Book of Short Papers SIS 2020

Editors: Alessio Pollice, Nicola Salvati and Francesco Schirripa Spagnolo

Published by Pearson

www.pearson.com

ISBN 9788891910776

Contents

Specialized sessions

Solicited Sessions

Contributed papers and Posters

Modeling Well-Being through *PLS-SEM and K-M*
Misurare il benessere con il modello PLS-SEM e K-Means

Venera Tomaselli[1], Mario Fordellone[2], Maurizio Vichi[3]

Abstract The concept of "fair and sustainable" well-being (BES) is employed to estimate the effects of 2008-2016 economic crisis on voting behaviour. Many studies suggest overcoming Gross Domestic Product as the only indicator to measure economic growth and development. So, in addition to strictly economic indicators, social welfare and performance of institutions, and public services indicators are used. With the aim to build a composite indicator of well-being, based on BES elementary indicators and domains, *Partial Least Squares-Structural Equation Modeling* and *K-means* clustering simultaneous method is employed to model the well-being classifying territorial micro-areas.

Abstract *Il benessere "equo e sostenibile" (BES) è impiegato per stimare gli effetti della crisi economica 2008-2016 sul comportamento di voto. Numerosi studi suggeriscono il superamento del prodotto interno lordo come unico indicatore per misurare la crescita e lo sviluppo economico. Pertanto, oltre ad indicatori strettamente economici, l'analisi tiene conto degli indicatori BES relativi al benessere sociale ed alle prestazioni delle istituzioni e dei servizi pubblici. Con l'obiettivo di costruire un indicatore composito di benessere, basato su indicatori elementari e domini del BES, il metodo simultaneo Partial Least Squares-Structural Equation Modeling e K-means clustering è utilizzato per l'analisi del benessere e la classificazione di micro aree territoriali.*

Key words: composite indicators, *Partial Least Squares-Structural Equation Models and K-means* method, well-being, territorial micro-areas

[1] Venera Tomaselli, Department of Political and Social Sciences, University of Catania IT; e-mail: venera.tomaselli@unict.it (*corresponding author*).
[2] Mario Fordellone, Department of Statistical Sciences, University "La Sapienza", Rome IT: e-mail: mario.fordellone@uniroma1.it.
[3] Maurizio Vichi, Department of Statistical Sciences, University "La Sapienza", Rome IT: e-mail: maurizio.vichi@uniroma1.it.

1 Introduction

The concept of well-being is very complex and takes into account economic resources, public policies, quality of the environment, and many other topics. Well-being indicators, indeed, are proxies of unobservable, latent dimensions.

By questioning the use of Gross Domestic Product (GDP), the study aims to measure well-being through a multidimensional approach on the basis of the relationships among elementary indicators and domains proposed in the analysis of well-being through BES in order to define the territorial differences of well-being more reliably.

The indicators have been selected from 12 BES domains, grouped in turn into 6 topics: economy and labour market, education, environment, institutional performance, safety and quality of services. The crucial issues are:

- is the economic well-being level the only useful and meaningful measurement to explain the variability of territorial micro-areas?
- are indicators closely related to economic well-being and indicators of well-being related, for instance, to the quality of services and opportunities available in the territory, do not necessarily are correlated?
- is the well-being a composite concept based on a scaling of specific priorities?

In the following section 2, the well-being measurement is discussed; in section 3 the PLS-SEM and K-Means method is presented and in sections 4 the results obtained simultaneously through well-being composite indicator and classification method for the territorial micro-areas are delivered.

2 Well-being measurement

The measures of well-being, economic progress, and social welfare are adopted as drivers for designing public policies by decision makers (Jayawickreme et al., 2012; Layar, 2011; Sachs, 2012). The measures more accurately depict changes not only in individual living standards (Helliwell et al., 2012) but simultaneously also in comprehensive national economic growth (Diener et al., 2009).

Since around 2000 the Organisation for Economic Cooperation and Development (OECD) embarked on a global project to measure the well-being and progress of societies not just and not only through the economic performances. The project was involved in setting up and supporting the Commission on the Measurement of Economic Performance and Social Progress (CMEPSP), established by the President of France, Nicolas Sarkozy in 2008 and led by Stiglitz, Sen and Fitoussi.

The limits of GDP are reviewed in the report of the Commission because the GDP is not believed as a standard of the well-being of societies. The GDP, indeed, does not address economic inequality, happiness, quality of life, wellness, and other crucial societal parameters, and does not integrate environmental services into economic decisions (Stiglitz et al., 2010).

More recently, other scholars (Ven, 2015; Fleurbaey, 2015) have called for a new generation of multifaceted and more comprehensive well-being measures, better able to describe actual living standards and useful for a more accurate design of policies improving efficiency in resources assignment.

A series of measures of well-being, inspired at the Nussbaum-Sen approach to human capabilities and subjective well-being (Nussbaum and Sen, 1993), have been proposed in attempt to go beyond GDP with the aim to broaden the scope of effects in the assessment of policies. For instance, the Human Development Index by UNDP or the Better Life Initiative launched by OECD (2015) and many other approaches are based on the income, health, and education measurements of the countries' performance (for a review, see Fleurbaey and Blanchet, 2013).

From a technical view, many methods are employed to measure the well-being level through composite indicators. Nevertheless, no method is universally valid to select indicators based on theory-driven criteria, measure properly the concept, aggregate and normalise a set of input variables and define a weighting system (OECD, 2008).

The aim is to simplify the analysis of the multidimensional concept according with a formative or reflective measurement model, where elementary indicators are causes or effects of latent variables, respectively (Michalos, 2014; Simonetto, 2012).

Only some studies have focused on the construction of well-being composite indexes to evaluate and compare well-being specifically across the Italian provinces. Mazziotta and Pareto (2019) have obtained a global well-being index by aggregating 11 composite indices with AMPI (Mazziotta and Pareto, 2016) and have ranked Italian provinces for each dimension of well-being and have given an overall ranking.

Also Calcagnini and Perugini (2019) have proposed a composite indicator of well-being for the Italian provinces (NUTS-3) based on the methodology of the regional Index of Regional Quality of Development (QUARS) to analyse the extent to which the socio-economic heterogeneity in individual and contextual features within region affects the well-being among adjacent provinces.

In Italy, indicators of well-being have being used more and more for policy-making reasons at national, but also regional or local levels involving public institutions. From a theoretical point of view, the relationship of well-being assessment with policy-making process in healthcare, education, and training, or local services is the rationale to analyse the well-being measures at local level.

Since the policies of local governmental authorities have a direct and huge impact on the socio-economic context, the assessment of living standards at provincial level allows to evaluate economic, environmental, and social needs of the citizens at any level of government in order to implement and design decentralised policies. Following this, for territorial micro-areas the economic dimension could not be so crucial.

3 PLS–SEM-KM for composite indicator building

Partial Least Squares (PLS) methodologies are algorithmic tools with analytic proprieties aiming at solving problems about the stringent assumptions on data, e.g., distributional assumptions that are hard to observe in real life.

Tenenhaus et al. (2005) try to better clarify the terminology used in the PLS field through an interesting review of the literature, focusing the attention on the Structural Equation Models standpoint.

Given the $n \times J$ data matrix $\mathbf{X}$, the $n \times K$ membership matrix $\mathbf{U}$, the $K \times J$ centroids matrix $\mathbf{C}$, the $J \times P$ loadings matrix $\boldsymbol{\Lambda} = [\boldsymbol{\Lambda}_{\mathrm{H}}, \boldsymbol{\Lambda}_{\mathrm{L}}]$, and the errors matrices $\mathbf{Z}$, $\mathbf{E}$ and $\mathbf{D}$, the PLS-SEM-*K*M model simultaneously identify a SEM model for variables and a partitioning *K*M model for units according to the following three equations:

$$\begin{aligned} \mathbf{H} &= \mathbf{H}\mathbf{B}^{\mathrm{T}} + \boldsymbol{\Xi}\boldsymbol{\Gamma}^{\mathrm{T}} + \mathbf{Z}, \\ \mathbf{X} &= \mathbf{Y}\boldsymbol{\Lambda}^{\mathrm{T}} + \mathbf{E} = \boldsymbol{\Xi}\boldsymbol{\Lambda}_{\mathrm{H}}^{\mathrm{T}} + \mathbf{H}\boldsymbol{\Lambda}_{\mathrm{L}}^{\mathrm{T}} + \mathbf{E}, \\ \mathbf{X} &= \mathbf{U}\mathbf{C}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\mathrm{T}} = \mathbf{U}\mathbf{C}\boldsymbol{\Lambda}_{\mathrm{H}}\boldsymbol{\Lambda}_{\mathrm{H}}^{\mathrm{T}} + \mathbf{U}\mathbf{C}\boldsymbol{\Lambda}_{\mathrm{L}}\boldsymbol{\Lambda}_{\mathrm{L}}^{\mathrm{T}} + \mathbf{D} \end{aligned} \tag{1}$$

subject to constraints: *(i)* $\boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\Lambda} = \mathbf{I}$; and *(ii)* $\mathbf{U} \in \{0,1\}$, $\mathbf{U}\mathbf{1}_K = \mathbf{1}_n$. Thus, the PLS-SEM-KM model includes the PLS-PM and the clustering *K*M equations (i.e., $\mathbf{X} = \mathbf{U}\mathbf{C}$ and then, $\mathbf{Y} = \mathbf{X}\boldsymbol{\Lambda}$ becomes $\mathbf{Y} = \mathbf{U}\mathbf{C}\boldsymbol{\Lambda}$).

The proposed methodology shows some important advantages with respect to the other proposed approaches for both cluster analysis and composite indicator construction: it is a new simultaneous approach.

In fact, identifies the best homogenous partition of the objects, represented by the best statistical relationships among latent and observed variables (Fordellone and Vichi, 2018; 2020).

4 Results: well-being composite indicator

The conceptual structure (Giovannini et al., 2012) of the BES considers 9 domains related to aspects that directly influence well-being, plus 3 instrumental or context domains.

In each well-being domain (ISTAT, 2018), elementary indicators and a synthesis through composite indicators related to each domain are integrated.

Our dataset consists in 109 units (Italian provinces) and 16 indicators organized in 9 different domains. Table 1 shows the relationships among MVs and LVs and the loadings estimated through PLS-SEM-KM.

Through the application of PLS-SEM-KM model, we have identified 3 homogeneous groups of Italian provinces. The optimal number of clusters identified corresponds to the maximum value of the pseudo-*F* function (around 1.1).

From this analysis, we can see that the theoretical polarity associated to each observed variable (see Table 1) is well described by the measurement PLS approach.

Whereas only the *Overcrowding of prisons* variable shows a non-significant loading, in fact also the sign is not correct.

Table 1: Measurement models matrix estimated through PLS-SEM-KM.

MVs	LV1	LV2	LV3	LV4	LV5	LV6	LV7	LV8	LV9	WB
MV1	0.795	0	0	0	0	0	0	0	0	0.721
MV2	0	0.750	0	0	0	0	0	0	0	0.619
MV3	0	-0.678	0	0	0	0	0	0	0	-0.917
MV4	0	0	-0.876	0	0	0	0	0	0	-0.938
MV5	0	0	-0.887	0	0	0	0	0	0	-0.881
MV6	0	0	0.920	0	0	0	0	0	0	0.870
MV7	0	0	-0.917	0	0	0	0	0	0	-0.914
MV8	0	0	0	-0.669	0	0	0	0	0	-0.569
MV9	0	0	0	0	0.050	0	0	0	0	0.202
MV10	0	0	0	0	0	-0.373	0	0	0	-0.417
MV11	0	0	0	0	0	-0.216	0	0	0	-0.140
MV12	0	0	0	0	0	-0.434	0	0	0	-0.412
MV13	0	0	0	0	0	0	0.105	0	0	0.226
MV14	0	0	0	0	0	0	0	0.277	0	-0.088
MV15	0	0	0	0	0	0	0	0.691	0	0.704
MV16	0	0	0	0	0	0	0	0	0.748	0.822

In Table 2 the path coefficients' matrix of the estimated structural-PLS model is shown.

Table 2: Structural model matrix estimated by PLS-SEM-KM.

	LV1	LV2	LV3	LV4	LV5	LV6	LV7	LV8	LV9	WB
LV1-Health	0	0	0	0	0	0	0	0	0	0.324
LV2-Education and training	0	0	0	0	0	0	0	0	0	0.053
LV3-Work-life balance	0	0	0	0	0	0	0	0	0	0.666
LV4-Economic well-being	0	0	0	0	0	0	0	0	0	0.247
LV5-Policy	0	0	0	0	0	0	0	0	0	0.117
LV6-Safety	0	0	0	0	0	0	0	0	0	0.163
LV7-Cultural heritage	0	0	0	0	0	0	0	0	0	0.067
LV8-Environment	0	0	0	0	0	0	0	0	0	0.192
LV9-Innnovation and research	0	0	0	0	0	0	0	0	0	0.198
Well-being composite indicator	0	0	0	0	0	0	0	0	0	0

From the structural model we can see that the Italian provincial well-being is highly affected by the LV3-*Work-life balance* construct (0.666), followed by LV1-*Health* (0.324) and LV4-*Economic well-being* (0.247) constructs. Whereas, the LV6-*Cultural heritage* (0.067) and LV2-*Education* (0.053) constructs are

dimensions with non-significant coefficients. The overall fit of the structural model is good with a $R^2 = 0.74$. In terms of clustering results, the three clusters identified by the PLS-SEM-KM algorithm describe three groups of well-being, i.e., *high*, *medium*, and *low* levels. In this way is very easy to classify the Italian provinces through a "BES ranking".

5 Conclusions

The BES 2018 report by ISTAT has confirmed that in 2015 and in 2016 an improvement in many areas of well-being has been observed, even if territorial differences remain stable both in levels and dynamics.

The differences appear in some cases as real structural differences between North and South of Italy. Then, the different well-being levels measured through a composite well-being indicator in the territories allow to explain the variability in territorial micro areas.

References

1. Diener, E., Lucas, R., Schimmack, U., & Helliwell, J. (2009). Well-Being for public policy. Oxford, UK: Oxford University Press. doi.org/10.1093/acprof:oso/9780195334074.001.0001.
2. Fleurbaey, M. (2015). Beyond income and wealth. Rev. Income Wealth, 61(2), 199-219.
3. Fleurbaey, M., & Blanchet, D. (2013). Beyond GDP: Measuring welfare and assessing sustainability. Oxford University Press.
4. Fordellone, M., & Vichi, M. (2018). Structural Equation Modeling and simultaneous clustering through the Partial Least Squares algorithm. arXiv preprint: arXiv:1810.07677v1.
5. Fordellone, M., & Vichi, M. (2020). Finding groups in structural equation modeling through the partial least squares algorithm. Comput Stat Data An doi: https://doi.org/10.1016/j.csda.2020.106957
6. Giovannini, E., Morrone, A., Rondinella, T., & Sabbadini, L. L. (2012). L'iniziativa CNEL-ISTAT per la misurazione del Benessere Equo e Sostenibile in Italia. Autonomie locali e servizi sociali, 1, 125-136.
7. Helliwell, J. F., Layard, R., & Sachs, J. (2012). Some policy implications. In J. F. Helliwell, R. Layard & J. Sachs (Eds.) World Happiness Report (pp. 90-107), New York, NY: The Earth Institute, Columbia University
8. ISTAT (2018). BES 2018: Il benessere equo e sostenibile in Italia. Istituto Nazionale di Statistica.
9. Jayawickreme, E., Forgeard, M. J., & Seligman, M. E. (2012). The engine of well-being. Rev Gen Psychol, 16(4), 327-342. doi.org/10.1037/a0027990.
10. Mazziotta, M., & Pareto, A. (2016). On a generalized non-compensatory composite index for measuring socio-economic phenomena. Soc Indic Res, 127(3), 983-1003.
11. Mazziotta, M., & Pareto, A. (2019). Use and misuse of PCA for measuring well-being. Soc Indic Res, 142(2), 451-476.
12. Nussbaum, M. & Sen, A. (1993). The Quality of Life. Oxford: Oxford University Press.
13. OECD (2015). How's life? 2015 Measuring well-being. Paris, FR: OECD Publishing. (Available online: http://dx.doi.org/10.1787/9789264121164-en).
14. Stiglitz, J. E., Sen, A., & Fitoussi, J. P. (2010). Mismeasuring our lives: why GDP doesn't add up. New York, NJ: The New Press.
15. Tenenhaus, M., Esposito Vinzi, V., Chatelin, Y. M., & Lauro, N. C. (2005). PLS path modelling. Comput Stat Data An, 48(1), 159-205.
16. Ven, P. (2015). Introduction to the symposium on new measures of well-being: Perspectives from statistical offices. Rev Income Wealth, 61(1), 1-3.