Electronic Research Archive

https://www.aimspress.com/journal/era

ERA, 32(8): 4772–4795.
DOI: 10.3934/era.2024218
Received: 09 June 2024
Revised: 14 July 2024
Accepted: 01 August 2024
Published: 05 August 2024

Research article

An application of structural equation modeling with partial least squares to analyze life satisfaction by degrees of urbanization

Jose A. Sánchez-Martí[1,*] and Miguel A. Esteban-Yago[2]

[1] Department of Applied Economics III, University of Sevilla, Sevilla 41018, Spain
[2] Department of Applied Economics, University of Murcia, Murcia 30100, Spain

* **Correspondence:** Email: jsanchez24@us.es.

Abstract: The objective conditions of people's lives have historically played a greater role than subjective perceptions. However, this trend is shifting, accompanied by an ongoing debate about life satisfaction and urbanization. This research has two objectives, the first is to test the hypothesis of the urban paradox. The second is to analyze the differences in the determinants of life satisfaction by urbanization levels. In order to determine the degrees of urbanization, a geographic grid was used to distinguish between rural, intermediate, and urban areas. In addition, a multi-group analysis in structural equation models with partial least squares was conducted to analyze the determinants of life satisfaction. The model has been developed using three general determinants: economic situation, health, and social relations, which has been applied to a case study of the Region of Murcia. The database used for this study includes 2462 samples. The population living in rural areas showed less life satisfaction. Regarding the determinants, it was found that health acts as a partial mediator of social relations. Differences were also observed in the intensity of the determinants according to the degree of urbanization. Thus, the degree of urbanization acts as a moderator of life satisfaction. This research is noteworthy because it provides valuable insights for guiding regional public policies.

Keywords: life satisfaction; degree of urbanization; rural happiness paradox; urban paradox; PLS-SEM; geographic grid; moderating effect

1. Introduction

Since ancient times, there has been a broad consensus that the fundamental objective of the public policies of any state should be the welfare of its citizens. Nevertheless, a greater focus has been placed on the objective circumstances of well-being than on individuals' subjective perceptions of their own lives. This situation appears to have reversed since the latter third of the 20th century.

One of the most extensively researched topics has been the search for the underlying factors or determinants of life satisfaction (LS), subjective well-being (SWB), or happiness. The pioneering study by Diener [1], from the field of psychology, synthesizes the preceding research in relation to theoretical approaches, measures, and determinants of SWB. Subsequent studies, including those by Kahneman and Krueger [2], Diener et al. [3], and Veenhoven [4], among many others, have contributed to the advancement of this field.

Most of the analyses indicate a direct relationship between a person's economic level or financial situation and LS or happiness, both in comparisons between different countries and in those within the same country. However, Easterlin [5] highlighted that at a certain level of income, happiness does not grow at the same rate, and even stops growing. This phenomenon is known as "Easterlin's paradox", and it is susceptible to diverse methodological approaches, including those employing a temporal, cross-sectional, or life-cycle perspective [6]. In addition to this paradox, other works have examined the relationship between age and LS, showing that it adopts a U-shape [7, 8]. Moreover, there is a large body of research that has focused on other characteristics such as marital status, health status and unemployment, among others [9–11].

The relationship between urbanization and LS is currently the subject of an ongoing debate. Given that income levels are typically higher in urban areas than in rural areas, one would expect to find higher levels of LS among the urban population. However, after analyzing many countries, Easterlin et al. [12] observed that as countries become more advanced, the SWB gap (encompassing the concepts and measures of both happiness and LS) between the two areas dissipates and even reverses. This phenomenon has been described as the 'urban paradox' [13] or the "rural happiness paradox" [14]. Nevertheless, some studies have shown that there are no statistically significant differences in LS between rural and urban populations [15, 16]. In addition, Lenzi and Perucca [17] found that the population living in the intermediate areas of urbanization were the most satisfied. Research that goes beyond a descriptive study and explores the causes of these differences is scarce and does not provide a complete explanation of the paradox [14, 18–21].

This research is situated within this context, although it differs from most of the existing literature in that it is applied at the regional level, specifically in the Region of Murcia (Spain). This region is characterized by a per capita income below the national average and a risk of poverty and social exclusion levels that exceed the national average [22, 23]. Moreover, the rural population of the Region of Murcia has a higher risk of poverty and social exclusion than the urban population, as well as worse living conditions [24, 25].

The objective of this research is twofold. The first objective is to analyze whether the urban paradox is present in the Region of Murcia. Unlike the existing literature, we have the advantage of georeferenced information to verify its presence. This enables greater precision when classifying each sample's degree of urbanization. The second objective is to analyze the determinants of LS in rural and urban populations. Specifically, three general determinants are considered: economic situation, health, and social relations. These determinants are characterized by a broad set of indicators and variables.

In general, research has employed ordered logit [11, 14] or ordinary least squares (OLS) [26, 27] to analyze the determinants of LS according to the degree of urbanization. The choice between one technique or the other depends on the discipline from which it is approached. Economists typically conceptualize LS from an ordinal perspective, whereas psychologists and sociologists adopt a cardinal stance [28]. In this study, the partial least squares structural equation modeling (PLS-SEM) technique was employed. Although there is a large body of literature that analyzes LS using structural equation models, there is a scarcity of research that focuses on this topic. Additionally, the literature reviewed does not focus on the general population, but on a specific age group. Therefore, this study represents a novel application of this technique in a general population. The purpose is to establish a model that includes the general determinants of LS and to analyze the changes in their relationships according to the degree of urbanization. This is referred to as a multi-group analysis (MGA-PLS). The software used to execute the PLS-SEM technique is the *SEMinR* (2.3.2) library of *R* statistical software [29,30].

The advantage of employing this technique in lieu of traditional methodologies is its capacity to incorporate unobserved variables or constructs into the model. For example, one of the general determinants of LS is social relationships of an individual. However, it is not possible to measure this directly; rather, it can be assessed through certain indicators or observed variables, such as the frequency of calling friends or attending events. Furthermore, it enables the quantification of measurement errors associated with the indicators, which represent the unexplained variance [31].

This article has seven sections. Following the introduction, Section 2 reviews the existing literature on LS and degrees of urbanization. Section 3 develops the hypotheses of the proposed model, and Section 4 describes the data and methodology employed. Section 5 presents the main results of the study, Section 6 discusses the results obtained, and Section 7 synthesizes the main conclusions and their academic and policy implications.

2. Literature background

Economic growth and development transform a rural society into a primarily urban society. When a country is less developed, urban areas have greater potential to generate economic activity. This is due to the location of activities with higher levels of productivity, better salaries, and better access to public services. In essence, urban areas achieve a higher standard of living for their inhabitants. The increased attractiveness of urban areas also causes the population to move from rural to urban areas. In principle, this should lead to an overall increase in SWB [32]. However, the relative advantages of urban areas diminish as development intensifies. The higher cost of urban living can be accompanied by other problems or negative externalities such as pollution, noise, congestion, crime, and lower levels of social capital. At the same time, rural disadvantages decrease with improved communications, infrastructure, and greater accessibility to public services [33].

Morrison [13] analyzed the urban paradox in different countries around the world, comparing the situation of rural areas to large and small cities. This author concluded that in less developed countries, the highest levels of well-being were observed in large cities, followed by smaller cities and, finally, rural areas. Conversely, when per capita income levels were higher, the order was reversed, with large cities exhibiting a decline in well-being.

Most studies have confirmed the urban paradox in European countries and the European Union. Veenhoven [15] and Shucksmith et al. [16] found no significant differences in LS between rural and urban populations. However, they indicated that rural residents in less wealthy countries reported lower LS than the urban population, while in wealthier countries, the opposite was true, with rural dwellers reporting higher LS than urban dwellers. Sørensen [18] reported that rural population had significantly higher LS than urban population when socioeconomic factors were constant. This was consistently observed in three groups of European Union countries defined by their level of Gross

Domestic Product (GDP). Nevertheless, Requena [26] revealed that the rural population in European countries with lower GDP per capita reported lower levels of happiness than their urban counterparts, while the rural population in countries with higher GDP per capita reported the highest levels of happiness.

Research on the influence of urbanization on LS can be classified into two groups. The first uses a subjective variable, such as the perception of the degree of urbanization. The second employs an objective variable, such as population density and population size.

Dijkstra and Papadimitriou [34] used a subjective classification of urbanization in which respondents self-reported their perceptions of urbanization levels on a three-point scale: cities, semi-dense areas, and rural areas. In Western Europe, the rural population reported higher LS than their urban counterparts, although the observed differences were not statistically significant. Brereton et al. [27] investigated individual well-being and rural quality of life in Ireland using self-reported LS levels. The authors discovered a high level of LS in Irish rural areas and identified the main problems perceived in rural areas, such as access to healthcare and public transportation. Weckroth and Kemppainen [35] employed distinct conceptualizations of urban or rural areas derived from two sources: the European Social Survey and the statistical office of the European Union (Eurostat). They concluded that self-reported areas of residence were closely associated with LS. Specifically, the researchers observed that the urban paradox occurred in countries with higher per capita income, while in countries with lower per capita income, there was no significant difference between residents of large cities and rural residents. Nevertheless, regional GDP and other urbanization indicators were not found to be significant.

Regarding the objective measures, Bernini and Tampieri [36] examined the role of urbanization in Italy. These authors delimited the rural population by population size. Their findings indicated that family satisfaction played a more significant role in explaining happiness in urban areas, while satisfaction with health, friendships, and the environment were more influential in rural areas. Lenzi and Perucca [17] analyzed the relationship between urbanization and LS in European regions during the period 2004-2011. The authors established that LS was higher in regions with intermediate levels of urbanization, particularly in Western European countries, and lower in more urbanized regions. Conversely, the authors noted that rural areas exhibited higher LS only when they were located in urbanized regions. This is because these populations can benefit from the positive externalities of the large cities in the region without suffering the potential disadvantages.

To analyze the causes of the rural happiness paradox, Okulicz and Maya [19] included several control variables in the model, which were grouped into socio-economic factors and urban environment factors. The results showed that the negative relationship between happiness and urbanization level remained stable when the categories of control variables were considered, although the socio-economic variables achieved greater statistical significance than the environmental variables. The same results were confirmed by Piper [20] for European capitals. Sørensen [18] pointed out that variables related to social interaction and control over one's life only partially explained the differences in LS between rural and urban populations.

Sørensen [14] tested three causes of the rural happiness paradox: spatial location satisfaction, social capital, and access to nature. In this case, the perspective was different from previous research, as he analyzed these three causes separately for each degree of urbanization. The results showed that only strong solidarity and natural amenities were statistically significant. Additionally, Morrison and

Weckroth [21] focused on human values to explain the paradox. These authors chose a different method of analysis than Sørensen [14], opting to include an interaction term between metropolitan areas and human values in the model. As a result, metropolitan residents felt more identified with individualistic values, while non-metropolitan residents showed a greater affinity with community values.

Studies examining the relationship between urbanization and SWB using structural equation models are scarce. Moreover, these studies do not analyze the general population but focus on specific age groups or populations, such as minors [37], college students [38], or older adults [39, 40]. These studies used a multi-group analysis strategy, comparing the results of their models according to a binary variable representing the degree of urbanization. Schilling and Wahl [40] opted for a different strategy, including an exogenous and binary variable representing the degrees of urbanization in their model.

3. Hypothesis development

A wide variety of determinants have been used in the literature. Layard [10] synthesized up to seven major domains: family relationships, financial situation, employment, community and friends, health, personal freedoms, and personal values. However, this author pointed out that, except for financial situation and health, the rest of the determinants correspond to the quality of social relationships. Additionally, Dolan et al. [9] provided a comprehensive review of the literature on the various determinants used. In essence, this research considered three general determinants: economic situation, health, and social relations. The purpose of this study was to test a series of hypotheses by examining the role of each general determinant and the possible relationships existing among them.

A household's material standard of living is considered a universal determinant of LS [41]. The relationship between income and LS has been extensively studied in the literature, both at the macroeconomic level, with GDP per capita [12, 42], and at the microeconomic level, with household income [18, 43]. Ferrer-i-Carbonell [44] emphasized the importance of relative income, that is, the comparison of a person's income with the income of their reference group. This author established that an increase in an individual's income accompanied by an equal increase for their reference group did not result in a significant change in their SWB. Other authors have incorporated additional variables. Bhuiyan and Szulga [45] suggested that household savings were a determinant of LS. Apart from the objective economic situation, Hayo and Seifert [46] found a high correlation between subjective perceptions of a person's economic situation and their LS.

H1: Economic situation affects LS.

Layard [10] identified family relationships as a crucial element of SWB. The literature analyzing SWB according to degrees of urbanization has primarily focused on two relationship indicators: the number of children in the household and marital status. However, current social relationships are not only face-to-face but also conducted online. In this context, Zhou et al. [47] examined the impact of online social relationships on LS in older adults. Their findings indicated that these online interactions had a positive effect on reducing loneliness and increasing LS. Additionally, Masuda et al. [48] advocated including leisure time in the assessment of LS.

H2: Social relationships affect LS.

The importance of health in SWB has been contrasted in several studies [49, 50]. Peiró [11] evidenced that poor or very poor health had a profound impact on both happiness and LS. Other authors have employed chronic illness or health issues as supplementary indicators of poor health [36, 41].

H3: Health affects LS.

Helliwell and Putnam [8] conducted an analysis of the social context of SWB assessment and concluded that social capital had a positive effect on physical health and SWB. Social relationships can influence SWB directly and indirectly through health. In this context, Yip et al. [51] reported that in rural China, these two relationships were more pronounced when social capital was based on trust and reciprocity with other individuals, as opposed to social capital related to membership in an organization. Furthermore, Wang et al. [52] demonstrated that social capital had a direct positive relationship with LS but also an indirect relationship through mental health.

H4: Social relationships affect health.

Furthermore, the economic situation affects an individual's health. Although this may depend on the development of the welfare state in each region, income may be necessary to pay for medications, treatment, or medical care. Several studies have shown that financial security, or its perception, affects psychological well-being [53,54]. Moreno-Agostino et al. [55] employed structural equation modeling to investigate the determinants of LS among older adults. Their findings indicated that health played a mediating role between a person's economic situation and LS.

H5: Economic situation affects health.

Finally, Sørensen [14] indicated that the relationship between social interactions and natural amenities with LS was conditioned by the degree of urbanization. Furthermore, Morrison and Weckroth [21] demonstrated that human values were a determinant of LS. However, this relationship depended on the degree of urbanization. Therefore, this moderating effect can be generalized to all other determinants of LS, including the relationships between them.

H6: The degree of urbanization is a moderator of LS.

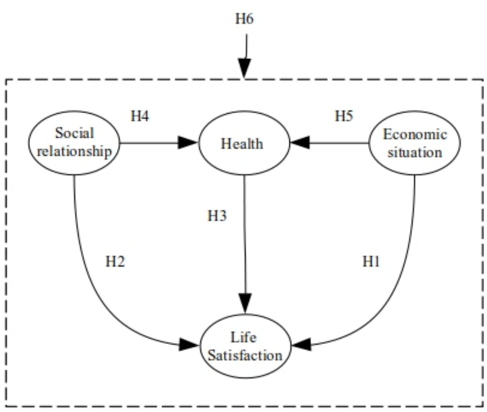

Figure 1. Hypotheses.

Figure 1 depicts the theoretical model with the corresponding hypotheses. The moderating effect is represented by dotted lines, the ovals refer to latent variables or constructs, and the arrows indicate the proposed hypotheses.

4. Materials and methods

4.1. Rural and urban identification

One of the criteria for delimiting rural areas is the size of the population of a spatial unit. However, relying exclusively on this criterion can lead to error since the surface area of the spatial unit is not taken into account. For this reason, it is generally accompanied by the criterion of population density, which is measured as the population per square kilometer (km^2) of a given spatial unit. It is important to note that population density can be influenced by the administrative boundaries of a spatial unit [56]. To mitigate this potential bias, a geographic grid can be employed. This grid comprises regular square cells of 1 km^2. By standardizing population density calculations and eliminating the influence of administrative boundaries, the resulting data can be more accurately interpreted.

This research followed the Eurostat methodology based on geographic grids to calculate the population density of each cell and classify them as urban or rural [57]. Urban areas are defined as a set of contiguous cells with a population of at least 5000 inhabitants and a population density of at least 300 inhabitants per km^2. Two urban categories are then created: intermediate areas and urban areas. Urban areas are characterized by a higher population density and population size than intermediate areas. To qualify as urban, a cell must have a population of at least 1500 inhabitants per km^2 and a population of at least 50,000 inhabitants. The identification of urban areas is based on contiguous cells that meet these criteria. Unlike intermediate areas, diagonals are not considered and gaps are filled in. Finally, rural cells are those that are neither intermediate nor urban areas.

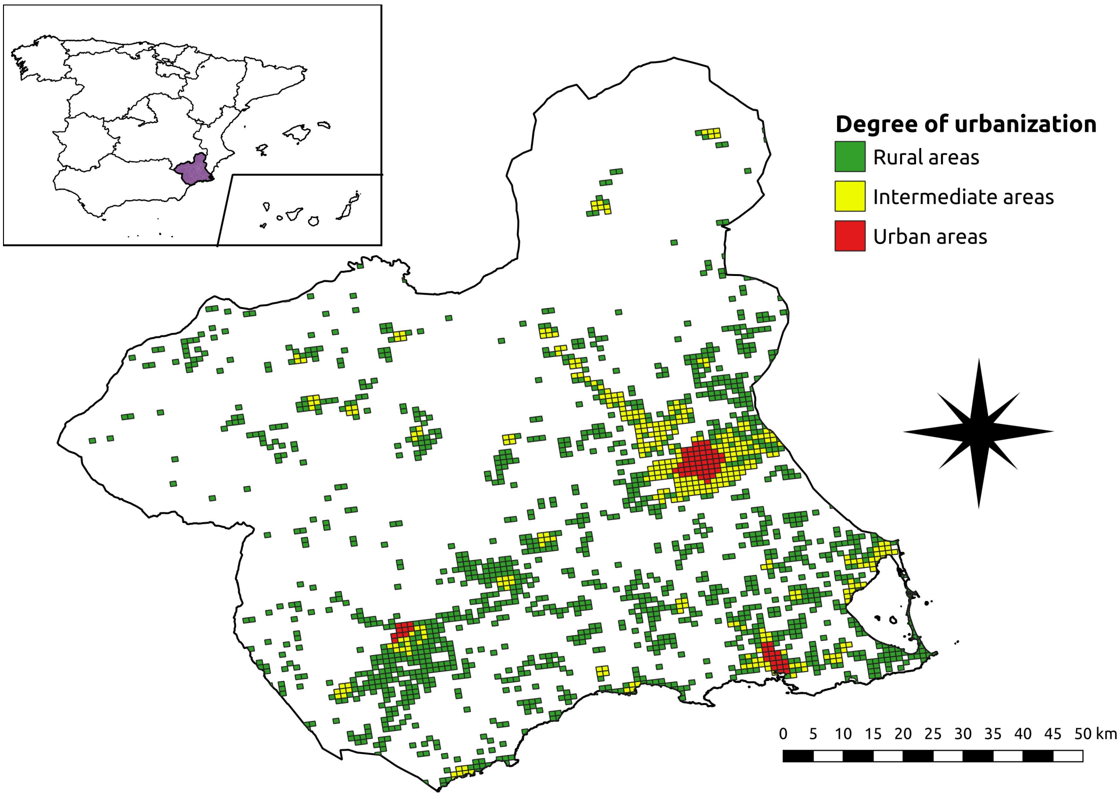

Figure 2. Region of Murcia and degree of urbanization.

Once each cell has been classified into a certain degree of urbanization, these are applied to a specific spatial reference unit. Using Eurostat's methodology, these are municipalities. However, municipalities have the disadvantage of administrative boundaries. In the Region of Murcia, the municipalities are heterogeneous and range greatly in size. To mitigate this potential bias, we selected the spatial unit closest to the population: the population nucleus [58]. Thus, population nuclei that did not intersect an urban cell were classified as rural population nuclei.

In short, rural areas have been identified as rural population nuclei and small rural municipalities, the latter with fewer than 5000 inhabitants. The data used were taken from the 2011 Population and Housing Census. The geographic grid provided by the National Statistics Institute [59] only considers the population residing in private households. Figure 2 shows a map of the degrees of urbanization of the Region of Murcia following the geographic grid methodology. A map of the different autonomous communities of Spain is in the upper left corner, with the Region of Murcia in purple.

4.2. Data

Table 1. Indicators.

Construct	Indicator	Type	Min	Max	Description
Economic situation (ECO)	ECO1	Ordinal[a]	0	3	Sum of the components of the AROPE
	ECO2	Binary	0	1	The household makes ends meet with great difficulty
	ECO3	Binary	0	1	Being unable to face unexpected expenses of 650 euros with a household's own means
	ECO4	Continuous	0	1	The income needed to move out of monetary poverty expressed as a decimal
Social relationship (SR)	SR1	Binary	0	1	Non-attendance at events
	SR2	Ordinal[b]	1	6	Contacts the family by phone
	SR3	Ordinal[b]	1	6	Contacts friends by phone
Health	Health1	Ordinal[c]	0	3	Having mediocre, poor, or very poor health
	Health2	Binary	0	1	Having a chronic disease
	Health3	Ordinal[d]	0	2	Daily activities are limited by a health problem
LS	LS	Ordinal[e]	1	5	In general, how satisfied are you with your life?

[a] No, one, two, three. [b] Daily, weekly, monthly, annual, never. [c] Good or very good health, mediocre, poor, very poor. [d] No limit, limited, severely limited. [e] Very satisfied, quiet satisfied, satisfied, dissatisfied, very dissatisfied.

The data for this research were obtained from the Living Conditions Survey of the Region of Murcia. The universe was private households in the Region of Murcia. The procedure used was probabilistic sampling with proportional allocation to the population of each census section, and the selection of private households was carried out by simple random sampling. Questionnaires were collected between 2017 and 2018. The initial database size was 2482 samples; after eliminating some questionnaires with missing data, the database contained 2462 samples. This represents the largest survey of this type conducted in this territory, quadrupling the number of observations collected by the Living Conditions Survey of the National Statistics Institute [60] for the Region of Murcia. Additionally, the data is geolocalized, which is a requisite for a more precise analysis of LS in relation to urbanization. The composition of this sample according to degrees of urbanization was as follows: 14.22% live in rural areas, 62.63% in intermediate areas, and 23.15% in urban areas.

Table 1 summarizes the different indicators, their association with each construct, the type of variable, the minimum and maximum values, and a brief description. All indicators were based on [60]. It should be noted that the LS scale is a five-point Likert scale, not an eleven-point scale, where 1 represents "very satisfied" and 5 represents "very dissatisfied". This same indicator with a five-point scale has been used in [61]. All the variables were expressed in terms of deprivation, with a higher value indicating a more negative aspect.

4.3. Statistical analysis

Structural equation models can be divided into two techniques: PLS-SEM and covariance-based SEM (CB-SEM) [62]. The main reason for using PLS-SEM is its exploratory nature (CB-SEM has a confirmatory purpose) since there is still no developed theory on LS and its relationship with degrees of urbanization. PLS-SEM also has a predictive purpose [63]. However, the aim of this research is to explain the constructs that are determinants of LS and analyze the changes in the relationships depending on degrees of urbanization. The characteristics of the data are also crucial in determining the suitability of one technique over another. PLS-SEM is a non-parametric technique, which means that the data can exhibit a non-normal distribution [64]. Consequently, metric, ordinal, and binary indicators can be incorporated into the model [31, 65]. As a non-parametric technique, parameter estimates were accompanied by a confidence interval, which was calculated using a bootstrap procedure [66]. Specifically, at least 10,000 subsamples are recommended [67].

One of the commonly asserted benefits of PLS-SEM is that a large sample size is not necessary. Nevertheless, it is crucial to consider the recommended sample size. The most straightforward approach, as evidenced in the literature, is to multiply the maximum number of relations a construct receives by ten. Hair et al. [31] proposed an alternative method that considers the significance level and the minimum R^2 for a statistical power of 80%. Kock and Hadaya [68] suggested the inverse square root method, which is similar to the previous one but replaces the minimum R^2 with the minimum path coefficient of the model. Following these three methods, a sample size of 30, 124, or 619, respectively, was advisable [29].

A PLS-SEM comprises two distinct components: a measurement model and a structural model. Also called outer and inner, respectively [69]. The measurement model analyzes the relationships between an observed variable or indicator and an unobserved variable or construct. The nature of a construct is either formative or reflective, depending on its relationship with its indicators [70]. In a reflective construct, the indicators are a consequence of the construct, whereas in a formative construct, the indicators cause the construct. In the literature, the reflective construct is also referred to as "Mode A" and the formative construct "Mode B" [62, 71]. Therefore, the constructs were conceptualized as reflective. The reason was that this model assesses a determinant of LS through a series of indicators, which provide information about its characteristics. Consequently, the change in a determinant is reflected in all its indicators.

The structural model alters the perspective and assesses the interrelationships among the constructs. In this instance, the constructs may be either endogenous or exogenous [31]. In the present model, social relations and economic situation were exogenous constructs because they did not receive any dependency relationship, while health and LS are endogenous constructs. The weighting scheme for path estimation was path weighting [66, 69].

To evaluate the reflective measurement model, it is necessary to examine the indicator loadings, construct's convergent validity, internal consistency reliability, and discriminant validity. An indicator loading (λ) measures the strength of the relationship between an indicator and its construct. Each loading must be statistically significant and reach a value of at least 0.4, with an ideal value greater than 0.7 [73, 77]. Convergent validity analyzes the proportion of variance shared by the indicators of a construct [72]. For this purpose, the average variance extracted (AVE) was calculated, which is defined as the mean of the squared loadings. Specifically, values greater than 0.5 are considered adequate. The mathematical formulation is as follows:

$$AVE = \frac{\sum_{i=1}^{n} \lambda_i^2}{n} \tag{4.1}$$

where n is the number of indicators of a construct.

Internal consistency reliability examines whether the indicators measure the same latent variable. Acceptable values should be between 0.7 and 0.9, and values exceeding 0.95 are problematic. However, as this is an exploratory study, values between 0.6 and 0.7 are acceptable [73]. There are three indicators to evaluate internal consistency reliability: Cronbach's α, composite reliability A (ρ_a), and composite reliability C (ρ_c). Cronbach's α is the most conservative indicator and usually represents the lower bound of internal consistency reliability, while ρ_c is more liberal and typically represents the upper bound. Cronbach's α differs from ρ_c in that all indicator loadings are equal and it is sensitive to the number of indicators. Unlike the previous indexes, ρ_a does not evaluate the loadings. Instead, it evaluates the weights of the constructs. The value of ρ_a is usually between the two extremes [74].

Finally, the discriminant validity analyzes the uniqueness of each construct. That is, it must measure factors that are different from those of other constructs. For this purpose, the Fornell-Lacker criterion and the HTMT test were employed. The Fornell-Lacker criterion determines that the square root of the AVE must be greater than any of the correlations with another construct [75]. In the HTMT test, values should not exceed 0.9 when the constructs are different from each other or 0.85 when the constructs are similar [76].

The structural model is initially evaluated to determine whether there are collinearity issues by calculating the variance inflation factor (VIF) for each construct, with an ideal value of less than 5 [29]. Secondly, the statistical significance of each path coefficient (β) is evaluated. Path coefficients are analogous to a standardized regression coefficient, they are interpreted as the change that exists in the endogenous latent variable when there is an increase of one standard deviations in the exogenous latent variable, ceteris paribus [77]. Thirdly, the explanatory power of the model is analyzed by R^2 and the effect size by f^2. In general, an R^2 of 0.25 is considered weak, 0.5 is moderate, and 0.75

is substantial [63]. However, in some social sciences, the R^2 is typically lower. For instance, in the reviewed literature, the highest R^2 values were between 0.15 and 0.25. The effect size f^2 is calculated by comparing the change in the R^2 produced by omitting a particular construct. An f^2 of 0.02 is considered small, 0.15 is medium, and 0.35 is large [78]. Mathematically, it can be expressed as follows:

$$f^2 = \frac{R^2\, included\ -\ R^2\, excluded}{1 - R^2\, included} \tag{4.2}$$

Although the purpose of this research is exploratory, it is advisable to evaluate the predictive power of the model. For this purpose, the *PLSpredict* method was performed, which is based on k-fold cross-validation [79]. This concept divides the total sample equally into k subgroups, distinguishing between $k-1$ training subgroups and a holdout subgroup. Then, the estimation model uses only the training subgroups to predict the values of the holdout subgroup [79, 80]. To ensure consistency of results, this process can be repeated r times. The k-fold cross-validation was conducted with $k = 10$ and the process was repeated 10 times. Finally, the degree of prediction error was evaluated by calculating the difference between the root-mean-square error (RMSE) and mean absolute error (MAE) indicators obtained by the PLS model and a linear model [80]. If the RMSE or MAE is lower in the PLS model than in the native model, the model has a high degree of predictive power. If the prediction errors are larger in the PLS model, the model lacks predictive power [29].

Furthermore, this research assessed the mediating effect of health and the moderating effect of the degree of urbanization on LS. A mediating effect occurs when an exogenous variable influences an endogenous variable through a third variable [81, 82]. In this instance, health functions as a mediating construct in the model. Under this circumstance, the total effect of an exogenous construct is the sum of the direct effect on the endogenous construct and the indirect effect through the mediating construct. In order to evaluate the health mediating effect, the statistical significance and sign of the direct and indirect effects were studied. In this sense, the mediating effect is only indirect when the direct effect is not statistically significant, complementary when both effects are significant and of the same sign, and competitive when they are both significant but have different signs [83]. Subsequently, the strength of the mediating effect was analyzed through the variance accounted for (VAF), which represents the weight of the indirect effect over the total effect. This can also be expressed mathematically:

$$VAF = \left(\frac{Indirect\, effect}{Total\, effect}\right) \times 100 \tag{4.3}$$

A VAF of less than 20% indicates the absence of mediating effect, and a VAF of 80% signifies full mediating effect. The remaining cases are classified as partial mediating effect [31].

A moderation effect occurs when the relationship between two variables depends on a third variable [81, 84]. There are two analytical approaches to examining a moderating effect: the addition of an interaction term to the model and multi-group analysis [29]. The rationale for employing MGA-PLS can be attributed to two key factors. Primarily, the degree of urbanization is conceptualized as a general moderator, exerting an influence on all relationships within the model. Secondly, it is a categorical variable. MGA-PLS is based on dividing the sample into groups, estimating a model for each group, and then analyzing whether there are statistically significant differences between the path coefficients of these estimated models [66, 85].

5. Results

5.1. Life satisfaction by degree of urbanization

Did the paradox of rural happiness exist in the Region of Murcia? Table 2 presents the results of the LS scale and its mean, with the LS scale ranging from very satisfied (1) to very dissatisfied (5).

Table 2. Life satisfaction and urbanization.

	Very satisfied	Quite satisfied	Satisfied	Dissatisfied	Very dissatisfied	Average dissatisfied
Total	12.10	24.98	45.37	14.46	3.09	2.714
Rural areas	11.71	20.57	45.43	17.71	4.57	2.829
Intermediate areas	12.78	25.75	46.69	12.71	2.08	2.656
Urban areas	10.53	25.61	41.75	17.19	4.91	2.804

The data indicated that there was no discernible negative linear relationship between degrees of urbanization and LS. When the first three LS categories were included, residents of intermediate areas were the most satisfied with their lives, with 85.22%, compared to 77.71% and 77.89% for residents of rural areas and urban areas, respectively. Residents of the intermediate areas exhibited the lowest percentages of dissatisfied and very dissatisfied responses. However, the urban population had the lowest percentage of individuals who reported being "very satisfied" and the highest percentage of individuals who reported being "very dissatisfied". When calculating the mean of the scores, the rural population showed the highest average of dissatisfied responses (2.829), followed closely by their counterparts in urban areas (2.804). The population in the intermediate areas had the lowest average of dissatisfied (2.656).

A hypothesis test was performed to determine the average of dissatisfied responses according to degrees of urbanization. Since degrees of urbanization is not a normal variable, a non-parametric test, like the Kruskal-Wallis test, was chosen. The homoscedasticity requirement was met since the Fligner-Killeen test reached a p-value of 0.329. Thus, there was no evidence against equality of variances. The Kruskal-Wallis test yielded a p-value of 0.001, indicating the presence of significant differences in the distribution of LS according to degrees of urbanization. The between-group comparison of the Mann-Whitney test revealed that significant differences (p-value < 0.05) existed between intermediate areas and rural areas and intermediate areas and urban areas. The contrast between rural and urban areas was not statistically significant.

5.2. Assessment of the measurement model

Table 3 shows the results of the reliability analysis, including the loadings, t-student values, and confidence intervals for each indicator. The results were satisfactory, as all the indicators were significant, with p-values close to 0.001. Moreover, all the loadings were greater than 0.7, except SR2, which had a loading of 0.648.

Table 4 presents the results of the internal consistency reliability and the convergent validity. The internal consistency reliability results were adequate. Given the exploratory nature of the investigation, the threshold was set at 0.6. With this reference, all the constructs showed Cronbach's α, ρ_a, and ρ_c values greater than 0.6, and none of the values exceeded 0.95. The convergent validity also gave good results; the AVE was greater than 0.5 in all constructs. The social relationships construct showed the lowest values for the indicators examined.

Table 3. Indicator reliability.

Indicator	Loading	t-student	CI 2.5%	CI 97.5%
ECO1	0.862	122.388	0.848	0.876
ECO2	0.712	49.023	0.681	0.738
ECO3	0.732	67.165	0.710	0.753
ECO4	0.821	80.613	0.800	0.840
SR1	0.775	42.883	0.739	0.810
SR2	0.648	23.444	0.589	0.698
SR3	0.785	47.755	0.749	0.814
Health1	0.889	151.077	0.877	0.900
Health2	0.822	82.155	0.801	0.841
Health3	0.892	133.679	0.878	0.904
LS	1.000		1.000	1.000

Note: Performed with 10,000 bootstrapping repetitions, all the loadings are significant at 0.001.

Table 4. Internal consistency reliability and convergent validity.

Construct	Cronbach's α	ρ_c	ρ_a	AVE
ECO	0.788	0.864	0.795	0.615
SR	0.607	0.781	0.624	0.545
Health	0.839	0.902	0.868	0.754
LS	1.000	1.000	1.000	1.000

Subsequently, the discriminant validity of the model was examined using the Fornell-Lacker criterion and the HTMT test. Table 5 provides a matrix whose diagonal corresponds to the AVE of each construct squared, while the rest of the table displays the correlations between the constructs. In this case, no correlation was greater than the value of the diagonal. Therefore, the Fornell-Lacker criterion reported that the model met the discriminant validity criterion. Table 6 presents the results of the HTMT test. None of the values exceeded 0.85, and they were substantially below that threshold. Thus, the model successfully met the assumption of discriminant validity.

Table 5. Discriminant validity: Fornell-Lacker criterion results.

	ECO	SR	Health	LS
ECO	0.784			
SR	0.383	0.739		
Health	0.242	0.383	0.868	
LS	0.476	0.323	0.321	1.000

Table 6. Discriminant validity: HTMT results.

	ECO	SR	Health	LS
ECO				
SR	0.498			
Health	0.284	0.505		
LS	0.536	0.380	0.340	

5.3. Assessment of the structural model

To evaluate the structural model, the initial step was to assess the presence of collinearity. Table 7 shows the VIF for each of the constructs. All the values exhibited a range of 1.172 to 1.310, well below the threshold of 5. Therefore, there were no collinearity problems in the model.

Table 7. VIF results.

	ECO	SR	Health	LS
ECO				
SR				
Health	1.172	1.172		
LS	1.187	1.310	1.187	

Table 8 shows the path coefficients, the *t*-student value, the confidence intervals, and the f^2 value. All the path coefficients were statistically significant, with *p*-values approaching 0.001, confirming the proposed hypotheses. In addition, all the path coefficients had a positive sign. The highest path coefficient was for the relationship from economic situation to LS, with 0.392. It was followed by the path coefficient of the effect of social relations on health, with 0.341. The path coefficient of the direct effect of health was also significant, with 0.187. The effect of individuals' economic situation on health and the direct effect of social relations were last.

Table 8. Assessment of the path coefficients.

Indicator	Beta	*t*-student	CI 2.5%	CI 97.5%	f^2
ECO → LS	0.392	21.510	0.356	0.428	0.180
SR → LS	0.101	4.932	0.062	0.141	0.011
Health → LS	0.187	9.493	0.147	0.225	0.041
SR → Health	0.341	16.002	0.300	0.383	0.117
ECO → Health	0.111	4.906	0.067	0.156	0.013

Note: Performed with 10,000 bootstrapping repetitions, all the path coefficients are significant at 0.001.

The explanatory power of the model was examined using R^2 and f^2. The LS obtained an R^2 of 0.280, while health yielded an R^2 of 0.158. Overall, only the R^2 of LS could be considered weak. However, given the field of study, this R^2 could be considered satisfactory. Regarding the f^2, the results followed the same order as those of path coefficient intensity. Nevertheless, only the f^2 of the direct effect of the economic situation was considered to be of medium size, while the indirect effect of social relations and the direct effect of health were categorized as small. The rest were insignificant. Figure 3 shows the graphical representation of the model.

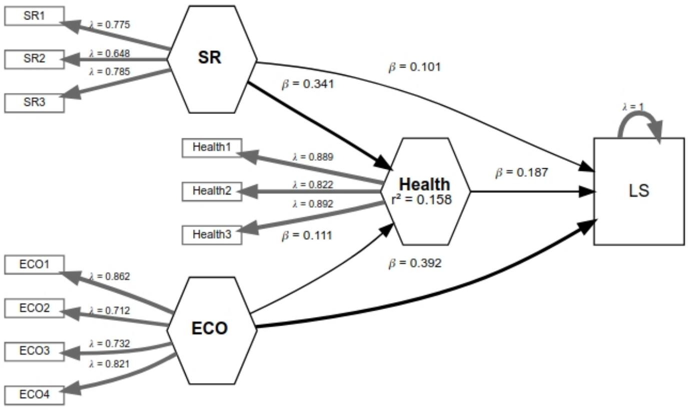

Figure 3. Results of the model.

Table 9 shows the total effect of each relationship and its confidence interval. All the total effects were statistically significant, with p-values close to 0.001. The two indirect effects were statistically significant and complementary. Further analysis was conducted using the VAF. In this analysis, health was identified as a partial mediator of social relationships, with a VAF of 38.8%. However, the strength of the mediating effect of health on economic situation was very low, as the VAF remained at 5.1%.

Table 9. Assessment of the total effects.

Indicator	Total effect	CI 2.5%	CI 97.5%
ECO → LS	0.413	0.377	0.449
SR → LS	0.165	0.128	0.202
Health → LS	0.187	0.147	0.225

Note: Performed with 10,000 bootstrapping repetitions.

Finally, Table 10 shows the predictive power of the model using a k-fold cross-validation process. The values of both RMSE and MAE in the PLS model were higher than those of the native model. This was true for all the exogenous indicators, although the difference for LS was minimal. Consequently, the model did not demonstrate predictive power.

Table 10. Assessment of the model's predictive power.

	RMSE PLS out	MAE PLS out	RMSE LM out	MAE LM out
Health1	0.622	0.478	0.601	0.448
Health2	0.434	0.371	0.427	0.363
Health3	0.545	0.419	0.534	0.407
LS	0.816	0.655	0.812	0.651

Note: Performed using 10 folds and 10 repetitions.

5.4. Rural-urban analysis

After analyzing the general model, its behavior was studied according to degrees of urbanization. However, the MGA-PLS technique is limited to comparisons between two groups, while the data are organized in three categories: rural areas, intermediate areas, and urban areas. Therefore, it was necessary to group the analysis into two categories versus one. Table 11 presents the path coefficients by groups and the *p*-values for the three different models. The first model (Model 1) contrasted rural (Group 1) and urban (Group 2) areas, including intermediate and urban areas. The second model (Model 2) compared rural and intermediate areas (Group 1) with urban areas (Group 2). The third model (Model 3) contrasted rural and urban areas (Group 1) with intermediate areas (Group 2).

Table 11. MGA-PLS results.

	Model 1			Model 2			Model 3		
Indicator	Group 1	Group 2	*p*-value	Group 1	Group 2	*p*-value	Group 1	Group 2	*p*-value
ECO → LS	0.426	0.388	0.222	0.382	0.426	0.860	0.423	0.373	0.074
SR → LS	0.072	0.103	0.702	0.095	0.127	0.771	0.106	0.095	0.388
Health → LS	0.157	0.193	0.764	0.169	0.238	0.949	0.208	0.171	0.182
SR → Health	0.306	0.346	0.737	0.331	0.370	0.797	0.345	0.336	0.405
ECO → Health	0.127	0.109	0.395	0.125	0.068	0.140	0.091	0.125	0.770

Note: Performed with 2000 bootstrapping repetitions.

In Model 1, no significant differences were observed among the path coefficients. Therefore, it could not be concluded with certainty that the intensity of the path coefficients differed between rural and urban areas. However, this changed in the remaining models. In Model 2, a slight difference emerged with the direct effect of health on LS, showing a *p*-value approaching 0.05. This revealed that health had a greater impact on LS for inhabitants of urban areas than for inhabitants of rural and intermediate areas. In Model 3, there was also a slightly significant difference in the direct effect of individuals' economic situation on LS, with a *p*-value of less than 0.1. In this case, the economic situation affected LS to a greater extent for residents of rural areas and urban areas than those of intermediate areas. These results indicated that the degree of urbanization can be considered a partial moderator of LS, given the differences observed in the intensity of two of the three direct effects.

6. Discussion

In contrast to the findings of Requena [26] and Sørensen [18], the results did not indicate the urban paradox existed in Murcia. Those residing in rural areas exhibited the lowest LS, closely followed by those in urban areas. These differences were not statistically significant, in line with other research [16, 34]. However, it should be noted that the relationship between degrees of urbanization and LS depends on the economic development of the territory in question [12]. The less economic development there is, the lower the LS among the rural population. Thus, it is consistent with the fact that the rural population of Murcia perceives less LS. However, this paper indicates that individuals residing in intermediate areas report greater LS, which is in alignment with the findings of Lenzi and Perucca [17]. Nevertheless, this paper refutes the findings of Morrison [13], which posit that the LS in small cities is between rural and urban areas, except at one point in the process of economic development, where the LS of the three areas coincide. In any case, the negative externalities derived from living in an urban area, as evidenced by numerous studies, may play a role in the Region of Murcia since the urban areas do not present the highest LS.

The results of the model are in accordance with the findings of the existing literature. The most important determinant in LS was the economic situation [41]. Also, a strong relationship between social relations and health was found [8]. Health was a partial mediator of the effect of social relations on LS. In contrast to the findings of Moreno-Agostino et al. [55], the strength of the mediating effect of health on economic status was very weak. Nonetheless, the explanatory power of the model was low, and it had almost no predictive power. This is in line with research in this field, where the models obtain very low R^2 values. The R^2 value obtained in this research is among the highest in the literature reviewed.

The results of the multi-group analysis to compare the intensity of the theoretical relationships indicated some significant differences according to degrees of urbanization. The degree of urbanization does not act as a mediator among the determinants of LS but as a partial moderator. The two path coefficients showing differences were the direct effects of health and economic situation on LS.

There were significant differences in the direct relationship between health and LS, indicating that as the degree of urbanization increases, so does its intensity. Thus, the largest effect of health on LS was among residents of urban areas. These results differ from those presented by Bernini and Tampieri [36], who argued that the lack of health services in rural areas is an explanatory factor for its greater importance. However, given that these are self-reported health variables, biases may exist. In this context, individuals with greater access to medical care and a higher level of education are more likely to self-diagnose poor health or perceive they are ill [86]. Therefore, the urban population places greater value on health, while rural populations adapt to worse health situations.

The significance of economic situation depends on the degree of urbanization. According to Brereton et al. [27], income is more important among urban residents than rural residents, mainly due to housing prices. Nevertheless, the outcomes of this study diverge to some extent, as income exerts a more pronounced influence on LS in rural and urban areas than in intermediate areas. Housing prices present a greater challenge in large cities than in intermediate and rural areas. However, given the higher rates of poverty among the rural population in the Region of Murcia [24], it is feasible that residents in rural areas value income more highly than their counterparts in intermediate areas.

7. Conclusions

This research contributes to the study of LS according to degrees of urbanization through an innovative methodology that combines geographic grid and structural equation modeling. In this case, three facts stand out. First, we have made an objective and precise identification of the degree of urbanization to which the sample belongs. Second, the causal relationships among the main determinants have been studied. Third, the changes in the relationships among the main determinants have been analyzed according to three degrees of urbanization and for the population as a whole.

As a result, the rural population of Murcia is less satisfied with their lives than the urban population. Therefore, the paradox of rural happiness does not exist in the Region of Murcia. As there is no literature on this field at the national level, it has not been possible to confirm whether this trend extends to the country as a whole or only to certain less-developed regions. In any case, this study indicates that the highest LS is found in the population residing in intermediate areas. Consequently, it is recommended that future articles consider more categories of urbanization than the simple dichotomy between rural and urban, besides being as accurate as possible. Including more categories would make the debate more complex, but it would enrich the information necessary to understand the relationship between degrees of urbanization and LS.

All the hypotheses proposed are statistically significant, indicating that an individual's economic situation has the strongest effect on LS. However, this research empirically confirms the mediating role of health between social relations and LS. Therefore, this profound relationship cannot be ignored in future research on the determinants of LS.

Finally, it should be noted that differences are found in the model when applied to areas with different degrees of urbanization. Specifically, these differences are the direct relationships of economic situation and health on LS. In short, degrees of urbanization can be understood as a moderator of LS. Nevertheless, it remains to be confirmed whether the economic development of a territory is a first-order moderator in the Spanish context.

Stiglitz et al. [87] recommend incorporating SWB measures to assess an individual's quality of life. In this case, any public policy becomes meaningful when it improves the quality of life of its citizens. Thus, this study can serve as a guide for the formulation and evaluation of public policies. Analyzing the different determinants of LS can guide the public policies needed to improve SWB. Moreover, it can provide a subjective assessment of the efficacy of a public policy, namely whether an intervention has improved the quality of life of its beneficiaries.

One limitation of this research is the scarcity of available variables. Consequently, LS is measured by a single indicator construct. This may contribute to the model's lack of predictive power [88]. Another limitation is the insufficient maturity of this field, which is evident by the fact that all the models analyzed have very low explanatory power. In conclusion, further research is needed to clarify the underlying causes of the discrepancies and paradoxes observed in the literature on LS and its distribution across the territory.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work is based on the survey carried out during the project "Living conditions and social welfare of the population at risk of poverty and social exclusion in the Region of Murcia", financed by the Anti Poverty Network of Region of Murcia (Spain), who bear no responsibility for the analysis or interpretations presented here.

Conflict of interest

The authors declare there is no conflicts of interest.

References

1. E. Diener, Subjective well-being, *Psych. Bull.*, **95** (1984), 542–575. https://doi.org/10.1037/0033-2909.95.3.542
2. D. Kahneman, A. B. Krueger, Developments in the measurement of subjective well-being, *J. Econ. Perspect.*, **20** (2006), 3–24. http://doi.org/10.1257/089533006776526030
3. E. Diener, S. Oishi, L. Tay, Advances in subjective well-being research, *Nat. Hum. Behav.*, **2** (2018), 253–260. https://doi.org/10.1038/s41562-018-0307-6
4. R. Veenhoven, *Happiness in Nations: Subjective Appreciation of Life in 56 Nations 1946–1992*, Erasmus University Rotterdam, Rotterdam, Netherlands, 1993.
5. R. A. Easterlin, Does economic growth improve the human lot? Some empirical evidence, in *Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz*, (eds. P. A. David and M. W. Reder), Academic Press, (1974), 89–125. https://doi.org/10.1016/B978-0-12-205050-3.50008-7
6. R. A. Easterlin, Income and happiness: Towards a unified theory, *Econ. J.*, **111** (2001), 465–484. http://www.jstor.org/stable/2667943
7. D. G. Blanchflower, A. J. Oswald, Well-being over time in Britain and the USA, *J. Public Econ.*, **88** (2004), 1359–1386. https://doi.org/10.1016/S0047-2727(02)00168-8
8. J. F. Helliwell, R. D. Putnam, The social context of well-being, *Philos. Trans. R. Soc. London. Ser B Biol. Sci.*, **359** (2004), 1435–1446. https://doi.org/10.1098/rstb.2004.1522
9. P. Dolan, T. Peasgood, M. White, Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being, *J. Econ. Psych.*, **29** (2008), 94–112. https://doi.org/10.1016/j.joep.2007.09.001
10. R. Layard, *Happiness: Lessons from a New Science*, The Penguin Press, New York, 2005.
11. A. Peiró, Happiness, satisfaction and socio-economic conditions: Some international evidence, *J. Socio-Econ.*, **35** (2006), 348–365. https://doi.org/10.1016/j.socec.2005.11.042
12. R. A. Easterlin, L. Angelescu, J. S. Zweig, The impact of modern economic growth on urban-rural differences in subjective well-being, *World Dev.*, **39** (2011), 2187–2198. https://doi.org/10.1016/j.worlddev.2011.04.015
13. P. S. Morrison, Wellbeing and the region, in *Handbook of Regional Science*, (eds. M. M. Fisher and P. Nijkamp), 2nd edition, Springer, (2020), 779–798. https://doi.org/10.1007/978-3-642-36203-3-16-1
14. J. F. L. Sørensen, The rural happiness paradox in developed countries, *Soc. Sci. Res.*, **98** (2021), 102581. https://doi.org/10.1016/j.ssresearch.2021.102581
15. R. Veenhoven, How satisfying is rural life? Fact and value, in *Changing Values and Attitudes in Family Households, Implications for Institutional Transition in East and West*, (eds. J. Cecora), Society for agricultural policy research in rural society, (1994), 41–51.
16. M. Shucksmith, S. Cameron, T. Merridew, F. Pichler, Urban-rural differences in quality of life across the European Union, *Reg. Stud.*, **43** (2009), 1275–1289. https://doi.org/10.1080/00343400802378750

17. C. Lenzi, G. Perucca, Are urbanized areas source of life satisfaction? Evidence from EU regions, *Pap. Reg. Sci.*, **97** (2018), 105–122. https://doi.org/10.1111/pirs.12232

18. J. F. L. Sørensen, Rural-urban differences in life satisfaction: Evidence from the European Union, *Reg. Stud.*, **48** (2014), 1451–1466. https://doi.org/10.1080/00343404.2012.753142

19. A. Okulicz-Kozaryn, J. M. Mazelis, Urbanism and happiness: A test of Wirth's theory of urban life, *Urban Stud.*, **55** (2018), 349–364. https://doi.org/10.1177/0042098016645470

20. A. T. Piper, Europe's capital cities and the happiness penalty: An investigation using the European Social Survey, *Soc. Indic. Res.*, **123** (2015), 103–126. https://doi.org/10.1007/s11205-014-0725-4

21. P. S. Morrison, M. Weckroth, Human values, subjective well-being and the metropolitan region, *Reg. Stud.*, **52** (2017), 325–337. https://doi.org/10.1080/00343404.2017.1331036

22. A. Losa, M. A. Esteban-Yago, E. Gadea, M. B. García-Romero, M. Á. Sánchez-García, J. A. Sánchez-Martí, *Condiciones de Vida y Bienestar Social de la Población en Riesgo de Pobreza y Exclusión Social en la Región de Murcia*, EAPN-Región de Murcia, Murcia, España, 2020.

23. A. Losa, M. A. Esteban-Yago, E. Gadea, M. B. García-Romero, M. Á. Sánchez-García, J. A. Sánchez-Martí, *Población en Riesgo de Pobreza y Exclusión Social en la Región de Murcia: Un Análisis Territorial sobre Condiciones de Vida y Bienestar Social*, EAPN-Región de Murcia, Murcia, España, 2020.

24. J. A. Sánchez-Martí, M. A. Esteban-Yago, A. Losa, Risk of poverty and social exclusion among the urban and rural population in the Region of Murcia, Spain, *Stud. Appl. Econ.*, **40** (2022), 1–17. https://doi.org/10.25115/eea.v40i2.7912

25. J. Colino, F. Martínez-Carrasco, A. Losa, J. M. Martínez-Paz, A. Pérez-Morales, J. A. Albaladejo, *Las Zonas Rurales de la Región de Murcia*, Consejo Económico y Social de la Región de Murcia, Murcia, España, 2022.

26. F. Requena, Rural-urban living and level of economic development as factors in subjective well-being, *Soc. Indic. Res.*, **128** (2016), 693–708. https://doi.org/10.1007/s11205-015-1051-1

27. F. Brereton, C. Bullock, J. P. Clinch, M. Scott, Rural change and individual well-being: The case of Ireland and rural quality of life, *Eur. Urban Reg. Stud.*, **18** (2011), 203–227. https://doi.org/10.1177/0969776411399346

28. A. Ferrer-i-Carbonell, P. Frijters, How important is methodology for the estimates of the determinants of happiness?, *Econ. J.*, **114** (2004), 641–659. https://doi.org/10.1111/j.1468-0297.2004.00235.x

29. J. F. Hair, G. T. M. Hult, C. M. Ringle, M. Sarstedt, N. P. Danks, S. Ray, *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook*, Springer, Switzerland, 2021.

30. J. F. Hair, A. Alamer, Partial least squares structural equation modeling (PLS-SEM) in second language and education research: Guidelines using an applied example, *Res. Method. Appl. Linguist.*, **1** (2022), 100027. https://doi.org/10.1016/j.rmal.2022.100027

31. J. F. Hair, G. T. M. Hult, C. M. Ringle, M. Sarstedt, *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, Sage, Los Angeles, CA, 2014.

32. R. Veenhoven, M. Berg, Has modernisation gone too far? Modernity and happiness in 141 contemporary nations, *Int. J. Happiness Dev.*, **1** (2013), 172–195. https://doi.org/10.1504/IJHD.2013.055645

33. M. J. Burger, P. S. Morrison, M. Hendriks, M. M. Hoogerbrugge, Urban-rural happiness differentials across the world, in *World Happiness Report 2020*, (eds. J. F. Helliwell, R. Layard, J. Sachs and J. E. De Neve), Sustainable Development Solutions Network, (2020), 66–93.

34. L. Dijkstra, E. Papadimitriou, Using a new global urban-rural definition, called the degree of urbanisation, to assess happiness, in *World Happiness Report 2020*, (eds. J. F. Helliwell, R. Layard, J. Sachs and J. E. De Neve), Sustainable Development Solutions Network, (2020), 146–151.

35. M. Weckroth, T. Kemppainen, (Un) Happiness, where are you? Evaluating the relationship between urbanity, life satisfaction and economic development in a regional context, *Reg. Stud., Reg. Sci.*, **8** (2021), 207–227. https://doi.org/10.1080/21681376.2021.1925146

36. C. Bernini, A. Tampieri, The mediating role of urbanization on the composition of happiness, *Pap. Reg. Sci.*, **101** (2022), 639–658. https://doi.org/10.1111/pirs.12671

37. S. Savahl, S. Adams, M. Florence, F. Casas, M. Mpilo, D. Isobell, et al., The relation between children's participation in daily activities, their Engagement with family and friends, and subjective well-being, *Child Indic. Res.*, **13** (2020), 1283–1312. https://doi.org/10.1007/s12187-019-09699-3

38. J. Wang, J. Zhao, Y. Wang, Self-efficacy mediates the association between shyness and subjective well-being: The case of Chinese college students, *Soc. Indic. Res.*, **119** (2014), 341–351. https://doi.org/10.1007/s11205-013-0487-4

39. X. Mao, W. Han, Living arrangements and older adults' psychological well-being and life satisfaction in China: Does social support matter?, *Fam. Relat.*, **67** (2018), 567–584. https://doi.org/10.1111/fare.12326

40. O. Schilling, H. W. Wahl, Family networks and life-satisfaction of older adults in rural and urban regions, *Koln. Z. Soziol. Sozpsych.*, **54** (2002), 304–307. https://doi.org/10.1007/s11577-002-0041-x

41. P. Diego-Rosell, R. Tortora, J. Bird, International determinants of subjective well-being: Living in a subjectively material world, *J. Happiness Stud.*, **19** (2018), 123–143. https://doi.org/10.1007/s10902-016-9812-3

42. B. Stevenson, J. Wolfers, Subjective well-being and income: Is there any evidence of satiation?, *Am. Econ. Rev. Pap. Proc.*, **103** (2013), 598–604. http://dx.doi.org/10.1257/aer.103.3.598

43. B. S. Frey, A. Stutzer, Happiness, economy and institutions, *Econ. J.*, **110** (2000), 918–938. https://doi.org/10.1111/1468-0297.00570

44. A. Ferrer-i-Carbonell, Income and well-being: An empirical analysis of the comparison income effect, *J. Public Econ.*, **89** (2005), 997–1019. https://doi.org/10.1016/j.jpubeco.2004.06.003

45. M. F. Bhuiyan, R. S. Szulga, Extreme bounds of subjective well-being: Economic development and micro determinants of life satisfaction, *Appl. Econ.*, **49** (2017), 1351–1378. https://doi.org/10.1080/00036846.2016.1218426

46. B. Hayo, W. Seifert, Subjective economic well-being in Eastern Europe, *J. Econ. Psych.*, **24** (2003), 329–348. https://doi.org/10.1016/S0167-4870(02)00173-3

47. D. Zhou, Y. Xu, P. Ai, The effects of online social interactions on life satisfaction of older chinese adults: New insights based on a longitudinal approach, *Healthcare*, **10** (2022), 1964. https://doi.org/10.3390/healthcare10101964

48. Y. J. Masuda, J. R. Williams, H. Tallis, Does life satisfaction vary with time and income? Investigating the relationship among free time, income, and life satisfaction, *J. Happiness Stud.*, **22** (2021), 2051–2073. https://doi.org/10.1007/s10902-020-00307-8

49. R. Margolis, M. Myrskylä, Family, money, and health: Regional differences in the determinants of life satisfaction over the life course, *Adv. Life Course Res.*, **18** (2013), 115–126. http://dx.doi.org/10.1016/j.alcr.2013.01.001

50. A. Steptoe, A. Deaton, A. A. Stone, Subjective wellbeing, health, and ageing, *Lancet*, **385** (2015), 640–648. https://doi.org/10.1016/S0140-6736(13)61489-0

51. W. Yip, S. V. Subramanian, A. D. Mitchell, D. T. S. Lee, J. Wang, I. Kawachi, Does social capital enhance health and well-being? Evidence from rural China, *Soc. Sci. Med.*, **64** (2007), 35–49. https://doi.org/10.1016/j.socscimed.2006.08.027

52. X. Wang, P. Wang, P. Wang, M. Cao, X. Xu, Relationships among mental health, social capital and life satisfaction in rural senior older adults: A structural equation model, *BMC Geriatr.*, **22** (2022), 73. https://doi.org/10.1186/s12877-022-02761-w

53. M. A. Shields, S. W. Price, Exploring the economic and social determinants of psychological well-being and perceived social support in England, *J. R. Stat. Soc. Ser. A: Stat. Soc.*, **168** (2005), 513–537. https://doi.org/10.1111/j.1467-985X.2005.00361.x

54. J. Yeo, Y. G. Lee, Understanding the association between perceived financial well-being and life satisfaction among older adults: Does social capital play a role?, *J. Fam. Econ. Issues*, **40** (2019), 592–608. https://doi.org/10.1007/s10834-019-09634-2

55. D. Moreno-Agostino, J. De la Fuente, M. Leonardi, S. Koskinen, B. Tobiasz-Adamczyk, A. Sánchez-Niubò et al., Mediators of the socioeconomic status and life satisfaction relationship in older adults: A multi-country structural equation modeling approach, *Aging Ment. Health*, **25** (2021), 585–592. https://doi.org/10.1080/13607863.2019.1698513

56. E. Reig, F. J. Goerlich, I. Cantarino, *Delimitación de Áreas Rurales y Urbanas a Nivel Local*, Fundación BBVA, Bilbao, España, 2016.

57. Eurostat, *Eurostat Regional Yearbook 2012*, Publications Office of the European Union, Luxembourg, 2012.

58. *National Statistics Institute*, Concept Glossary, 2024. Available from: https://www.ine.es/DEFIne/en/concepto.htm?c=4930&tf=&op=30261.

59. *Population and Housing Census*, Madrid: National Statistics Institute, 2011.

60. *Living Conditions Survey*, Madrid: National Statistics Institute, 2023.

61. *Korean Labor and Income Panel Study*, Seoul: Korea Labor Institute, 2024.

62. W. W. Chin, The partial least squares approach for structural equation modeling, in *Modern Methods for Business Research*, (eds. G. A. Marcoulides), Lawrence Erlbaum Associates, (1998), 295–336. https://doi.org/10.1007/978-3-642-36203-3-16-1

63. J. F. Hair, C. M. Ringle, M. Sarstedt, PLS-SEM: Indeed a silver bullet, *J. Mark. Theory Pract.*, **19** (2011), 139–152. https://doi.org/10.2753/MTP1069-6679190202

64. W. Reinartz, M. Haenlein, J. Henseler, An empirical comparison of the efficacy of covariance-based and variance-based SEM, *Int. J. Res. Mark.*, **26** (2009), 332–344. http://dx.doi.org/10.1016/j.ijresmar.2009.08.001

65. M. Sarstedt, J. F. Hair, M. Pick, B. D. Liengaard, L. Radomir, C. M. Ringle, Progress in partial least squares structural equation modeling use in marketing research in the last decade, *Psych. Mark.*, **39** (2022), 1035–1064. https://doi.org/10.1002/mar.21640

66. J. Henseler, C. M. Ringle, R. R. Sinkovics, The use of partial least squares path modeling in international marketing, in *New Challenges to International Marketing*, (eds. R. R. Sinkovics and P. N. Ghauri), Emerald Group Publishing Limited, (2009), 277–319. https://doi.org/10.1108/S1474-7979(2009)0000020014

67. S. Streukens, S. Lezai-Werelds, Bootstrapping and PLS-SEM: A step-by-step guide to get more out of your bootstrap results, *Eur. Manag. J.*, **34** (2016), 618–632. http://dx.doi.org/10.1016/j.emj.2016.06.003

68. N. Kock, P. Hadaya, Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods, *Inf. Syst. J.*, **28** (2018), 227–261. https://doi.org/10.1111/isj.12131

69. M. Tenenhaus, V. E. Vinzi, Y. M. Chatelin, C. Lauro, PLS path modeling, *Comput. Stat. Data Anal.*, **48** (2005), 159–205. https://doi.org/10.1016/j.csda.2004.03.005

70. K. A. Bollen, *Structural Equations with Latent Variables*, John Wiley & Sons, New York, 1989.

71. E. E. Rigdon, Rethinking partial least squares path modeling: In praise of simple methods, *Long Range Plan.*, **45** (2012), 341–358. https://doi.org/10.1016/j.lrp.2012.09.010

72. J. F. Hair, W. C. Black, B. J. Babin, R. E. Anderson, *Multivariate Data Analysis*, 7th edition, Pearson Prentice Hall, New York, 2010.

73. J. F. Hair, J. J. Rischer, M. Sarstedt, C. M. Ringle, When to use and how to report the results of PLS-SEM, *Eur. Bus. Rev.*, **31** (2018), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

74. T. K. Dijkstra, J. Henseler, Consistent partial least squares path modeling, *MIS Q.*, **39** (2015), 297–316. https://www.jstor.org/stable/26628355

75. C. Fornell, D. F. Larcker, Structural equation models with unobservable variables and measurement error, *J. Mark. Res.*, **18** (1981), 382–388. https://doi.org/10.2307/3150980

76. J. Henseler, C. M. Ringle, M. A. Sarstedt, A new criterion for assessing discriminant validity in variance-based structural equation modeling, *J. Acad. Mark. Sci.*, **43** (2015), 115–135. https://doi.org/10.1007/s11747-014-0403-8

77. J. Benitez, J. Henseler, A. Castillo, F. Schuberth, How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research, *Inf. Manag.*, **57** (2020), 103168. https://doi.org/10.1016/j.im.2019.05.003

78. J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd edition, Lawrence Erlbaum Associates, Hillsdale, NJ, 1988.

79. G. Shmueli, S. Ray, J. M. Velasquez, S. Babu, The elephant in the room: Predictive performance of PLS models, *J. Bus. Res.*, **69** (2016), 4552–4564. http://dx.doi.org/10.1016/j.jbusres.2016.03.049

80. G. Shmueli, M. Sarstedt, J. F. Hair, J. H. Cheah, H. Ting, S. Vaithilingam, C. M. Ringle, Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict, *Eur. J. Mark.*, **53** (2019), 2322–2347. https://doi.org/10.1108/EJM-02-2019-0189

81. R. M. Baron, D. A. Kenny, The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations, *J. Personal. Soc. Psych.*, **51** (1986), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

82. C. Nitzl, J. L. Roldan, G. Cepeda, Mediation analysis in partial least squares path modeling: Helping researchers discuss more sophisticated models, *Ind. Manag. Data Syst.*, **116** (2016), 1849–1864. https://doi.org/10.1108/IMDS-07-2015-0302

83. X. Zhao, J. G. Lynch, Q. Chen, Reconsidering Baron and Kenny: Myths and truths about mediation analysis, *J. Consum. Res.*, **37** (2010), 197–206. https://doi.org/10.1086/651257

84. M. Sarstedt, J. Henseler, C. M. Ringle, Multigroup analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results, in *Measurement and Research Methods in International Marketing (Advances in International Marketing, Vol. 22)*, (eds. M. Sarstedt, M. Schwaiger and C. R. Taylor), Emerald Group Publishing Limited, (2011), 195–218. https://doi.org/10.1108/S1474-7979(2011)0000022012

85. M. Klesel, F. Schuberth, B. Niehaves, J. Henseler, Multigroup analysis in information systems research using PLS-PM: A systematic investigation of approaches, *SIGMIS Database*, **53** (2022), 26–48. https://doi.org/10.1145/3551783.3551787

86. A. Sen, Health: Perception versus observation, *BMJ*, **324** (2002), 860–861. https://doi.org/10.1136/bmj.324.7342.860

87. J. E. Stiglitz, A. Sen, J. P. Fitoussi, *Report by the Commission on the Measurement of Economic Performance and Social Progress*, Commission on the Measurement of Economic Performance and Social Progress, Paris, 2009.

88. A. Diamantopoulos, M. Sarstedt, C. Fuchs, P. Wilczynski, S. Kaiser, Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective, *J. Acad. Mark. Sci.*, **40** (2012), 434–449. https://doi.org/10.1007/s11747-011-0300-3

AIMS Press