***Controlling biases***

***Using of Likert-scales of 11 points***

*Batista-Foguet et al. (2009, p. 580) conclude that “the 11-point scale should be used (due to) its higher measurement quality in our European and Latin American sample”. These same authors evidence that “the tradition of using 11-point scales in Europe appears in academic and educational settings from primary school upwards” (Batista-Foguet et al. (2009, p. 575). As we performed our analysis in the Spanish context, we decided to use this scale, which is really frequent in this cultural environment.*

***Common method bias***

*Following Chang et al (2010) and Podsakoff and Organ (1986), we adopted ex-ante measures at the research design stage to control for the potential bias associated with the analysis of data obtained from the same respondents and in the same measurement context, using the same item context and similar item characteristics:*

*a) To reduce potential respondent misunderstandings, the questionnaire was validated by a panel of experts from the academic (5 members), professional (5 members) and public administrations (2 members), who were asked to assess the content of the questions, as well as the way in which these questions were presented, and also asked to evaluate the comprehension level of the indicators. The five university professors belonged to four research areas: general management (2), marketing (1), accounting (1) and statistics (1). The five professionals were managers of nonprofit organizations (2), managers from two Aragonese associations for nonprofit organizations (2) and an auditor specialized in nonprofit accounting (1). Finally, we included two public servants, the registrar of nonprofit organizations in Aragon (1) and the responsible for the monitoring office for nonprofit organizations in this region (1). The feedback from the panel was included in a revised version of the questionnaire.*

*b) We also provided respondents with the opportunity to complete the survey guarantying complete confidentiality, making specific reference to current data protection legislation. It was highlighted that there were being no right or wrong answers, respondents should answer the survey questions as honestly as possible.*

*c) Additionally, we considered other ex-ante remedies to prevent potential bias in our data analysis (Chang et al., 2010; Podsakoff, et al., 2012). We conducted a pilot survey to validate and test our measures in one nonprofit organization. After receiving verbal and written feedback from respondents, we reviewed and polished some of the items in our measures. Secondly, we focused our survey on the managers of Aragonese nonprofit organizations and we gathered our data using an on-line survey. We also reassured them about the confidentiality of the data. Participants were made aware of the fact that the survey data was for research purposes only rather than commercial or political.*

*d) Although ex-ante remedies are needed to reduce problems in the comprehension phase of the survey process; they are not sufficient to enable us to completely avoid potential selection and common method biases (Chang et al., 2010). Hence, we applied ex-post statistical control strategies to test for common method bias. To control for common method variance, we used Harman’s single-factor test. All the observed variables were subjected to principal components analysis, and the first un-rotated component explains less than 14.398 percent of the variance. Furthermore, we estimated a single-factor confirmatory model for each dimension. The estimated models had a bad fit:*

*a) Economic dimension: (2[44]=474.537, RMSEA=0.283, SRMR=0.201 and CFI=0.341).*

*b) Social dimension: (2[119]=360.097, RMSEA=0.109, SRMR=0.087 and CFI=0.691).*

*c) Governance dimension: (2[299]=856.726, RMSEA=0.105, SRMR=0.106 and CFI=0.501)*

*Consequently, no single factor can explain the data structure. Based on the results of the Harman single-factor test, the goodness of fit of the single-factor confirmatory model and the proposed theoretical model we do not expect common method variance to be a significant problem in our research.*

*e) Finally, Chang et al (2010) consider that “common method variance is more likely to emerge in models that are overly simple”. In this study, we combine information from difference sources: a) The financial statements and activity reports deposited in the public register and, b) The information from the questionnaire. Additionally, we also highlight that the specified models are estimated step-by-step. In this sense, first, we estimate the economic, social and governance dimensions using confirmatory factor analysis. In a second step, and after obtaining the factor scores, we estimate the structural model which enables us to evaluate the relationship between the dimensions and the income structure of nonprofit organizations. Both aspects, the mixture of information from different sources together with the characteristics of the proposed model, reduce the impact of common bias variance.*

***Selection bias***

*Following Armstrong and Overton (1977), we compared early and late respondents of the questionnaire in the observed variables. The statistical tests for differences between means were not significant at 0.10, which suggests that this bias was not a serious problem in this research. To do so, we have split the sample considering the median (50%/50% of participants). Table 1 shows the results for the analysed variables.*

***Table 1.*** *t-test for independent samples (exogenous variables)*

|  |  |  |
| --- | --- | --- |
| **Dimension** | **Indicator** | **50/50****t-test****(pvalue)** |
| **Economic** | **Continuous activity** | The execution of the financial plan enables the detection of possible weaknesses | 0.793(0.429) |
| The nonprofit has a monitoring system in the execution of the financial plan | 0.276(0.783) |
| New courses of action are defined if deviations from the financial plan are detected | 0.289(0.773) |
| **Social** | **Social mission** | Users participate in the decision-making processes | 1.485(0.140) |
| Users share organizational ethics values | 1.637(0.104) |
| The organization has sufficient resources to lend its services | 0.836(0.404) |
| There are mechanisms to monitor the service quality  | -0.445(0.657) |
| Impact of the paid workers recruitment in the local environment | 1.061(0.290) |
| The organization promotes the recruitment of paid workers | 1.494(0.137) |
| The organization provides formation for their paid workers | 1.127(0.261) |
| Impact of the participation of volunteers in the local environment | -1.084(0.280) |
| The organization promotes the participation of volunteers | 1.248(0.214) |
| The organization provides formation for their volunteers | 1.247(0.214) |
| **Civil society** | Level of development of the organizational ethical ideology | 0.642(0.522) |
| Level of integration of local community and its environment  | -0.638(0.525) |
| Development of service quality surveys among users | -0.445(0.657) |
| Promotion of high levels of accessibility to the lent services | -0.091(0.927) |
| Promotion of the participation in social networks | 0.633(0.528) |
| Estimation of the social return on investment (SROI) | 0.204(0.838) |
| Local community participates in the decision-making processes | 1.485 (0.140) |
| **Governance** | **Autonomy** | The organization shares workers with other organizations | 0.792 (0.429) |
| The organization shares volunteers with other organizations | 0.339 (0.735) |
| The organization progressively improves the formation of their human resources | 1.211 (0.228) |
| The organization progressively attracts a higher number of users  | 1.133 (0.259) |
| The organization progressively has more financial resources | 1.572 (0.118) |
| **Decision-making processes** | The board of trustees consults users before making a decision | 0.530 (0.597) |
| The board of trustees consults workers before making a decision  | -0.137 (0.891) |
| The board of trustees consults volunteers before making a decision | -1.257 (0.210) |
| The board of trustees consults funders before making a decision  | 1.095 (0.275) |
| **Stakeholder** **orientation** | The organization consolidates the existing users  | 1.579 (0.116) |
| The organization improves the relationship with external stakeholders  | 1.459 (0.146) |
| The organization shares information with its stakeholders | 0.799 (0.425) |
| The organization promotes the stakeholders’ participation in the decision-making  | 1.148 (0.252) |
| The organization facilitates accessible information to its stakeholders | 1.572 (0.118) |
| The organization maintain a high degree of visibility in its local community | 1.122 (0.264) |

*It is possible to observe that there are no differences between both groups in the analysed variables, which would be an indicator about the lack of significant of this bias in the performed research.*

*Moreover, we have developed a second ex-post analysis. As we have information extracted from the financial statements and activity reports deposited by nonprofit organization in the public register, we have tested the existence of possible differences between those who participated in the questionnaire and those who do not participate. This test was developed in relation to the endogenous variables (diversification indices) and the different income sources (Table 2).*

***Table 2.*** *t test for independent samples (endogenous variables)*

|  |  |  |  |
| --- | --- | --- | --- |
| **Indicator** | **t-test (pvalue)** | **Indicator** | **t-test (pvalue)** |
| Private donations share 2008 | 0,969 (0,334) | Herfindahl-Hirschman Normalised Index 2008 | 1,535 (0,126) |
| Private donations share 2018 | 0,180 (0,857) | Herfindahl-Hirschman Normalised Index 2018 | 1,211 (0,228) |
| Government funding share 2008 | -0,154 (0,878) | Entropy 2008 | -0,463 (0,644) |
| Government funding share 2018 | -1,165 (0,246) | Entropy 2018 | 1,156 (0,25) |
| Commercial income share 2008 | -0,828 (0,409) | Concentration ratio C2 2008 | 0,828 (0,409) |
| Commercial income share 2018 | -0,546 (0,586) | Concentration ratio C2 2018 | 0,868 (0,387) |
| Income from investments share 2008 | 0,549 (0,584) | Volatility income sources 2008 | 0,961 (0,338) |
| Income from investments share 2018 | -0,599 (0,55) | Volatility income sources 2018 | -0,291 (0,771) |

*We observe that there are no significant differences between those companies that participated in the questionnaire and the companies that do not (pvalue>0.100). This results would be a second sign about the scarce significant of this bias in the studied sample.*

*Additionally, we have tested the existence of possible differences between the sample and the population in the variables that describe the sample (activity, rural/urban, activity and initial endowment). To do so, we have proposed a Chi-squared test to identify differences between the defined categories. We obtained that there are no significant differences between the sample and the population when these characteristics are analysed (pvalue Chi-squared – activity: 0.224; pvalue Chi-squared – rural/urban: 0.157; Chi-squared – age: 0.199; pvalue Chi-squared – initial endowment: 0.201).*

***Sample size bias***

*Kline (2015) expose the absence of absolute standards in the SEM literature about the relationship between sample size and model complexity. Despite this asseveration, there is a rule-of-the-thumb that a minimum size of 100 cases is always required. In this sense, Anderson & Gerbing (1984, pp.170–171) conclude that for “three or more indicators per factor, a sample size of 100 will usually be sufficient for convergence,” and a sample size of 150 “will usually be sufficient for a convergent and proper solution”. Other authors increase this figure to 200 cases, although this number is considered to be “conservative and surely simplistic” (Iacobucci, 2010, p.92).*

*Consequently, the assessment of the sample size needs to take into account other indicators than can reveal the suitability of this statistical technic:*

*a) Kline (2015, p.112) introduces the ratio subject/parameter. This author established that ratio between 10:1 and 5:1 is a realistic target to assure the convergence of the model. The proposed model is based on three dimensions, but they have been estimated separately to assure the convergence of the achieved estimations and the suitability of this ratio.*

*b) MacCallum et al (1999) consider than “samples in the range of 100-200 are acceptable with well-determined factors”, understanding that factors will be well-determined if the communalities are in the range 0.5. In table 4, it is possible to observe that the different communalities are above this minimum value.*

***Table 4.*** *Communalities*

|  |  |  |
| --- | --- | --- |
| **Dimension** | **Communalities** | **Correlations** |
| Economic Dimension 1 | [0.797-0.999] | [0.826-0.926] |
| Economic Dimension 3 | [0.588-0.694] | [0.621-0.714] |
| Social Dimension 1 | [0.502-0.791] | [0.463-0.788] |
| Social Dimension 3 | [0.500-0.959] | [0.498-0.823] |
| Governance dimension 1 | [0.519-0.614] | [0.512-0.627] |

*Moreover, if we analyse the correlation matrices, we also observe that the different items are highly correlated with also reveals that the sample size is reasonable for the proposed measurement models.*

*c) Additionally, we have reestimated the proposed model selecting only 100 cases. Table 4 shows the value of the different indicators. If we compare the results in Table 5 and the results presented in Table 3, we observe that there are no significant variations in the main goodness-of-fit indices.*

***Table 5.*** *Global goodness-of-fit indicators for N=100*

|  |  |  |  |
| --- | --- | --- | --- |
| **N=100** | **Economic** **dimension** | **Social** **dimension** | **Governance** **dimension** |
| Chi-square | 9.477 | 68.423 | 13.184 |
| RMSEA | 0.041 | 0.056 | 0.104 |
| CFI | 0.997 | 0.954 | 0.914 |
| SRMR | 0.028 | 0.058 | 0.060 |

*Finally, the estimation step-by-step of the proposed models together with the rate of participation in the questionnaire (43%) enables us to use formative indicators guaranteeing the global fit. Moreover, SEM methodology guarantees the stability of the parameters in the model and it is suitable to answer the aim of this study.*

**References**

Anderson, J.C., & Gerbing, D.W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, *49*(2), 155-173.<http://dx.doi.org/10.1007/BF02294170>

Armstrong, J.S., & Overton, T.S. (1977). Estimating nonresponse bias in mail surveys. *Journal of Marketing Research, 14*(3), 396-402. <http://dx.doi.org/10.2307/3150783>

Batista-Foguet, J.M., Saris, W., Boyatzis, R., Guillén, L., & Serlavós, R. (2009). Effect of response scale on assessment of emotional intelligence competencies. *Personality and Individual Differences, 46*(5-6), 575-580. <http://dx.doi.org/10.1016/j.paid.2008.12.011>

Chang S., van Wittleloostuijn, A., & Eden, L. (2010). Common method variance in international business research, *International Business Studies, 41*, 178-184.<http://dx.doi.org/10.1057/jibs.2009.88>

Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of consumer psychology, 20*(1), 90-98. <http://dx.doi.org/10.1016/j.jcps.2009.09.003>

Kline R.B. (2015). *Principles and Practice of Structural Equation Modeling*. Guilford publications, New York.

MacCallum, R.C., Widaman, K.F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84-99. <http://dx.doi.org/10.1037/1082989X.4.1.84>

Podsakoff, P.M., & Organ, D.W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management, 12*(4), 531-544. <http://dx.doi.org/10.1177/014920638601200408>

Podsakoff, P.M., MacKenzie, S.B., & Podsakoff, N.P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*, 539-569. <http://dx.doi.org/10.1146/annurev-psych-120710-100452>