Socio-Economic Planning Sciences xxx (xxxx) xxx

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

journal homepage: www.elsevier.com/locate/seps

A composite indicator for the waste management in the EU via Hierarchical Disjoint Non-Negative Factor Analysis

Carlo Cavicchia [a,b,*], Pasquale Sarnacchiaro [a], Maurizio Vichi [b]

[a] University of Rome Unitelma Sapienza, Italy
[b] University of Rome La Sapienza, Italy

ARTICLE INFO

Keywords:
Specific indicators
Waste
Environment
Factor Analysis
Hierarchy

ABSTRACT

In the last years, the quantity of information and statistics about waste management are more and more consistent but so far, few studies are available in this field. The goal of this paper is of producing a model-based Composite Indicator of “good” Waste Management, in order to provide a useful tool of support for EU countries’ policy-makers and institutions.

Composite Indicators (CIs), usually, are multidimensional concepts with a hierarchical structure characterized by the presence of a set of specific dimensions, each one corresponding to a subsets of manifest variables. Thus, we propose a CI for Waste Management in Europe by using a hierarchical model-based approach with positive loadings. This approach guarantees to comply with all the good properties on which a composite indicator should be based and to detect the main dimensions (i.e., aspects) of the Waste Management phenomenon.

In other terms, this paper provides a hierarchically aggregated index that best describes the Waste Management in EU with its main features by identifying the most important high order (i.e., hierarchical) relationships among subsets of manifest variables. All the parameters are estimated according to the maximum likelihood estimation method (MLE) in order to make inference on the parameters and on the validity of the model.

1. Introduction

We are more than 7.5 billion people on our planet, and we are producing waste every day. The constant expanding of population implies an increasing generation of waste, and although the management of the waste keeps improving in the EU, many estimates tell us that half of that waste is not collected, treated or safely disposed of. That is why policy-makers need consistent and useful tools to measure and monitor waste management. The challenge is composed by two different aspects: sustainable consumption and smarter waste management. In order to plan a coordinate action through the EU countries, a reliable measure of “good” waste management is needed.

A multidimensional phenomenon like waste management is described by a huge quantity of information useful for making strategical decisions and the demand for statistics on waste generation and treatment has grown considerably in recent years. This amount of information needs to be synthesized by studying relationships among manifest (i.e., observed) variables. It is important to find the relationships among dimensions and manifest variables in order to synthesize the information and have a response on the conduct of each country to achieve the priority goals set by Europe, reducing waste generation and maximizing recycling and re-using. Identifying these relationships could be fundamental to understand where each country should focus its actions and what impacts each action could have. It is worth understanding how an effective waste management might impact other important social aspects, such as for example poverty [1]. A “good” waste management is vital for global sustainable development; it is connected with the Sustainable Development Goals (SDGs). The SDGs constitute the core of the 2030 Agenda for Sustainable Development [2], and their aim is to guide global, regional and national actions regarding development for the next 15 years. The United Nations Industrial Development Organization (UNIDO) contributes to the achievement of the SDGs by supporting Member States in achieving inclusive and sustainable development. Since the interconnected nature of the SDGs, many of UNIDO’s activities contribute to more than one goal. In our specific field of interest the Goal 12, named “Ensure sustainable consumption and production patterns”, has some targets to reach by 2030; among others: substantially reduce waste generation through reduction, recycling and reuse, and reduce food losses along production and supply

* Corresponding author at: University of Rome Unitelma Sapienza, Italy.
E-mail address: carlo.cavicchia@unitelmasapienza.it (C. Cavicchia).

https://doi.org/10.1016/j.seps.2020.100832
Received 7 October 2019; Received in revised form 25 February 2020; Accepted 15 March 2020

Table 1
List of manifest variables.

	Label	Name		Label	Name
1	GenWas	Generation of waste[a]	6	MatRec	Share of material recovered and fed back into the economy (%)
2	GenRec	Generation of recyclable waste[a]	7	ImpnEU	Imports from non-EU countries of recyclable raw materials[a]
3	PriInv	Private investments as value added at factor cost related to circular economy (€ per capita)	8	ImpEU	Imports intra EU countries of recyclable raw materials[a]
4	JobCir	Jobs as persons employed related to circular economy (% persons employed)	9	RecMun	Recycling rate of municipal waste (% generated waste)
5	PatRec	Patents related to recycling and secondary raw materials (number per capita)	10	RecWas	Recycling rate of all waste excluding major mineral waste (% generated waste)

[a]tonnes per capita.

Table 2
List of countries.

Name	Abbreviation	Country	Abbreviation
Belgium	(BE)	Bulgaria	(BL)
Czech Republic	(CZ)	Denmark	(DK)
Germany	(DE)	Estonia	(EE)
Ireland	(IE)	Greece	(GR)
Spain	(ES)	France	(FR)
Croatia	(HR)	Italy	(IT)
Cyprus	(CY)	Latvia	(LV)
Lithuania	(LT)	Luxembourg	(LU)
Hungary	(HU)	Malta	(MT)
The Netherlands	(NL)	Austria	(AT)
Poland	(PL)	Portugal	(PT)
Romania	(RO)	Slovenia	(SI)
Slovakia	(SK)	Finland	(FI)
Sweden	(SE)	United Kingdom	(UK)

chains. Other Goals are involved by waste management, for instance, plastics are devastating for the planet and its inhabitant, especially marine environment. In many areas of our planet it is common to find waste burned instead of collected properly, it has a huge impact on methane and CO_2 emissions and consequentially, on climate change. The quantity of information and statistics about waste management are more and more consistent but so far, few studies are available in this field. The aim of this work is of producing a model-based measure of "good" waste management, in order to provide a useful tool for policy-makers and institutions.

A usual way to synthesize a big amount of information is obtained by using Composite Indicators (CIs), that is, non-observable latent variables, linear combinations of observed variables [3] that are used to describe a complex phenomenon, in our case the WM. CIs have been considered to have several advantages: they can summarize multidimensional situations and facilitate evidence-based decision-making; they can be easier to interpret than a list of separate indicators; they facilitate communication among policy-makers, the media and the general public. However, CIs also have some potential disadvantages: their construction is particularly difficult since it requires both a sound scientific base and political consensus; they may send misleading messages about policy if poorly constructed (i.e., unreliable) or misinterpreted [4]; they may lead to over-simplistic conclusions, on behalf of both the general public and political actors.

In this paper we propose a CI for waste management in Europe by using a model-based approach in order to avoid or minimize the potential disadvantages. The model for the WM CI has a hierarchical structure formed by factors associated to subsets of manifest variables with positive loadings, by identifying the most important high order relationships among such subsets. This approach guarantees to comply with all the good properties on which a composite indicator – summarizing a multidimensional phenomenon – should be based. Such properties are: model-based, statistically estimated (i.e., non-normative), with a hierarchical structure, scale-invariant, uni-dimensional, reliable and non-compensable.

The general goal is to find, via statistical data modeling, the hierarchically aggregated index that best represents the waste management and its parameters are estimated according to the maximum likelihood estimation method (MLE) in order to make inference on the parameters and on the validity of the model. It is worthy of remark that our proposal allows us to carry out an exploratory analysis where all the parameters are simultaneously estimated, these features differentiates this model from a other sequential procedures or confirmatory analysis generally used to detect second-order (i.e. hierarchical) factor models.

The paper is organized as follows. Section 2 briefly introduces a review of different methodologies for composite indicators construction. In Section 3, the manifest variables about waste management and recycling are introduced. In Section 4, the *Hierarchical Disjoint Non-negative Factor Analysis* (*HDNFA*) model with its estimations is recalled. The results of the waste management are presented in Section 5. A final discussion completes the paper in Section 6.

2. Literature

Composite Indicators (CIs) are non-observable latent variables which are able to summarize big amount of information. CIs are very useful to measure multidimensional phenomena (e.g., socio-economic) and many methodologies for the construction of CIs have been proposed through the years. CIs frequently have been criticized because the methods for their construction are not always statistical and mathematically rigorous and often they are based on theories which do not seem to have a solid foundation [5]. Often CIs are computed as the weighted mean of the manifest variables (e.g., Multidimensional Poverty Index). The weights given to the manifest variables represent the importance of each variable and they are chosen by the researcher subjectively or according to a known theory. In particular, many authors do not appreciate CIs determined by subjective weights on the manifest variables because this approach can lead to misinterpretation of the results [6].

Multidimensional Data Analysis (MDA) approaches, like Factorial Analysis (FA) [7,8] or Principal Component Analysis (PCA) [9,10], are considered valid in order to build CIs for multidimensional phenomena. In PCA and FA, the weights are computed by taking into account the statistical relations among manifest variables. They represent reflective models thus they can be used whether all the variables refer to a general latent concept. Another widely used methods for the construction of CIs are Structural Equation Models (SEM) [11–13], they are used in order to build a flexible system of composite indicators able to model causal relations among them.

In [3], the Factor Analysis is considered as a weighting method in order to combine manifest variables. FA has the advantage to define loadings that best reconstruct the manifest variables according to the estimation method chosen avoiding the subjective choice of a system of weights given by the researcher. However, some choices are needed, for instance the choice of type of rotation (varimax, equimax, orthomax, etc.). When the model presents a hierarchical structure a valid alternative method to construct a CI into the factor analysis framework is the Hierarchical Confirmatory Factor Analysis [14–16], but this methodology needs a priori information by hypothesizing the most relevant relations in the hierarchy associated to the CI. In order to avoid the limitation to define a priori the most relevant relations, a two levels hierarchical factor model with simultaneous estimations of all loadings in order to build a CI for complex and multidimensional phenomena based on reliable and concordant specific factors has been considered.

3. Data

Waste management is a complex phenomenon, described by a huge quantity of information and its importance is crucial for the environmental and for the human life in general. It is more and more important to find the way to measure it in order to provide support for decision making. The number of statistics and measures related to waste collection, recycling and circular economy is expanding every year and the need of building aggregated index to monitor the countries' behavior in terms of policy is even more important. Since the aim of our analysis is the identification of a system of non-negative loadings in order to define a two-levels hierarchy with the general composite indicator as root, a preliminary analysis is needed to detect both: variables that are discordant with the general latent construct (i.e., waste management) and variables with relations not statistically significant. If variables (both manifest and latent) have positive loadings, they contribute positively to the construction of the general composite indicator. However, if variables (both manifest and latent) have negative loadings, they must be reversed. This is crucial for the definition of a general composite indicator (GCI) which avoids a compensation effect among variables. Another essential part of the preliminary analysis consists in detecting the presence of not statistically significant manifest variables into the model. They are discarded from the analysis because are not relevant, and they can assume a confounding role for the analysis.

After the preliminary steps of the analysis, the final data-set considered in this paper is composed of 10 manifest variables (Table 1) and 28 units (i.e., countries) (Table 2). Many variables about the characteristic of countries have been considered in order to help the interpretation of the results. The variables into the data-set come from different sources: Eurostat, Joint Research Centre, Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs (DG GROW) and the European Patent Office, and they are regularly updated and available for free on Eurostat website.[1]

4. Hierarchical Disjoint Non-Negative Factorial Analysis

4.1. Model

Hierarchical Disjoint Non-Negative Factorial Analysis (*HDNFA*) is a factorial model that considers two typologies of latent unknown constructs: H specific factors and a single (nested) general factor. *HDNFA* is identified by the two simultaneous equations:

$$\mathbf{x} - \mu_{\mathbf{x}} = \mathbf{A}\mathbf{y} + \mathbf{e}_{\mathbf{x}} \tag{1}$$

$$\mathbf{y} = \mathbf{c}\mathbf{g} + \mathbf{e}_{\mathbf{y}} \tag{2}$$

where $\mathbf{A}$ is the $J \times H$ matrix of unknown specific factors loadings, $\mathbf{c}$ is the $H \times 1$ vector of unknown general factor loadings, $\mathbf{e}_{\mathbf{x}}$ and $\mathbf{e}_{\mathbf{y}}$ are a $J \times 1$ and a $H \times 1$ random vector of errors, respectively.

[1] https://ec.europa.eu/eurostat/web/circular-economy/indicators/main-tables.

Let include model (2) into model (1) and consider the loading matrix $\mathbf{A}$ is restricted to the product $\mathbf{A} = \mathbf{B}\mathbf{V}$ [17], where $\mathbf{B}$ is a diagonal matrix and $\mathbf{V}$ a row stochastic and binary matrix, the *HDNFA* model, for n multivariate observation, is defined

$$\mathbf{X} = \mathbf{g}\mathbf{c}'\mathbf{V}'\mathbf{B} + \mathbf{E}_{\mathbf{x}}. \tag{3}$$

The variance–covariance structure related to the model (3) is

$$\Sigma_{\mathbf{x}} = \mathbf{B}\mathbf{V}(\mathbf{c}\mathbf{c}' + \Psi_{\mathbf{y}})\mathbf{V}'\mathbf{B} + \Psi_{\mathbf{x}} \tag{4}$$

where

$$\Sigma_{\mathbf{y}} = \mathbf{c}\mathbf{c}' + \Psi_{\mathbf{y}} \tag{5}$$

such that

$$\mathbf{V} = [\mathbf{v}_{jh} : \forall \mathbf{v}_{jh} \in \{0,1\}] \tag{6}$$

$$\mathbf{V}\mathbf{1}_H = \mathbf{1}_J \tag{7}$$

$$\mathbf{B} = \mathrm{diag}(b_1, \ldots, b_J) \text{ with } b_j^2 > 0 \tag{8}$$

$$\mathbf{V}'\mathbf{B}\mathbf{B}\mathbf{V} = \mathrm{diag}(b_{\cdot 1}^2, \ldots, b_{\cdot H}^2) \text{ with } b_{\cdot h}^2 = \sum_{j=1}^{J} b_{jh}^2 > 0. \tag{9}$$

It is assumed that $\mathbf{y} \sim N_H(0, \Sigma_{\mathbf{y}})$ where $\Sigma_{\mathbf{y}}$ is the correlation matrix of the specific factors since they are standardized, and $\mathbf{e}_{\mathbf{x}} \sim N_J(0, \Psi_{\mathbf{x}})$, where $\mathrm{cov}(\mathbf{e}_{\mathbf{x}}) = \Psi_{\mathbf{x}}$ is the J-dimensional diagonal positive definite variance–covariance matrix of the error of model (1) and $\mathrm{cov}(\mathbf{e}_{\mathbf{x}}, \mathbf{y}) = 0$. Furthermore, $\mathbf{g}$ is the random general factor with mean 0 and variance $\sigma_{\mathbf{g}}^2 = 1$ denoting the composite indicator related to a reduced set of specific factors. In addition, $\mathbf{e}_{\mathbf{y}}$ is a non-observable $H \times 1$ random vector of errors. It is assumed that $\mathbf{g} \sim N(0,1)$ and $\mathbf{e}_{\mathbf{y}} \sim N_H(0, \Psi_{\mathbf{y}})$, where $\mathrm{cov}(\mathbf{e}_{\mathbf{y}}) = \Psi_{\mathbf{y}}$ is the H-dimensional diagonal positive definite variance–covariance matrix of the error of model (2). In addition it is assumed that errors in the two models are uncorrelated $\mathrm{cov}(\mathbf{e}_{\mathbf{x}}, \mathbf{e}_{\mathbf{y}}) = 0$; and errors and factors are uncorrelated, i.e., $\mathrm{cov}(\mathbf{e}_{\mathbf{x}}, \mathbf{g}) = 0$ and $\mathrm{cov}(\mathbf{e}_{\mathbf{y}}, \mathbf{g}) = 0$.

4.2. Estimation

Suppose that a random sample of $n > J$ multivariate observations of $\mathbf{x}$ is observed, the maximization of the log-likelihood with respect to $\boldsymbol{\mu}$ gives the sample mean, thus the reduced log-likelihood is as follows

$$L(\mathbf{x}_i, \mathbf{A}, \Psi_{\mathbf{x}}, \Psi_{\mathbf{y}}) = \tag{10}$$

$$= -\frac{nJ}{2}\log 2\pi - \frac{n}{2}\{\log|\mathbf{A}(\mathbf{c}\mathbf{c}' + \Psi_{\mathbf{y}})\mathbf{A}' + \Psi_{\mathbf{x}}| + \mathrm{tr}\{[\mathbf{A}(\mathbf{c}\mathbf{c}' + \Psi_{\mathbf{y}})\mathbf{A}' + \Psi_{\mathbf{x}}]^{-1}\mathbf{S}\}\}$$

where $\mathbf{S} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu}_{\mathbf{x}})'\Sigma_{\mathbf{x}}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_{\mathbf{x}})$

This is equivalent to the minimization of the discrepancy function

$$D(\mathbf{x}_i, \mathbf{A}, \Psi_{\mathbf{x}}, \Psi_{\mathbf{y}}) = \log|\mathbf{A}(\mathbf{c}\mathbf{c}' + \Psi_{\mathbf{y}})\mathbf{A}' + \Psi_{\mathbf{x}}| + \mathrm{tr}\{[\mathbf{A}(\mathbf{c}\mathbf{c}' + \Psi_{\mathbf{y}})\mathbf{A}' + \Psi_{\mathbf{x}}]^{-1}\mathbf{S}\}. \tag{11}$$

This is a discrete and continuous problem that cannot be solved by a quasi-Newton type algorithm, it is solved by a coordinate descent algorithm. A general composite indicator should be composed by consistent and reliable specific composite indicators; thus we require that loadings must be positive during the estimation of $\mathbf{Y}$ and $\mathbf{g}$. So the discrepancy function (11) is minimized with respect to $\mathbf{B}_h = \mathrm{diag}(\mathbf{b}_h)$ by

$$\hat{\mathbf{b}}_h = \hat{\Psi}_{\mathbf{x}h}^{-\frac{1}{2}}\mathbf{u}_{1h}(\lambda_{1h} - 1)^{\frac{1}{2}} \tag{12}$$

where λ_{1h} and $\mathbf{u}_{1h}$ are respectively the largest eigenvalue and the corresponding eigenvector of the variance–covariance matrix $\hat{\Psi}_{\mathbf{x}h}^{-\frac{1}{2}}\mathbf{S}_h\hat{\Psi}_{\mathbf{x}h}^{-\frac{1}{2}}$ corresponding to variables identified by $\mathbf{v}_{\cdot h}$, that corresponds to hth column of $\mathbf{V}$. It is important to notice that λ_{1h} and $\mathbf{u}_{1h}$ minimize the function

$$\|\mathbf{X}_h\hat{\Psi}_{\mathbf{x}h}^{-\frac{1}{2}} - \sqrt{\lambda_{1h}}\mathbf{y}_h\mathbf{u}_{1h}'\|^2 \tag{13}$$

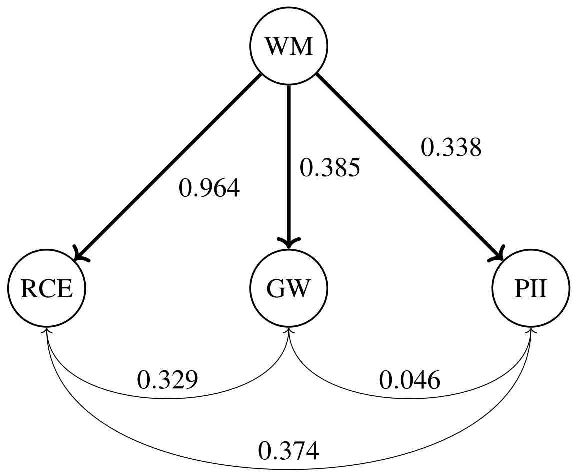

Fig. 1. Path diagram of two-levels hierarchy.

where $\mathbf{X}_h$ is the centered data sub-matrix (corresponding to variables identified by $\mathbf{v}_{\cdot h}$). That can be solved by an Alternate Non-Negative Least Squares algorithm, such that $\hat{\mathbf{y}}_h$ is estimated by a step of a normal Alternate Least Squares (ALS) algorithm while the estimations of $\hat{\mathbf{u}}_{1h}$ consists. thus given $\hat{\mathbf{u}}_{1h}$, $\hat{\mathbf{y}}_h$ is computed by

$$\hat{\mathbf{y}}_h = \mathbf{X}_h \hat{\Psi}_{\mathbf{x}h}^{-\frac{1}{2}} \hat{\mathbf{u}}_{1h} (\hat{\mathbf{u}}_{1h}' \hat{\mathbf{u}}_{1h})^{-1} \tag{14}$$

and given $\mathbf{y}_h$, $\mathbf{u}_{1h}$ is computed by

$$\hat{\mathbf{u}}_{1h} = \mathbf{X}_{h+} \hat{\Psi}_{\mathbf{x}h}^{-\frac{1}{2}} \hat{\mathbf{y}}_h (\hat{\mathbf{y}}_h' \hat{\mathbf{y}}_h)^{-1} \tag{15}$$

where $\mathbf{X}_{h+}$ is the set of passive variables. Thus, the non-negative solution of (13) with respect to $\mathbf{u}_{1h}$ will simply be the unconstrained least squares solution using only the variables corresponding to the passive set, setting the regression coefficients of the active set to zero.

5. Results

5.1. A composite indicator for waste management

The Hierarchical Disjoint Non-Negative Factor Analysis has been applied on the data-set (Section 3) composed by 10 manifest variables for the 28 EU countries. Some variables are taken into consideration during the analysis in order to enrich the information about countries and their performance in waste management (e.g., population, density of population, GDP per capita, etc.). Population has resulted as being the most appropriate element to normalize some manifest variables (see Table 1).

In order to formalize and analyze, in a general framework, the waste management indicator, we propose a hierarchically aggregated index that best represents the waste management in EU, via the statistical identification of reliable and uni-dimensional Specific CIs (SCIs).

The SCIs, which represent dimensions, measure specific concepts contributing into the definition of waste management. Furthermore, the general CI reconstructs the manifest variables via a set of composite indicators according to reflective relations. Few missing data were present in the studied data-set, they were MCAR (Missing Completely at Random) and they have been imputed by the K-nearest neighbors method by setting $K = 10$ and by using the euclidean distance.

In order to set the right polarity of the manifest variables, the *HDNFA* model has been applied many times by following a recursive strategy with increasing number of dimensions. The aim of this analysis is to get a CI for waste management in Europe through H reliable dimensions; and the best model (i.e., the model with the optimal number of dimensions) has been selected via the evaluation of the Bayesian Information Criterion (*BIC*, [18]). The optimal model is the model with 3 dimensions and the related factors, thus the one associated at the lowest value of *BIC* equal to 362.866. The partition and the loadings are reported in Table 3. It is worth observing that all the variables result statistical significant into the model, all the factors are reliable (Cronbach's α [19] higher than 0.70) and uni-dimensional (the variance of the second component of each subset is lower than 1).

Table 3
Results of the optimal model for defining dimensions of waste management.

Variables	RCE	GW	PII	Std error	$Pr(p > \|Z\|)$
1 — GenWas	0	0.682	0	0.138	0.000
2 — GenRec	0	0.876	0	0.091	0.000
3 — PriInv	0	0	0.841	0.102	0.000
4 — JobCir	0	0.461	0	0.168	0.014
5 — PatRec	0	0	0.841	0.102	0.000
6 — MatRec	0.640	0	0	0.145	0.000
7 — ImpnEU	0.401	0	0	0.172	0.030
8 — ImpEU	0	0.417	0	0.172	0.027
9 — RecMun	0.869	0	0	0.093	0.000
10 — RecWas	0.770	0	0	0.120	0.000
Uni-dimensionality	Yes	Yes	Yes		
Cronbach's α	0.751	0.705	0.828		
CI	0.964	0.385	0.338		

Table 4
Spearman's correlations among SCIs and with WM.

	RCE	GW	PII
RCE	1	0.43	0.48
GW	0.43	1	0.32
PII	0.48	0.32	1
WM	0.95	0.64	0.56

As reported in Table 3, the first factor is characterized by the presence of four variables and the most important one is “Recycling rate of municipal waste”, this factor is named “Recycling and Circular Economy - RCE”; the second factor is defined by other four variables as well and the most important is “Generation of recyclable waste”, the second factor is called “Generation of Waste - GW”; whereas the third factor, denominated “Private Investments and Innovation - PII”, is characterized by the presence of two variables: “Private investments as value added at factor cost related to circular economy” and “Patents related to recycling and secondary raw materials” with equal importance in term of loadings.

The factor RCE is the most important in the construction of the general composite indicator “Waste Management - WM”, with loading equal to 0.964, whereas the others two have similar loadings (0.385 for GW and 0.338 for PII). Thus, our model underlines that contribution of the countries' performance in recycling activities is crucial in the definition of their “Waste Management” performance. It is possible to observe the reflective hierarchical structure of the WM CI in Fig. 1.

An interesting result to underline is the low correlation between GW and PII, it seems that the countries that produce most waste are also the ones which less invest privately into circular economy. We can explain it by observing that the correlation between the manifest variables “Private investments as value added at factor cost related to circular economy” and “Generation of recyclable waste” is equal to 0.023. It is worth observing that the correlations between RCE and GW and between RCE and PII are significantly different from 0 because both GW and PII are explained by variables that are positively correlated with “Recycling rate of municipal waste” and “Recycling rate of all waste”.

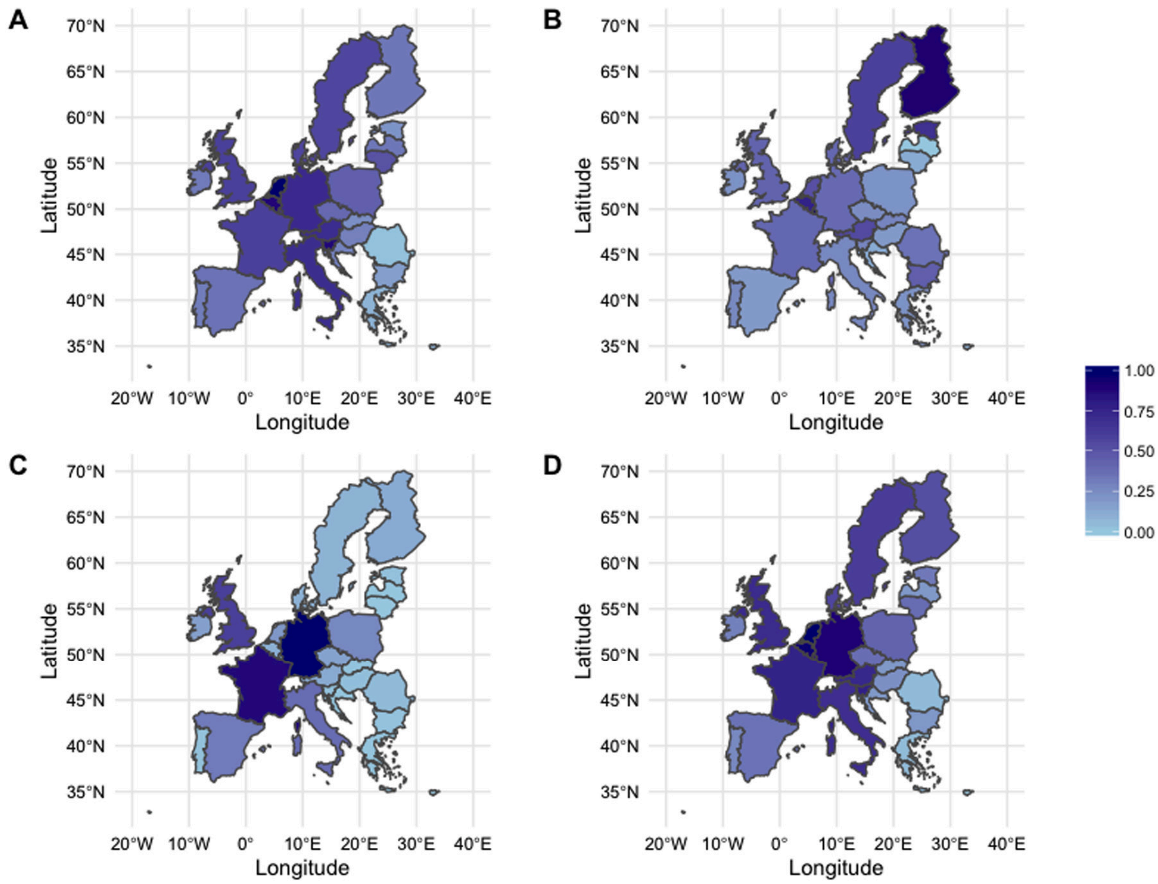

Fig. 2. Normalized scores of Eu Countries: (**A**) RCE, (**B**) GW, (**C**) PII and (**D**) WM.

5.2. Rankings and specific composite indicators

For each SCI is possible to rank the 28 EU countries and finally it is possible to observe the final ranking based on the CI "Waste Management", Table 5. The Netherlands results to be the best country in terms of recycling performances, it is reflected on RCE's ranking and on the final CI's ranking (see Fig. 2). Since the Spearman's correlations among SCIs are not high (as reported in Table 4), the behavior of countries is not equal for each aspect of WM. If we take into account only the group of the best countries (i.e., reported with (1) in Table 5) of each ranking, no country is present in all of them. Germany is present in the group of the best countries for RCE, PII and for the WM, but it is present in the group of intermediate countries (i.e., reported with (2) in Table 5) for GW. All the others, which are present in the group of the best countries for WM, are present in the group of the best countries only for one SCI. By taking into consideration the group of the worst countries (i.e., reported with (3) in Table 5) for WM, we can observe that seven countries (Portugal, Hungary, Slovakia, Croatia, Greece, Malta and Cyprus) are present in the group of the worst countries for all SCIs whereas only three (Ireland, Bulgaria and Romania) are present in the group of the worst countries for two rankings out of three.

Therefore, we can observe that the behavior of the best countries is quite different according to the three main aspects of WM whereas it is more stable for the countries which do not perform well in terms of WM. Our paper evaluates the performances of each country in terms of WM and of its main features. This study provides a crucial information in order to point policies and activities: in which aspects each country has to increase its performance trying to reach the better ones. For instance, The Netherlands has to improve in "Private investments as value added at factor cost related to circular economy" if it wants to reinforce its position as leader in WM.

6. Conclusion

In this paper, we propose a hierarchically aggregated CI for the multidimensional construct Waste Management, by detecting three important SCIs which represent dimensions, in order to provide a useful tool for policy-makers and institutions. The GCI is a general latent construct that best reconstruct the manifest variables. This approach has several advantages because the CI computed respects important properties: scale invariant, unidimensional, reliable and non-compensable. This model-based approach limit the choices of the researcher which often are not based on verified theory.

The three SCIs which characterize WM are: RCE, GW and PII. The most important is RCE, GW and PII contribute less and with almost equal importance. The GCI is important in order to compare the countries' behaviors in terms of WM. Moreover it is possible to evaluate countries' performances for each specific aspect in order to understand why some countries perform better than others and in which way it is possible to improve their performances.

In conclusion, this study provides a useful tool to measure the "goodness" of WM in EU with its main aspects by identifying the most important relationships among manifest variables. The goal is provide a support for EU countries' actions and policies.

Table 5
Rankings based on SCIs and GCI according to these thresholds, (1): normalized score > 0.60, (2): normalized score > 0.30 and < 0.60, (3): normalized score < 0.30.

RCE	GW	PII	WM
NL(1)	LU(1)	DE(1)	NL(1)
BE(1)	FI(1)	FR(1)	BE(1)
SI(1)	BE(1)	UK(1)	DE(1)
DE(1)	EE(1)	IT(2)	LU(1)
AT(1)	SE(1)	ES(2)	SI(1)
IT(1)	AT(2)	MT(3)	FR(1)
LU(2)	NL(2)	PL(3)	AT(1)
UK(2)	BL(2)	NL(3)	UK(1)
FR(2)	UK(2)	IE(3)	IT(1)
DK(2)	DK(2)	CZ(3)	SE(2)
SE(2)	FR(2)	AT(3)	DK(2)
LT(2)	DE(2)	FI(3)	FI(2)
PL(2)	RO(2)	BE(3)	PL(2)
CZ(2)	SI(3)	DK(3)	LT(2)
ES(2)	SK(3)	SE(3)	CZ(2)
LV(2)	IT(3)	LU(3)	ES(2)
FI(2)	PT(3)	RO(3)	EE(2)
IE(2)	CZ(3)	HU(3)	IE(3)
PT(3)	IE(3)	EE(3)	PT(3)
HU(3)	GR(3)	SI(3)	HU(3)
HR(3)	PL(3)	PT(3)	BL(3)
SK(3)	MT(3)	BL(3)	SK(3)
EE(3)	ES(3)	SK(3)	LV(3)
BL(3)	HU(3)	GR(3)	HR(3)
CY(3)	HR(3)	HR(3)	RO(3)
GR(3)	LT(3)	LT(3)	GR(3)
RO(3)	CY(3)	LV(3)	MT(3)
MT(3)	LV(3)	CY(3)	CY(3)

Acknowledgments

The authors acknowledge the support from Noesi Evolution S.r.l. that funded the research.

References

[1] Gutberlet J. Waste, poverty and recycling. Waste Manag 2010;30:171–3.
[2] Transforming our world: the 2030 agenda for sustainable development. 2015, United Nations – Sustainable Development Knowledge Platform.
[3] OECD. The OECD-JRC handbook on practices for developing composite indicators, paper presented at the OECD committee on statistics. 2004.
[4] OECD-JRC. Handbook on constructing composite indicators. methodology and user guide. OECD Publisher; 2008.
[5] Mazziotta M, Pareto A. Methods for constructing composite indices: One for all or all for one? Riv Ital Econ Demogr Stat 2013;67(2):67–80.
[6] Nardo M, Saisana M, Saltelli A, Tarantola S. Tools for composite indicators building. Report EUR 21682, Ispra, Italy: European Commission (Join Research Centre); 2005.
[7] Anderson T, Rubin H. Statistical inferences in factor analysis, in: Proceedings of the third symposium on mathematical statistics and probability, vol. 5, 1956, p. 111–50.
[8] Horst P. Factor analysis of data matrices. 1965.
[9] Pearson K. On lines and planes of closest fit to systems of points in space. Phil Mag 1901;2(11):559–72.
[10] Hotelling H. Analysis of a complex of statistical variables into principal components. J Educ Psychol 1933;24(6). 417–441 and 498–520.
[11] Joreskog K. A general method for analysis of covariance structure. Biometrika 1970;57:239–51.
[12] Bollen K. Structural equations with latent variables. 1989.
[13] Kaplan D. Structural equation modeling: foundations and extensions. 2000.
[14] Joreskog K. A general approach to confirmatory maximum-likelihood factor analysis. Psychometrika 1969;34(2):183–202.
[15] Joreskog K. Structural analysis of covariance and correlation matrices. Psychometrika 1978;43(4):443–77.
[16] Joreskog K. A general approach to confirmatory maximum likelihood factor analysis with addendum. In: Jöreskog K, Sörbom D, editors. Advances in factor analysis and structural equation models. 1979.
[17] Vichi M. Disjoint factor analysis with cross-loadings. Adv Data Anal Classif 2017;11(3):563–91.
[18] Schwarz G. Estimating the dimension of a model. Ann Statist 1978;6(2):461–4.
[19] Cronbach L. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16(3):297–334.

Dr. Carlo Cavicchia obtained his Ph.D. in Statistical Sciences in 2020 at "La Sapienza – University of Rome" and he is currently Research Assistant in Statistic at "Unitelma Sapienza – University of Rome". His research interests are about composite indicators, dimensional reduction and clustering. He won Research Grant for Junior Researcher at "La Sapienza – University of Rome" during his first Ph.D. year. He has presented his works in many conferences.

Professor Pasquale Sarnacchiaro is Associate Professor in Statistics at the University of Rome Unitelma Sapienza, where he teaches a basic course of Statistics for undergraduate students and an advanced course on multidimensional methods for graduate students. Since 2009, he is secretary of the group "Statistics for the evaluation and the quality of services" a standing group of the Italian Statistical Society. He is Guest Editor of Social Indicator Research for the special issue "SocioEconomic indicators for performance evaluation and quality assessment. Statistical Methods and Applications". He is author of many papers published on International Journals of Statistics and applied Statistics (Journal of applied Statistics, Food Quality and Preference, Journal of Mental Health, Environment International, Australian and New Zealand Journal of Statistics, Corporate Social Responsibility and Environmental Management).

Professor Maurizio Vichi is full professor of statistics and director of the Statistic Sciences Department at "La Sapienza – University of Rome". He is Vice-President and past-President of FENStatS. He is Deputy-Chair of European Statistics Advisory Committee (ESAC), Member of Board of the European Master in Official Statistics and Elected Member of the International Statistical Institute. He has been President of Italian Statistic Society and President of International Federation of Classification Societies. He is editor of journal "Advances in Data Analysis and Classification" and of series "Studies in Theoretical and Applied Statistics" and "Studies in Classification, Data Analysis and Knowledge Organization".