Book of the Short Papers

Editors: Francesco Maria Chelli, Mariateresa Ciommi, Salvatore Ingrassia, Francesca Mariani, Maria Cristina Recchioni

SIS Società Italiana di Statistica

UNIVERSITÀ POLITECNICA DELLE MARCHE

Dipartimento di Scienze Economiche e Sociali DISES

LIUC Università Cattaneo

LIUC Università Cattaneo | BUSINESS ANALYTICS AND DATA SCIENCE HUB

Pearson

CHAIRS

Salvatore Ingrassia (Chair of the Program Committee) - *Università degli Studi di Catania*
Maria Cristina Recchioni (Chair of the Local Organizing Committee) - *Università Politecnica delle Marche*

PROGRAM COMMITTEE

Salvatore Ingrassia (Chair), Elena Ambrosetti, Antonio Balzanella, Matilde Bini, Annalisa Busetta, Fabio Centofanti, Francesco M. Chelli, Simone Di Zio, Sabrina Giordano, Rosaria Ignaccolo, Filomena Maggino, Stefania Mignani, Lucia Paci, Monica Palma, Emilia Rocco.

LOCAL ORGANIZING COMMITTEE

Maria Cristina Recchioni (Chair), Chiara Capogrossi, Mariateresa Ciommi, Barbara Ermini, Chiara Gigliarano, Riccardo Lucchetti, Francesca Mariani, Gloria Polinesi, Giuseppe Ricciardo Lamonica, Barbara Zagaglia.

ORGANIZERS OF INVITED SESSIONS

Pierfrancesco Alaimo Di Loro, Laura Anderlucci, Luigi Augugliaro, Ilaria Benedetti, Rossella Berni, Mario Bolzan, Silvia Cagnone, Michela Cameletti, Federico Camerlenghi, Gabriella Campolo, Christian Capezza, Carlo Cavicchia, Mariateresa Ciommi, Guido Consonni, Giuseppe Ricciardo Lamonica, Regina Liu, Daniela Marella, Francesca Mariani, Matteo Mazziotta, Stefano Mazzuco, Raya Muttarak, Livia Elisa Ortensi, Edoardo Otranto, Ilaria Prosdocimi, Pasquale Sarnacchiaro, Manuela Stranges, Claudia Tarantola, Isabella Sulis, Roberta Varriale, Rosanna Verde.

FURTHER PEPOPLE OF LOCAL ORGANIZING COMMITTEE

Elisa D'Adamo, Christian Ferretti, Giada Gabbianelli, Elvina Merkaj, Luca Pedini, Alessandro Pionati, Marco Tedeschi, Francesco Valentini, Rostand Arland Yebetchou Tchounkeu

Technical support: Matteo Mercuri, Maila Ragni, Daniele Ripanti

PUBLISHED BY PEARSON
WWW.PEARSON.COM
ISBN 9788891935618AAVV

Contents

Statistical Machine Learning for environmental applications

Statistical Process Monitoring for Complex Data in Industry 4.0

Advances in Data Science and Statistical Learning [IMS Invited Session]

ENBIS Session: System Maintenance, Boosting algorithms for regression, and Research Excellence

Educational Data mining: methods for complex data in students' assessment

Spatial and Spatio-Temporal Modeling: Theory and Applications

Statistical Framework for Measuring the Sustainability of Tourism

Statistical learning for well-being analysis

Statistics in Society 1

Assessment and Education

Bayesian methods and applications 1

Health and mortality

Mixture Models

Sampling methods and analysis of survey data

Space-time statistics

Clustering and classification 1

Dynamic models and time series

Environmental learning and indicators

Health statistics 2

Migrants in Italy and return migration

Sustainability assessment

Bayesian methods and applications 2

Clustering and classification 2

Economics and labour markets

Environmental modeling

Multivariate data analysis 2

Statistics in Society 2

Bayesian methods and applications 3

Functional Data Analysis

Machine Learning and text mining

Multivariate data analysis 3

Nonparametric statistical methods

Regression modeling

Preface

This book includes the contributions presented at the Intermediate Meeting of the Italian Statistical Society (SIS) "SIS 2023 - Statistical Learning, Sustanaibility and Impact Evolution" held in Ancona at the Università Politecnica delle Marche, from June 21th to 23th of 2023.

The new challenges of digitalization, innovation and sustainability are showing the crucial role of data-driven approaches in supporting decision-making processes. Methodologies resulting from the integration of different know-how seem to be a reliable way to deal with the increasing need to measure the impact of the policies and to forecast scenarios. This meeting welcomed any attempt to face new challenges.

The conference registered more than 250 presentations, including 3 keynote speakers in 3 plenary sessions and 72 presentations in 24 invited sessions, all dealing with specific themes in methodological and/or applied statistics and demography. Furthermore, more than 180 contributions, with one or more authors, have been spontaneously submitted to the Program Committee and arranged in 30 contributed sessions.

The numerous participation of researchers in the conference shows how the challenges of sustainability, in its broadest sense, are of interest to both methodological and applied statistics.

With the publication of this book, we wish to offer to all members of the Italian Statistical Society, all international academics, researchers, Ph.D. students, and all interested practitioners, a good snapshot of the on-going research in the statistical and demographic fields.
We aim to provide all members of the Italian Statistical Society - as well as international academics, researchers, Ph.D. students, and interested practitioners - with a comprehensive overview of the ongoing research in the fields of statistics and demography.

We extend our heartfelt gratitude to all the contributors for submitting their works to the conference and to the researchers for their outstanding job in serving as referees and discussants with precision and timeliness.

A special appreciation goes to the Scientific and Organizational Committees for their tremendous efforts in managing all the organizational aspects, as well as to the Università Politecnica delle Marche and the Department of Economic and Social Science for making this event possible.

Finally, we wish to express our gratitude to the publisher Pearson Italia for all the support received.

A latent variable approach to Millennials' knowledge of green finance

Maria Iannario[a], Alessandra Tanda[b], and Claudia Tarantola[b]

[a]Via L. Rodinò, 22, Naples, University of Naples Federico II; maria.iannario@unina.it
[b]Via San Felice, 5, Pavia, University of Pavia,; alessandra.tanda@unipv.it, claudia.tarantola@unipv.it

Abstract

In this paper, we consider a latent variable model for Millennials' knowledge of green finance. We draw inspiration from a survey conducted last year among students at Italian universities, as part of the COST project 'Fintech and Artificial Intelligence in Finance'. In the analysis, we consider a cumulative link model that accounts for heterogeneity in the response process. The model presented is discussed in a Bayesian framework.

Keywords: Heterogeneity of variance, MCMC, *numeracy*, ordinal responses, scale effects

1. Introduction

In the latest years, the attention devoted to climate change consequences has grown considerably together with the awareness on the need to pursue more sustainable growth and perform a transition to greener economies (16). The interest in this topic comes from many directions. Companies adopt greener practices to reduce their environmental impact to attract investors and reduce their cost of funding, also for reputational concerns, see (23) and (30) among others. Investors place a lot of emphasis on ESG performance in their investment decisions in order to drive economic growth according to their values and to achieve long-term sustainable performance; see e.g. (8), (10) and (29). Policymakers developed guidelines and action plans to promote sustainable growth and manage the risks of climate change, accompanying companies to the transition (15); (13); (27); and, in addition, civil society has become more aware and attentive to sustainability and to the impact of their choices as citizens.

Among consumers and potential investors, Millennials find themselves in the position to effectively contribute to greener growth, through their purchase and investment choices. Millennials are individuals born between 1980 and 2000 (26); they generally show a strong sensitivity towards sustainability (4) and different risk attitudes and investment habits than the older generations (3). However, as the literature shows, this focus on sustainability issues does not always translate into more sustainable purchasing (4; 25). The authors of (25) claim that the sustainable behaviour of Millennials depends on their specific characteristics, including their ecological knowledge or *eco-literacy* (20).

Starting from these findings rooted in the managerial and consumers' behaviour literature, we extend the idea to the financial sphere of decisions. As the availability of funds is crucial to the transition and to the development of a greener economic system (15), it is important that savings are redirected towards those activities that are able to pursue a more balanced and inclusive growth, also contributing to the achievement of the Sustainable Development Goals (SDGs), issued by the UN (see https://sdgs.un.org). Globally, Millennials are estimated to have contributed to a surge in the amount invested in ESG funds (24).

Nevertheless, their contribution may currently be hampered by a low level of knowledge in finance and, specifically, 'green finance'. The level of financial knowledge and education among households remains very low, especially among the young investors. Bank of Italy reports that in 2020 the average level of financial literacy in Italy was 11.2 (out of 21); see (11) and (14). Additionally, young people (less than 35 years old in the survey) show a lower level of financial literacy than their older peers. Young people scored less than the average, especially for the component 'financial attitude' that investigates the ability to plan financial decisions from a long-term perspective. In an international comparison, Italy scores poorly compared to other countries and this further underlines the critical situation of the young Italian generation in terms of financial literacy. Furthermore, the Edufin Committee provides evidence on the lack of knowledge of sustainability in finance and 'green finance' in Italy (14), reporting that around 40% of the surveyed population has never heard the word 'ESG' and 1/5 never heard about 'sustainable finance'. Similarly, Millennials perform slightly better for ESG, with about 35% having never heard of 'ESG'; 26% have never heard of 'sustainable finance'. To foster Millennials' participation as investors in the transition to sustainable growth, it is necessary to understand the factors that determine their knowledge and behaviour.

Based on the previous considerations, this paper addresses the assessment of the degree of knowledge of 'green finance' by Millennials and the factors that determine this outcome. We examine a set of data collected via a survey distributed among students at Italy's largest university. The aim of the survey was to evaluate students' financial literacy. To assess their knowledge, respondents reported scores to items on ordinal rating scales. The observed score can be considered as a discretization of an underlying unobserved (latent) continuous variable, with every possible score for the ordinal response corresponding to an interval of the latent variable. The analysis presented here has been performed by means of a latent trait model accounting for heterogeneity in the response process (22). We work in a Bayesian framework, gaining flexibility in specifying the model and enhancing richness and accuracy in providing parameter estimates, see (17) for an application in a student evaluation context. A further advantage of the Bayesian setup is the possibility to use the same framework and approach even if the sample size is small (as in the examined case). For a general discussion of the advantages of the use of a Bayesian approach for ordinal data see, e.g, (21) and references therein. For some milestones regarding Bayesian approach in the context of ordinal data see (1), (2), (12), (18) and (19).

To the best of our knowledge, no previous study investigates the degree of awareness of the existence of 'green finance' among Millennials and whether this knowledge depends on their specific characteristics and aptitudes, such as numeracy.

The paper is organized as follows. The next section provides a brief description of the considered latent trait model in its location-scale version. Section 3. introduces the examined survey data and presents the Bayesian estimates of the examined model. Section 4. presents some preliminary findings deriving from the Bayesian analysis of the examined model.

2. Latent trait model description

Let $(x_i, y_i), i = 1, \ldots, n$; be a data set of size n, where the covariates x_i's are assumed to be non-stochastic in R^p. The i-th measurement y_i is the realization of a random variable $Y \sim G(y)$ conditioned on x_i. Variable Y takes values in the finite set $\chi = \{1, \ldots, k\}$. Following (22), we assume that there exists an unobserved latent random variable Y_i^* such that when $\alpha_{j-1} < Y_i^* \leq \alpha_j$, then $Y_i = j\,, j \in \chi$. Here $-\infty = \alpha_0 < \alpha_1 < \ldots < \alpha_k = +\infty$ are the thresholds in the continuous support of the latent variable Y^*.

The i-th rating of Y^* linearly depends on $p \geq 1$ covariate(s) through $\boldsymbol{x}_i$ as $Y_i^* = \boldsymbol{x}_i\boldsymbol{\beta} + \sigma_i\epsilon_i$, $i = 1, 2, \ldots, n$, where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$ are the covariates coefficients. In the latent regression σ_i is the standard deviation of the noise variable $\epsilon \sim F_\epsilon(.)$, which may depend on covariates yielding $\sigma_i = \exp(\boldsymbol{z}_i\boldsymbol{\gamma})$. Here $\boldsymbol{z}_i$ is a row vector of the matrix $\boldsymbol{Z}$ which includes all the $q \geq 1$ relevant covariates and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_q)'$ are the related covariates coefficients.

The probability mass function of Y_i, for $j = 1, 2, \ldots, k$, can be expressed as

$$\begin{aligned} Pr\left(Y_i = j \mid \boldsymbol{\theta}, \boldsymbol{x}\right) &= Pr\left(\alpha_{j-1} < Y_i^* \leq \alpha_j\right) \\ &= F_\epsilon\Big[(\alpha_j - \boldsymbol{x}_i\boldsymbol{\beta})/\sigma_i\Big] - F_\epsilon\Big[(\alpha_{j-1} - \boldsymbol{x}_i\boldsymbol{\beta})/\sigma_i\Big]. \end{aligned}$$

Among the alternative choices for $F_\epsilon(.)$ we focus on the logit link function for robustness properties.

Since we do not have relevant prior information, following the approach prosed by (6) and (7), we use non-informative priors on all parameters of interest, letting the data guide the behavior of the posterior distributions. More precisely on the covariate coefficients we assign improper uniform priors, $unif(-\infty, +\infty)$, while on the thresholds we consider Student-t priors with 3 degrees of freedom. This ensures that the tails are wide, while the distribution is still proper with finite mean and variance (here set –without loss of generality– equal to 0 and 2.5 respectively). In order to obtain posterior samples we rely on Markov Chain Monte Carlo (MCMC) method. In particular, we use `R` package `brms` (6), which implements a Hamiltonian MCMC using Stan (5). The ordering of the intercepts is ensured via the `order class` in Stan. More precisely the joint prior distribution is truncated to support over points satisfying the ordering constraints.

3. Data

The data come from the survey on 'Knowledge and Use of Fintech Products' administered as part of the European project CA19130 Fintech and Artificial intelligence in Finance. The sample consists of 385 Italian Millennials, of whom 0.511 are women, 0.563 are from Southern Italy, 0.210 are studying economics at university. The distribution of the dependent variable on the level of knowledge of green finance (ranging from 0 =I don't know to 5 =I know it perfectly) is in Figure 1.

Among the *numeracy* variables considered are the self-assessment of how good the respondent perceives himself/herself to be with fractions (0.388 Millennials give a good rating) and the correct answers to the Berlin numerical test (multiple-choice format) (9) in which the respondents who correctly formulated all answers to the four questions are indicated by *top* (0.035 respondents). Difficulties with financial language and stress in relation to financial decisions are expressed by 0.703 and 0.696 of the respondents, respectively.

The Bayesian estimates of the location and scale parameters are reported in Table 1 (posterior mean, MCMC Standard Error and 95% credible intervals). Standard convergence diagnostics have been considered. Possible interactions among covariates were tested, but they were related to not statistically significant parameters.

We run in parallel 4 chains of 2000 iteration with a burnin period of 1000 iteration each. The Bayesian estimate of the standard deviation is obtained from the posterior samples of log-disc (log-discrimination) with disc corresponding to the inverse of the standard deviation. More precisely, for every iteration t, $t = 1, 2, \ldots, T$, we transformed $\log(disc)^t$ to sd^t with $sd^t = 1/exp(log(disc))$; the Bayesian estimates of sd is then obtained as the average of all sd^t.

4. Preliminary results

Preliminary results show that the knowledge of 'sustainable finance' or 'green finance' is associated with a set of financial education items and *numeracy*. We also observe that knowledge of Fintech innovations, such as Crowdfunding, Roboadvisor and advanced technologies (AI) is associated with knowledge of 'green finance'. These two elements, financial technology, and sustainability are, in fact, often coupled together by policymakers to foster more inclusive growth. It is therefore not surprising that Millennials who are more attentive to Fintech innovations are also more aware of green finance issues.

In addition, Millennials attending university courses in the field of economics have a greater knowledge of green finance. This could be due to the inclusion of ESG and sustainability topics in courses

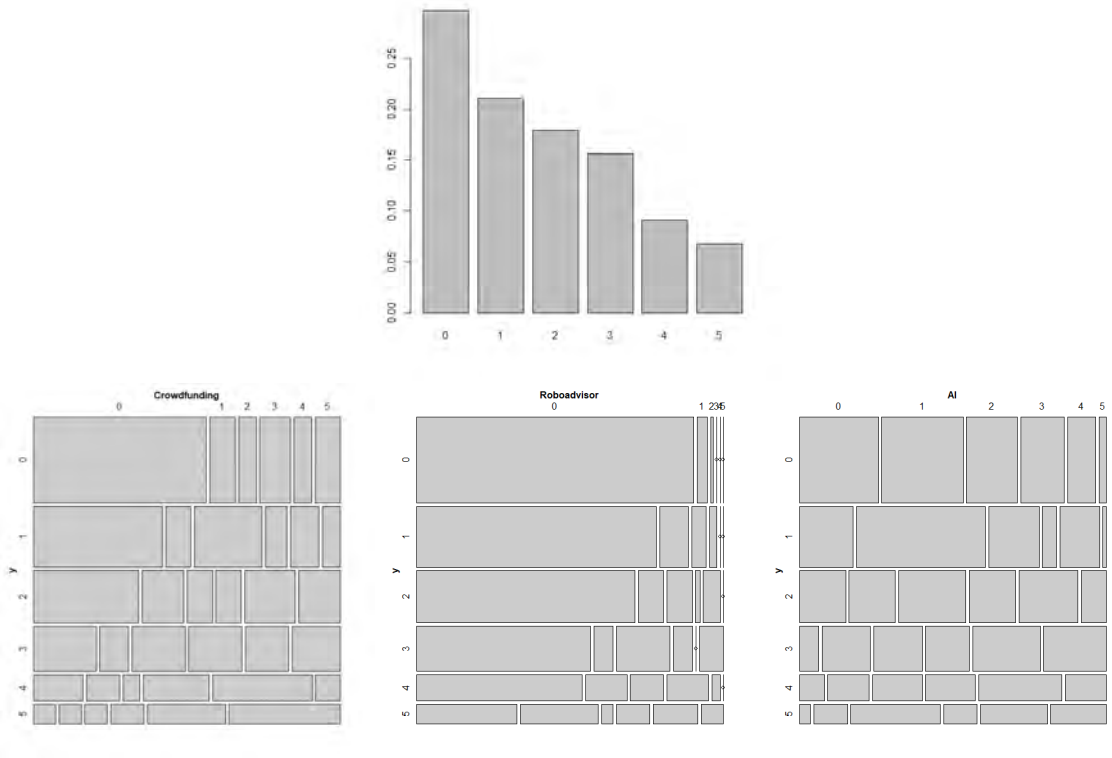

Figure 1: First row: Frequency distribution of the ordinal variable. Second row: Mosaic plot for the dependent variable vs Crowdfunding (first), Roboadvisor (second), and Artificial Intelligence - AI (third).

taken by university students in the field of finance. Numerical skills provide two seemingly counterintuitive results. Those who have a better ability to use fractions have a greater knowledge of 'green finance'. But the better performers in the Berlin numeracy test have less knowledge of our variable of interest Y.

Numeracy is a strong component of financial education and a basic skill for taking good financial decisions (28). However, in the field of 'green finance', this does not seem to be enough. In our analysis, the Millennials who score highest on *numeracy* tests have the least knowledge of green finance. This result may be determined by those Millennials studying in a STEM degree programme, who have excellent numerical skills, but have not yet encountered the topic of 'green finance'. We then evaluate the effect of financial knowledge in terms of familiarity with financial language and decisions.

Respondents who find financial language more difficult are also less knowledgeable about 'green finance'. At the same time, those who are more stressed when making financial decisions are more familiar with green finance. This could be due to the fact that stressed people might learn more before making decisions and, by doing so, become more aware of the different instruments available in the financial markets. This aspect probably deserves further investigation.

Our results are relevant for policy makers and for the design of future financial education initiatives. Further efforts should be devoted to increasing awareness of these key sustainability issues, along with general financial knowledge and *numeracy*.

Acknowledgments This work acknowledges research support by COST Action CA19130 'Fintech and Artificial Intelligence in Finance - Towards a transparent financial industry' (FinAI), supported by COST (European Cooperation in Science and Technology).

Table 1: Bayesian estimation for the location-scale model; i.e., posterior mean estimates, standard deviations and 95% Credible Intervals (CI) for its parameters

	Estimate(Sd)	l-95% CI	u-95% CI
$\hat{\alpha}_1$	0.80(0.97)	-0.97	2.92
$\hat{\alpha}_2$	5.72(1.62)	3.10	9.56
$\hat{\alpha}_3$	10.20(2.47)	6.20	15.73
$\hat{\alpha}_4$	15.64(3.61)	9.83	23.70
$\hat{\alpha}_5$	21.86(5.10)	13.78	33.36
Crowdfunding	1.00(0.32)	0.47	1.68
Roboadvisor	1.54(0.52)	0.62	2.69
AI	0.92(0.35)	0.33	1.72
economics	2.82(1.19)	0.72	5.33
fractions	0.99(0.34)	0.44	1.74
top	-6.26(2.90)	-12.62	-1.36
financial-language	-1.47(0.47)	-2.51	-0.69
stress-fin	0.90(0.42)	0.19	1.79
log_disc_info	1.23(0.06)	1.11	1.35
log_disc_franctions	1.13(0.04)	1.05	1.22
log_disc_sud	1.26(0.15)	0.99	1.57

References

[1] Albert, J. H. and Chib, S.: Bayesian Analysis of Binary and Polychotomous Response Data. *JASA*, **88**, 422, 669–679 (1993)

[2] Albert, J. and Chib, S.: Bayesian methods for cumulative, sequential and two step ordinal data regression models. Technical report (1997)

[3] Beck, J. J., Garris III, R. O.: Managing personal finance literacy in the United States: A case study. Educ. Sci., **9**, 129 (2019)

[4] Bernardes, J. P., Ferreira, F., Marques, A. D., Nogueira, M.: Millennials: is 'green' your colour?. In IOP Conference Series: Materials Science and Engineering (Vol. 459, No. 1, p. 012090). IOP Publishing (2018)

[5] Betancourt, M., and Girolami, M.: Hamiltonian Monte Carlo for Hierarchical Models. arXiv 1312.0906. http://arxiv.org/abs/1312.0906 (2013)

[6] Bürkner, P.: brms: An R Package for Bayesian Multilevel Models Using Stan.J. Stat. Softw., **80**, 1–28 (2017)

[7] Bürkner, P., Vuorre M.: Ordinal Regression Models in Psychology: A Tutorial. AMPPS, **2**, 77-101 (2019)

[8] Chauhan, Y., Kumar, S. B.: Do investors value the nonfinancial disclosure in emerging markets?. Emerg. Mark. Rev., **37**, 32-46 (2018)

[9] Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., Garcia-Retamero, R.: Measuring risk literacy: The Berlin Numeracy Test. JDM, **7**, 25-47 (2012).

[10] Cornell, B. ESG preferences, risk and return. Eur. Financial Manag., **27**, 12-19 (2021)

[11] D'Alessio, G., De Bonis, R., Neri, A., Rampazzi, C.: Financial literacy in Italy: The results of the Bank of Italy's 2020 survey. Politica economica, **37**, 215-252 (2021)

[12] Dellaportas, P. Smith, A.F.M.: Bayesian Inference for Generalized Linear and Proportional Hazards Models via Gibbs Sampling. JRSS C, 42, 443-459 (1993)

[13] ECB: Climate-related risk and financial stability. July 2021. Available at https://www.ecb.europa.eu/pub/pdf/other/ecb.climateriskfinancialstability202107 87822fae81.en.pdf

[14] Edufin: Rapporto Edufin 2021. Available at https://www.quellocheconta.gov.it/it/news-eventi/rassegna/Rassegna-Stampa/news_138.html

[15] European Commission: The European Green Deal, COM(2019) 640 final, 11 December. $https : //ec.europa.eu/info/sites/info/files/european - green - deal - communication_en.pdf.$

[16] Hafner, S., Jones, A., Anger-Kraavi, A., Pohl, J.: Closing the green finance gap. A systems perspective.Environ. Innov. Soc. Transit., **34**, 26-60 (2020)
[17] Iannario, M., Kateri, M., Tarantola, C.: Modelling scale effects in rating data: a Bayesian approach. Manuscript (2023)
[18] Johnson, V. E.: On Bayesian analysis of multirater ordinal data: An application to automated essay grading. JASA, 91, 42-51 (1996)
[19] Johnson V.E., Albert J.H.: Ordinal Data Modeling. Springer-Verlag, New York (1999)
[20] Kanchanapibul, M., Lacka, E., Wang, X., Chan, H. K.: An empirical investigation of green purchase behaviour among the young generation. J. Clean. Prod., **66**, 528-536 (2014)
[21] Liddell, T. M., Kruschke, J.K.: Analyzing ordinal data with metric models: What could possibly go wrong? J. Exp. Soc. Psychol., **79**, 328–348 (2018)
[22] McCullagh, P.: Regression models for ordinal data (with discussion). *J. R. Stat. Soc. B*, **42**, 109–142 (1980)
[23] Miroshnychenko, I., Barontini, R., Testa, F.: Green practices and financial performance: A global outlook. J. Clean. Prod., bf147, 340-351 (2017)
[24] MSCI: Swipe to invest: the story behind millennials and ESG investing. (2020) Available at https://www.msci.com/documents/10199/07e7a7d3-59c3-4d0b-b0b5-029e8fd3974b
[25] Naderi, I., Van Steenburg, E.: Me first, then the environment: Young Millennials as green consumers. Young Consumers (2018)
[26] Ng, E. S., Schweitzer, L., and Lyons, S. T.: New generation, great expectations: A field study of the millennial generation. JBP, **25**, 281-292 (2010)
[27] OECD: Financial Markets and Climate Transition: Opportunities, Challenges and Policy Implications, OECD Paris, (2021) https://www.oecd.org/finance/Financial-Markets-and-ClimateTransition-Opportunities-challenges-and-policy-implications.htm
[28] Sunderaraman, P., Barker, M., Chapman, S., Cosentino, S.: Assessing numerical reasoning provides insight into financial literacy. Appl. Neuropsychol. Adult, **29**, 710-717 (2022)
[29] Van Duuren, E., Plantinga, A., Scholtens, B.: ESG integration and the investment management process: Fundamental investing reinvented. J. Bus. Ethics, **138**, 525-533 (2016)
[30] Yu, E. P. Y., Tanda, A., Luu, B. V., Chai, D. H.: Environmental transparency and investors' risk perception: Cross-country evidence on multinational corporations' sustainability practices and cost of equity. BSE, **30**, 3975-4000 (2021)