

University of Brescia

Department of Economics and Management

























Statistics for Health and Well-being



University of Brescia Department of Economics and Management 25 – 27 September 2019

ASA CONFERENCE 2019 Statistics for Health and Well-being

BOOK OF SHORT PAPERS

Maurizio Carpita and Luigi Fabbris *Editors*



ASA Conference 2019 - Book od Short Papers Statistics for Health and Well-being

University of Brescia, September 25-27, 2019 Maurizio Carpita and Luigi Fabbris (Editors)

ISBN: 978-88-5495-135-8

This Book is published only in pdf format.

Copyright © 2019 CLEUP sc Cooperativa Libraria Editrice University of Padova via G. Belzoni 118/3 35121 Padova info@cleup.it

INTRODUCTION

This Book includes a selection of 53 peer-reviewed short papers submitted to the Scientific Conference "Statistics for Health and Well-Being", held at the University of Brescia from 25 to 27 September, 2019.

The Conference, aimed at promoting applications that use statistical techniques and models suitable for health and well-being analyses, was organized by the ASA (Association for Applied Statistics) and the DMS StatLab (Data Methods and Systems Statistical Laboratory) of the Department of Economics and Management, University of Brescia.

The programme of the Conference included 25 parallel sessions with a total of 82 contributions with about 100 attendants, 4 plenary sessions (organised by ISTAT, the Italian National Statistical Institute, and USCI, the Statistical Union Italian Municipalities; SIS, the Italian Statistical Society, and ASA; AICQ-CN, the Italian Association for Quality Culture-North and Centre of Italy, and AISS, the Italian Academy for Six Sigma; and DBSPORTS, Big Data Analytics in Sports Project, respectively) and 4 special events (ISTAT and ASA Open Conference with the President of ISTAT, IASA Sensory Experiment, Visit to Capitolium, and Kick-off meeting ISI-SPG in Sports Statistics). Thank you very much to Eugenio Brentari, Chair of the Local Program Committee. For more information about the programme and other material visit the website www.sa-ijas.org/statistics-for-health-and-well-being/.

As co-chairs of the ASA Conference 2019, we are very grateful to the authors for submitting their interesting research with various real application of statistics in so many contexts of health and well-being, and to the members of the Scientific Committee for collaborating to the peer-reviewing process.

October, 2019

Co-chair Scientific Program Committee

Maurizio Carpita Luigi Fabbris

University of Brescia University of Padova

Conference session topics include, but are not limited to, the following areas of special interest:

Health and healthcare
Education and health
Health Psychology
Work and life balance
Economic well-being

Social relationships and social health

Welfare and well-being Safety and security Subjective well-being Environment and pollution

Innovation, research and creativity

Quality of health services

Equitable and sustainable well-being

Resilience and vulnerability Sport, Health and wellbeing

Sport analytics
Health and fitness
Sport psychology
Statistics and tourism

Food and beverage, health, well-being and life quality Qualitative and quantitative methods for sensory analysis

Psychology and food

Food and beverage industries and markets

Methods and models for health and well-being analysis

Technology for health analysis

Scientific Program Committee:

Luigi Fabbris (University of Padua, co-chair) Maurizio Carpita (University of Brescia, co-chair)

Giuseppe Arbia (SIS - Università Cattolica di Milano)

Rossella Berni (University of Florence)

Matilde Bini (SIS - European University of Rome) Giovanna Boccuzzo (University of Padova)

Eugenio Brentari (University of Brescia)

Vittoria Buratta (ISTAT)

Giulia Cavrini (University of Bolzano-Bozen)

Alessandro Celegato (AICQ-AISS, PSV Project Service and Value)

Giuliana Coccia (ISTAT)

Adriano Decarli (University of Milan)

Tonio Di Battista ('G. D'Annunzio' University of Chieti and Pescara) Simone Di Zio ('G. D'Annunzio' University of Chieti and Pescara)

Benito Vittorio Frosini (Sacred Heart Catholic University of Milan)

Antonio Giusti (University of Florence)

Silvia Golia (University of Brescia)

Maria Gabriella Grassia (Federico II University of Naples)

Maria Iannario (Federico II University of Naples)

Domenica Fioredistella lezzi (Tor Vergata University of Rome)

Michele Lalla (University of Modena and Reggio Emilia)

Fabio Lucidi (SIPSA - La Sapienza University of Rome)

Marica Manisera (University of Brescia)

Paolo Mariani (University of Milan-Bicocca)

Francesco Mola (University of Cagliari)

Antonio Mussino (La Sapienza University of Rome)

Luigi Odello (International Academy of Sensory Analysis)

Francesco Palumbo (Federico II University of Naples)

Maurizio Pessato (Assirm)

Alessandra Petrucci (University of Florence)

Alfonso Piscitelli (Federico II University of Naples)

Marco Trentini (Unione Statistica Comuni Italiani)

Fabio Vernau (Federico II University of Naples)

Domenico Vistocco (Federico II University of Naples)

Paola Zuccolotto (University of Brescia)

Local Program Committee:

Eugenio Brentari (University of Brescia, chair)

Maurizio Carpita (University of Brescia) Silvia Golia (University of Brescia) Marica Manisera (University of Brescia) Manlio Migliorati (University of Brescia) Anna Simonetto (University of Brescia)

Marika Vezzoli (University of Brescia)

Mariangela Zenga (University of Milano-Bicocca)

Paola Zola (University of Brescia)
Paola Zuccolotto (University of Brescia)



Visit to the Capitolium. Brescia, 26th September 2019

INDEX OF SHORT PAPERS

Giuseppe Alfonzetti, Laura Rizzi, Luca Grassetti, Michele Gobbato Observed expenditures vs estimated burden of health care: a comparative evaluation based on spatial analysis
Pietro Amenta, Antonio Lucadamo, Gabriella Marcarelli Computing ordinal consistency thresholds for pairwise comparison matrices
Ilaria Lucrezia Amerise, Agostino Tarsitano Household wealth and consumption in Italy: analysis by density-weighted quantile regression pag. 9
Fabrizio Antolinia, Francesco Giovanni Truglia Ecotourism and food geographic areas
Bruno Arpino, Silvia Bacci, Leonardo Grilli, Raffaele Guetto, Carla Rampichini Issues in prior achievement adjustment for value added analysis: an application to Invalsi tests in Italian schools
Silvia Bacci, Bruno Bertaccini, Alessandra Petrucci Museum preferences analysis: an item response model applied to network data
Chiara Bocci, Silvana Salvini Elderly with and without children: do they report different health conditions? pag. 25
Chiara Bocci, Laura Grassini, Emilia Rocco A multi-inflated hurdle regression model for the total number of overnight stays of Italian tourists in the years of the economic recession
Riccardo Borgia, Elena Castellari, Paolo Sckokai Family lifestyle habits: what is passed down from adults to children?
Elena Bortolato, Luigi Fabbris, Marco Vivian Quantity and mood of final open-ended comments on an Erasmus+ VET mobility questionnaire pag. 37
Rafaela Soares Bueno, Luiz Sá Lucas, Ana Carolina Sá Lucas Balancing multi-class imbalanced data into a training dataset using SCUT method pag. 41
Stefania Capecchi, Carmela Cappelli, Maurizio Curtarelli, Francesca Di Iorio Investigating well-being at work via composite indicators
Maurizio Carpita, Enrico Ciavolino, Paola Pasca Exploring the statistical structure of soccer team performance variables using the Principal Covariates Regression
Maurizio Carpita The mobile phone big data tell the story of the impact of Christo's The Floating Piers on the Lake Iseo
Daniela Caso, Maria Iannario, Francesco Palumbo Athletes' mental skills, personality and other drivers to assess the performance in a study on volleyball
Rosanna Cataldo, Maria Gabriella Grassia, Marina Marino Partial Least Squares Path Modelling approach for sustainability using qualitative information pag. 61
Carlo Cavicchia, Pasquale Sarnacchiaro, Maurizio Vichi A composite indicator via hierarchical disjoint factor analysis for measuring the Italian football teams' performances

Alessandro Lubisco, Stefania Mignani, Carlo Trivisano Assessment of game actions performance in water polo: a data analytic approach	oag. 141
Luiz Sá Lucas, Ana Carolina Sá Lucas, Rafaela Bueno Selecting features for Machine Learning in Alzheimer's diagnosticsp	oag. 145
Paolo Mariani, Andrea Marletta, Nicholas Missineo Missing values in social media: an application on Twitter data	oag. 149
Milica Maricic Application of multivariate statistics in sports: exploration of recall and recognition of Champions League sponsors	
Daria Mendola, Paolo Li Donni Short-run and long-run persistence of bad health among elderly	oag. 157
Vittorio Nicolardi, Caterina Marini Harmonised Administrative Databases: a new approach in the era of Big Data	oag. 161
Antonio Notarnicola, Vito Santarcangelo, Nicola Martullib, Francesco Abbondanza The blockchain for the certification of the dairy supply chain, the "Lucanum" basket and the products for well-being	
Omar Paccagnella, Ilaria Zanin Another look at the relationship between perceived well-being and income satisfaction p	oag. 169
Anna Parola, Francesco Palumbo Profile pattern of italians NEET by nonlinear PCAp	oag. 173
Anna Maria Parroco, Vincenzo Giuseppe Genova, Laura Mancuso, Francesca Giannone Assessing mental health therapeutic communities functioningp	oag. 177
Eugenio Pomarici, Alfonso Piscitelli, Luigi Fabbris, Raffaele Sacchi A pre-post sensory experiment on the effect of a seminar on olive oil preferences of consumers	
Luca Romagnoli, Luigi Mastronardi Understanding local administrations policies effects on well-being in Italian inner areas p	oag. 185
Vito Santarcangelo, Emilio Massa, Diego Carmine Sinitò, Giuseppe Scavone Intelligent systems to support patients	oag. 189
Anna Simonetto, Silvia Golia, Buirma Malo, Gianni Gilioli Food quality perception in children: a comparison between Bayesian Network and St Equation Modelling	
Federico M. Stefanini, Yura Loscalzo The studyholism comprehensive model: towards a bayesian reanalysis	oag. 197
Alessio Surian, Andrea Sciandra City Prosperity Index: a comparative analysis of Latin American and Mediterranean cities b well-being and social inclusion features	
Emma Zavarrone, Maria Gabriella Grassia, Rocco Mazza Invariance in the structural topic models	oag. 205
Paola Zola, Costantino Ragno, Paulo Cortez Inferring Twitter users home location based on trend topics	oag. 209

Investigating well-being at work via composite indicators

Stefania Capecchi^a, Carmela Cappelli^a, Maurizio Curtarelli^b, Francesca Di Iorio^a ^a Department of Political Sciences, University of Naples Federico II, Naples, Italy; ^b Prevention and Research Unit, EU-OSHA (European Agency for Safety and Health at Work), Bilbao, Spain

1. Introduction

Psychological or subjective well-being is a multifaceted concept covering several related phenomena, involving emotional responses, feelings and global judgements of satisfaction about life (Howell et al., 2007, among many others), and its multiple domains (e.g. housing, family life, work and so on) (World Health Organization, 2012; OECD, 2013).

Work has long been recognised as having important influences (both positive and negative, indeed) on health and well-being (Litchfield et al., 2016). In modern workplaces - alongside to physical, chemical and biological hazards, depending on the type of industry -, hazards are frequently related more to the way work is organised, to the working environment and the nature of work itself rather than to specific agents, and harm is therefore more psychological than physical (Litchfield et al., 2016). The literature provides a comprehensive account of the topic and of job characteristics (and of their lack) which are considered as psychosocial risk factors for workers (EU-OSHA, 2013, among many others). Nonetheless quantitative evidence about the effect of psychosocial risks on health and well-being is still relatively scarce. Furthermore, analyses of the measure of interactions between physical and psychosocial risk factors seem not frequently reported in the relevant literature.

A common practice in analysing survey data regarding workers' SAH is to consider a few drivers covering a wide range of psychosocial and physical risk factors customarily measured by means of scales administered to respondents: interviewees are usually asked to select a response category out of a list, answering questionnaires often made of a number of question batteries.

Using some of the evidence of the European Working Conditions Survey (EWCS), carried out by the European Foundation for the Improving of Living and Working Condition (Eurofound), we present an empirical analysis where the variable of interest is the self-assessed health (SAH) as a proxy of workers' well-being.

More specifically, this paper focuses on well-being at work in order to build synthetic indicators¹, instead of providing a collection of individual results, aiming at understanding which individual risk factors exert a stronger impact on workers' health at the EU28 level, and whether psychosocial risk factors do affect well-being as much as the physical ones.

After a brief sketch of the data employed and of the implemented procedure, results of the synthetic indicators, as obtained from two subsets of risk factors, are discussed, and few concluding remarks end the paper.

2. Data and methods

Data employed in this exercise come from the Sixth EWCS² which provides a wide-ranging

CLEUP Coop. Libraria Editrice ISBN: 978-88-5495-135-8

¹ Current literature, as summarized in OECD (2008) handbook, emphasizes several steps to achieve an effective and consistent composite indicator.

² Eurofound carried out the 6th wave of the survey in 2015 interviewing 43,850 employees and self-employed workers in 35 European countries: the 28 European Union Member States plus, namely, the candidate countries for EU membership (Albania, F.Y.R.of Macedonia, Montenegro, Serbia and Turkey), and Norway and Switzerland. At country level, the sample size ranges from 1,000 to 3,300 people according to the sample design. Data can be downloaded from http://discover.ukdataservice.ac.uk, while detailed description of survey design and report can be found in Eurofound (2017).

picture of Europe at work across countries, occupations, sectors and age groups.

Response variable of interest stems from question Q75 referring to self-reported health status³, as it is common in literature on the subject: "How is your health in general? Would you say it is: (1 Very good; 2 Good; 3 Fair; 4 Bad; 5 Very bad)".

Common individual characteristics here considered are gender, age and education level. Gender (from question Q2a) is expressed by the usual dummy variable where female = 1; age is expressed in years (Q2b). Education level (from the original Q106) is described by a dummy where holding a university degree =1 (tertiary). Given the high number of missing values with reference to net monthly earnings, to investigate the relationship with individual's economic status, information is derived from the answers to Q100: "Thinking of your household's total monthly income, is your household able to make ends meet" (make-ends-meet) rated on a six point wording scale from "Very easily" (1) to "With great difficulty" (6). With respect to job features, we introduce two dummies to distinguish full-time vs. part-time job (fulltime, where full-time=1) and permanent vs. non-permanent job (permjob, where permanent job=1). A dummy is used to consider the belonging of respondent to a country of EU12. Moreover, we reckon the number of working days per week (Q26, d4w) and the hours weekly (Q24) spent at work (whours).

With respect to risk factors at work as surveyed by the EWCS, we got two different sets⁴: the first one gathers all the physical risk factors (15 variables); the second one refers to the psychosocial risk factors, which include a list of variables related to the ways work is organised and managed as long as the social environment of workers (27 variables).

Missing values and "don't Know" responses have not been considered in the analyses; therefore, our target sample consists of 21,991 individuals at EU28 level.

Given the nature of the SAH variable, Ordered Probit models have been implemented, for the two subsets of risk factors separately and altogether (estimates are obtained using STATA14 where the dummy variables are treated as usual as factors). Although, with so many variables to be considered, the relationship between the type of risk and self-reported health was difficult to read and globally interpret. To cope with this problem, we have derived two composite indicators, obtained by Principal Component Analysis (Jolliffe, 2011), and then employed them as explanatory variables in a further model implementation.

3. Results

3.1 Derivation of the PCs and composite indicators

To better synthesize data, a Principal Components Analysis (PCA) has been performed distinctly on the two sets of variables, the one comprising the physical risk factors and the other referred to the psychosocial ones, and two composite indicators have been built. For homogeneity purposes, the analysis has been conducted on the correlation matrix.

For space constraints, we do not report in this paper tables and figures which may be available from Authors. To summarise, some variables show a strong correlation with health and well-being of workers. This is the case of variables related to positions or movements during work, and the same can be said for those related to work-life balance or to a positive and motivating work environment, in which workers have a sense of fulfilment with work, feel motivated, have a say, are consulted and participate in decisions, are supported by management and trust managers and, finally, experience good relationships with colleagues

PC1.1: Physical risk factors.

CLEUP Coop. Libraria Editrice ISBN: 978-88-5495-135-8

³ The proportion of those claiming a bad or very bad SAH is about 2%, while more than 77% report a positive o very positive evaluation with reference to SAH, and there are not prominent differences in the frequency distributions with regards to gender and work sector.

⁴ For brevity, we do not report here the lists of selected variables for the risks in the workplaces, as well as tables for detailed frequency distribution, break downs by gender, sector and type of contract, which are available from Authors.

For these set of covariates, the total inertia of the data is p=15. The eigenvalue associated with the first PC is 4.71 while the second largest is much smaller (1.53). Indeed, the variance of the first PC (PC1) accounts, alone, for 31.4% of the total inertia (4.71/15*100), therefore suggesting that one dimension, provided by the first PC, is enough to synthetize information.

The PC1 thus, ranges from better to worse physical working conditions, and it is indeed negatively correlated to SAH: the higher (with a couple of exceptions) the value of the covariates (which, given the direction of the scale employed for coding, indicates worse physical working conditions), the higher the value of the PC1. Along PC1 it is possible to identify at one end workers with a high exposure to physical risk agents (with positive high scores on PC1) and at the opposite end workers in desk-based jobs (low negative score on the PC1).

PC1.2: Psychosocial risk factors.

The synthetic indicator has been derived also for the second set of variables describing psychosocial risks. In this case, the total inertia of the data is p=26. The eigenvalue associated to the first PC is 6.17 while the second largest is 2.63. The first PC explains the 23.7% (6.17/26*100) of the total inertia. As in the previous case, also for the psychosocial covariates, one dimension (the first PC) captures most of information in the data, although in this case a second PC might be considered.

Most of the variables are positively correlated to PC1: the higher the value of these variables, which denote a positive and motivating work environment, the higher the value of PC1. In this case, higher values of the synthetic indicator denote better working conditions.

3.2 Ordered Probit models including composite indicators, discussion and limitations of the study

Based on the findings of the PCA, three Ordinal Probit models have been estimated (see Table 1), considering as explanatory the respondents' covariates and the first PC derived from the first set of variables, denoted as *PC1.1*, and from the second set of variables, denoted as *PC1.2*.

<u> </u>																		
SAH	Mod 1 (PC1.1 and PC1.2)			Mod 2 (PC1.1: Physical synthetic indicator)			Mod3 (PC1.2: Psychosocial synthetic indicator)											
										Coef.	Std. Err.		Coef.	Std. Err.		Coef.	Std. Err.	
										2.gender	-0.083	0.016	***	-0.108	0.016	***	-0.023	0.016
	Age	-0.031	0.001	***	-0.029	0.001	***	-0.030	0.001	***								
1.tertiary	0.071	0.017	***	0.076	0.017	***	0.073	0.017	***									
1.permjob	-0.096	0.021	***	-0.089	0.021	***	-0.085	0.021	***									
1.fulltime	0.066	0.025	***	0.085	0.025	***	0.075	0.025	***									
1.private	0.027	0.016		0.021	0.016		0.013	0.016	***									
Whours	0.001	0.001		-0.001	0.001		0.001	0.001										
d4w	-0.002	0.011		-0.007	0.011		-0.004	0.011										
endsmeet	0.112	0.007	***	0.146	0.007	***	0.130	0.007	***									
1.deu12	0.088	0.016	***	0.061	0.015		0.094	0.016	***									
Pc1.1	-0.063	0.004	***	-0.080	0.004	**												
Pc1.2	0.096	0.003	***				0.104	0.003	***									
/cut1	-3.980	0.087		-3.830	0.085		-3.855	0.086										
/cut2	-3.051	0.0737		-2.926	0.072		-2.934	0.073										
/cut3	-1.732	0.070		-1.642	0.069		-1.629	0.070										
/cut4	-0.111	0.069		-0.061	0.068		-0.021	0.069										

Table 1: Ordered Probit Models with Physical and Psychosocial risk factors indicators

In all the cases, the synthetic indicators built for the two sets of risk, either together (Mod1) or alone (Mod2 and Mod3), turn out to be significant, confirming that they provide an effective synthesis of the underlying variables which exert an impact on SAH. Also, the respondent related

CLEUP Coop. Libraria Editrice ISBN: 978-88-5495-135-8

characteristics remain significant in the same way, for all the models.

It is worth to stress that the added value of building these synthetic indicators relies on that they allow either for simplifying an analysis or for disentangling specific drivers of work-related well-being, with the additional advantage of removing redundant information.

Nevertheless, it is important to also underline some limitations of this exercise, which stem directly from the data used and the survey itself. First, the physical risk factors are not extensively surveyed in the case of the EWCS and therefore they refer only to a small subset of the sample. Moreover, European and national legislations have targeted this type of risk factors for several decades now resulting in their steadily decrease. Another point regards the limitations stemming from the questionnaire that seem to include too many questions (and variables) and some questions appear to be repetitive as they seek to grasp sometimes the same concept. This seems confirmed by the circumstance that only the first principal component in the two groups of selected variables is significant; in addition, no great contrasts are captured. All in all, a questionnaire including fewer and more targeted questions would allow for gathering a better quality information and would be a more cost-effective solution.

References

- EU-OSHA (2013). *Psychosocial risks and workers' health*. OSHWiki (contributor: Hupke M.) available at: https://oshwiki.eu/wiki/Psychosocial_risks_and_workers_health [accessed 24/7/2019]. Eurofound (2017). Sixth European Working Conditions Survey, 2015. [data collection]. 4th Edition. UK Data Service. SN: 8098, http://doi.org/10.5255/UKDA-SN-8098-4.
- Howell, R.T., Kern, M.L., Lyubomirsky, S. (2007). Health benefits: Meta-analytically determining the impact of well-being on objective health outcomes. *Health Psychology Review*, **1**(1), pp. 83-136
- Jolliffe, I. (2011). *Principal component analysis* (pp. 1094-1096). Berlin Heidelberg: Springer Litchfield, P., Cooper, C., Hancock, C., Watt, P. (2016). Work and Wellbeing in the 21st Century, *International Journal of Environmental Research and Public Health*, **13**(11), pp. 1065.
- OECD (2008). Handbook on constructing composite indicators. Methodology and user guide. OECD, Paris.
- OECD (2013). OECD Guidelines of Measuring Subjective Well-Being. OECD, Paris.
- World Health Organization (2012). *Measurement of and target-setting for well-being: an initiative by the WHO Regional Office for Europe*, World Health Organisation: Geneva, available at: http://www.euro.who.int/ data/assets/pdf file/0003/180048/E96732.pdf [retrieved on line, 20/06/2019].

CLEUP Coop. Libraria Editrice ISBN: 978-88-5495-135-8