data sets assigned by LB_{BCa} to agreement category l and by LB_p to category l' (with l, l' ranging from Poor to Almost Perfect) has been calculated for both BP_w (see Figs. 1 and 3) and AC_2 (see Figs. 2 and 4).

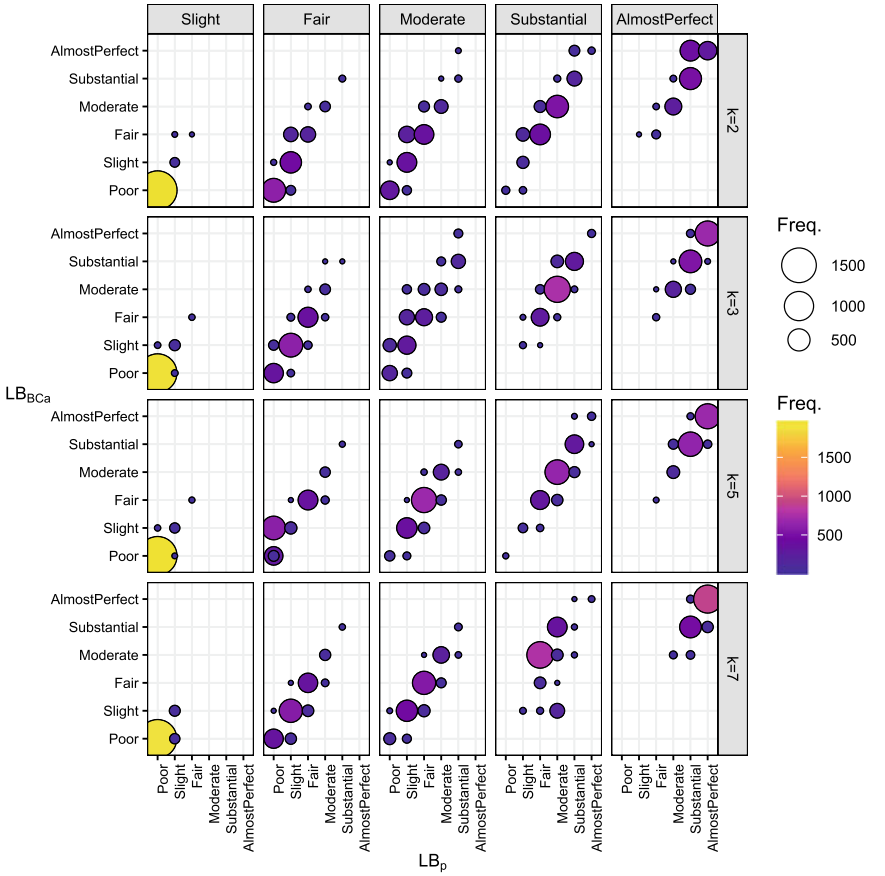


Fig. 3 Bubble charts of agreement characterizations provided by $\mathrm{LB_{BCa}}$ (left) and $\mathrm{LB_p}$ (bottom) for BP_w, stratified by true agreement levels (top) and number of rating categories k (right) with $n = 30$

The results reveal that the two benchmarking procedures generally assign the simulated data sets into the true agreement category or into lower ones and rarely the agreement level is overestimated. The agreement characterization is mainly the same between the two benchmarking procedures and very rarely, especially with $n = 5$ and $k \leq 3$, the two procedures characterize the agreement as belonging to more than 1-step apart categories. Moreover, $\mathrm{LB_{BCa}}$ generally characterizes the agreement into a higher category than $\mathrm{LB_p}$ with BP_w, vice-versa for AC_2.

With samples of $n = 30$ items, the performance improves for both benchmarking procedures and the frequency of discordant classifications between benchmarking procedures decreases.

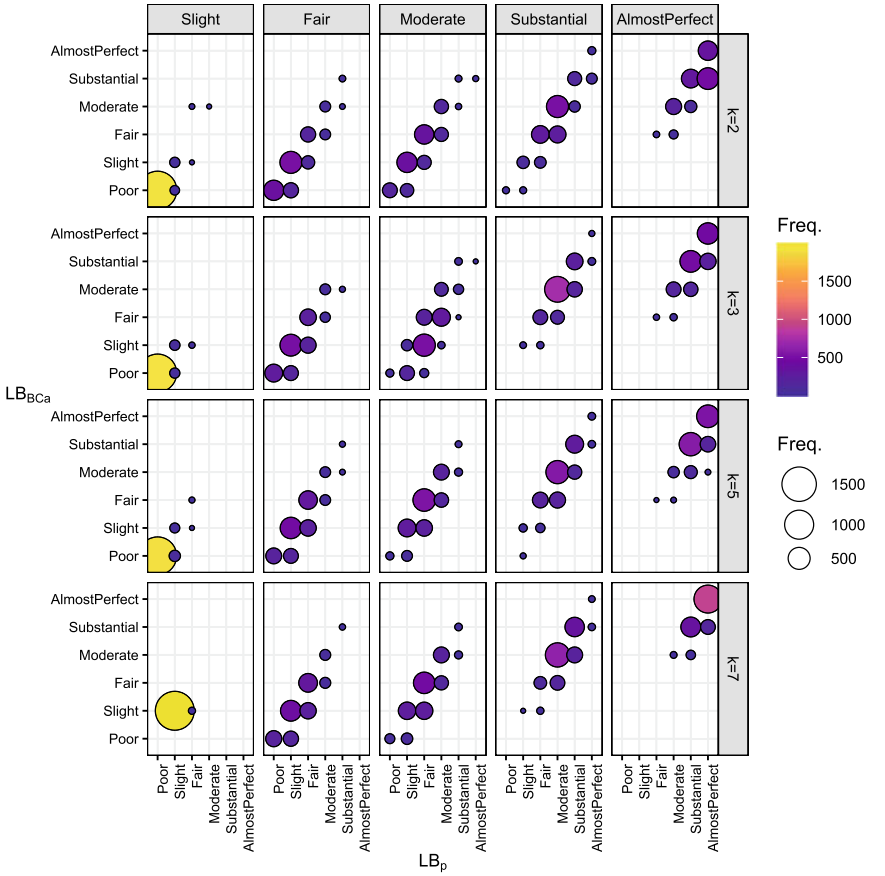


Fig. 4 Bubble charts of agreement characterizations provided by LB_{BCa} (left) and LB_p (bottom) for AC_2, stratified by true agreement levels (top) and number of rating categories k (right) with $n = 30$

The performance of both procedures is slightly worse when the true agreement category is Moderate; this result could be explained by the dimension of Landis and Koch benchmark scale [15] consisting of 6 categories of agreement (Table 2), maybe too many for a correct agreement characterization.

5 Conclusions

In this paper the agreement characterization performance of two non-parametric inferential procedures based on bootstrap resampling has been evaluated via Monte Carlo simulation.

The simulation results suggest that the two non-parametric procedures under study provide a correct agreement characterization even for very small samples of $n = 10$ items when a $k = 7$-point rating scale is adopted; vice-versa, for smaller rating scale dimensions, at least $n = 30$ items are needed. The misclassification is more frequently due to an underestimation of the extent of agreement. The characterizations perfectly agree in more than 50% of cases and for less than 5% of cases it is in categories of two or more steps apart.

Because of the less computational burden, benchmarking the lower bound of the percentile bootstrap confidence interval could be suggested for characterizing the extent of rater agreement for both BP_w and AC_2, as the performance of the benchmarking procedures generally differs from each other for no more than 5%.

Study results also reveal a small difference between coefficients due to the slightly lower frequency of misclassifications when adopting AC_2 suggesting that the choice of the agreement coefficient seems to be a more influencing factor than that of the non-parametric benchmarking procedure.

References

1. Bennett, E.M., Alpert, R., Goldstein, A.: Communications through limited-response questioning. Public Opin. Q. **18**(3), 303–308 (1954)
2. Brennan, R.L., Prediger, D.J.: Coefficient kappa: some uses, misuses, and alternatives. Educ. Psychol. Measur. **41**(3), 687–699 (1981)
3. Carpenter, J., Bithell, J.: Bootstrap confidence intervals: when, which, what? a practical guide for medical statisticians. Stat. Med. **19**(9), 1141–1164 (2000)
4. Cicchetti, D.V., Allison, T.: A new procedure for assessing reliability of scoring EEG sleep recordings. Am. J. EEG Technol. **11**(3), 101–110 (1971)
5. Cicchetti, D.V., Feinstein, A.R.: High agreement but low kappa: II. resolving the paradoxes. J. Clin. Epidemiol. **43**(6), 551–558 (1990)
6. Cohen, J.: A coefficient of agreement for nominal scales. Educ. Psychol. Meas. **20**(1), 37–46 (1960)
7. De Mast, J., Van Wieringen, W.N.: Measurement system analysis for categorical measurements: agreement and kappa-type indices. J. Qual. Technol. **39**(3), 191–202 (2007)
8. Efron, B., Tibshirani, R.J.: An introduction to the bootstrap. CRC Press (1994)
9. Feinstein, A.R., Cicchetti, D.V.: High agreement but low kappa: I. the problems of two paradoxes. J. Clin. Epidemiol. **43**(6), 543–549 (1990)
10. Guttman, L.: The test-retest reliability of qualitative data. Psychometrika **11**(2), 81–95 (1946)
11. Gwet, K.L.: Computing inter-rater reliability and its variance in the presence of high agreement. Br. J. Math. Stat. Psychol. **61**(1), 29–48 (2008)
12. Gwet, K.L.: Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Raters. Advanced Analytics, LLC (2014)
13. Holley, J.W., Guilford, J.P.: A note on the g index of agreement. Educ. Psychol. Measur. **24**(4), 749–753 (1964)
14. Janson, S., Vegelius, J.: On generalizations of the g index and the phi coefficient to nominal scales. Multivar. Behav. Res. **14**(2), 255–269 (1979)
15. Landis, J.R., Koch, G.G.: The measurement of observer agreement for categorical data. Biometrics 159–174 (1977)
16. Lee, J., Fung, K.: Confidence interval of the kappa coefficient by bootstrap resampling. Psychiatry Res. (1993)

17. Reichenheim, M.E.: Confidence intervals for the kappa statistic. Stand. Genomic Sci. **4**(4), 421–428 (2004)
18. Thompson, W.D., Walter, S.D.: A reappraisal of the kappa coefficient. J. Clin. Epidemiol. **41**(10), 949–958 (1988)
19. Zapf, A., Castell, S., Morawietz, L., Karch, A.: Measuring inter-rater reliability for nominal data-which coefficients and confidence intervals are appropriate? BMC Med. Res. Methodol. **16**(1), 93 (2016)

Author Index

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

E. Brentari et al. (eds.), *Models for Data Analysis*, Springer Proceedings in Mathematics & Statistics 402, https://doi.org/10.1007/978-3-031-15885-8