
How's life? An international classification based on life satisfaction and its determinants

Sergio Tezanos and Borja López-Noval*

Cátedra de Cooperación Internacional y con Iberoamérica and
Departamento de Economía, Universidad de Cantabria,

Av. los Castros SN, 39005 Santander, Spain

ORCID: 0000-0002-7202-5472

ORCID: 0000-0002-0550-8976

Email: sergio.tezanos@unican.es

Email: borja.lopeznoval@gmail.com

*Corresponding author

Abstract: Average life evaluations significantly vary across countries due to several factors, such as income, health, social support, freedom, generosity and corruption. In this paper we carry out an analysis of the joint distribution of average life satisfaction and five key determinants in 103 countries by means of a hierarchical cluster analysis. We build a life satisfaction taxonomy that identifies five groups of countries: two comprise relatively dissatisfied countries, one includes moderately satisfied countries, and the remaining two highly satisfied countries. The contribution of the taxonomy is twofold. First, it provides the first systematic classification of countries based on life satisfaction and its determinants and suggests that previous classifications are not as robust as it is usually assumed. Second, the taxonomy contributes to the discussion on the meaning and measurement of well-being. Interestingly, it shows that different configurations of the classification variables may be associated with similar levels of life satisfaction.

Keywords: life satisfaction; cross-country analysis; cluster analysis; multidimensional taxonomy; international classification.

Reference to this paper should be made as follows: Tezanos, S. and López-Noval, B. (2022) 'How's life? An international classification based on life satisfaction and its determinants', *Int. J. Happiness and Development*, Vol. 7, No. 1, pp.64–90.

Biographical notes: Sergio Tezanos is PhD in International Economics and Development by the University Complutense of Madrid and associate professor at the Department of Economics of the University of Cantabria (Spain) and deputy director of the Ibero-American Research Office in International Development & Co-operation of this university. His research fields are human development (with a focus on international development taxonomies), international cooperation policies (with emphasis on the effectiveness and allocation of Official Development Assistance, ODA), international migration (from developing countries to developed countries) and the epistemological analysis of "development studies".

Borja López-Noval is PhD in Economics by the University of Cantabria (Spain) where he currently teaches at the Department of Economics. His research focuses on the relationship between objective and subjective

well-being indicators, with a special attention to education, social capital and religiosity. He has also studied the historical foundations of the concept of quality of life.

1 Introduction

Since the United Nations (UN) launched the *World Happiness Report* in 2012 the wide public has got used to news such as “for the third year in a row, Finland has placed at the top of the list as the happiest country in the world”; and “this year, Afghanistan was named the unhappiest place” (Bloom, 2020). There is actually evidence suggesting that cross-cultural comparisons of life satisfaction measures have a large degree of validity (Blanchflower and Oswald, 2008; Helliwell et al., 2019).

At the cross-country level, there is a well-established happiness model that includes six seemingly ‘universal’ factors: income, health, social support, freedom, generosity and corruption (Helliwell and Wang, 2013). Overall, Helliwell and Wang’s model explains three-quarters of the variation in average life evaluations across countries. The success of the model is consistent with the evidence showing that the structure of the happiness equation is rather homogeneous across countries (e.g., Graham and Pettinato, 2001; Di Tella et al., 2003; Smith, 2003; Appleton and Song, 2008; Selim, 2008; Shields et al., 2009; Graham and Chattopadhyay, 2009; Tiefenbach and Kohlbacher, 2015; Reyes-García et al., 2016; Silver et al., 2017).

However, previous literature has identified several groups of countries that show distinct life satisfaction patterns, such as the countries from Latin America and the Caribbean, the East Asian countries and the ex-communist countries. Importantly, it must be noted that the methods used by previous research rely on the *a priori* definition of such groups of countries and researches limit to refer to cultural differences and, at most, to point out one particular characteristic of each group.

In light of the previous findings, in this paper we aim at building a multidimensional classification of countries based on life satisfaction and its main determinants. We use data from the *World Values Survey* and the *European Values Study* for 103 countries around the world in the period 1994–2014.

We run a hierarchical cluster analysis, which is a numerical technique suitable for identifying groups of similar countries in terms of a set of clustering variables. We identify five groups of countries with distinct patterns in the joint distribution of life satisfaction and its determinants.

The rest of the paper is structured as follows. In Section 2 we review the academic literature on the determinants of average life evaluations across countries and introduce some previously proposed country groupings. Section 3 presents the data and the clustering method used in our piece of research. In Section 4 we present the resulting international life satisfaction taxonomy and we compare it with previous studies. Section 5 concludes and identifies the strengths and limitations of our approach and possible avenues for future work.

2 Previous research

2.1 *Cross-country determinants of life satisfaction*

Most recent models aiming at explaining societal differences in average life evaluations have been proposed within the *World Happiness Report*. In particular, Helliwell and Wang (2013) established a parsimonious model that includes six arguably ‘universal’ factors: income, health, social support, freedom, generosity and corruption. The model explains three quarters of the observed variation in average life evaluations across 150 countries and along multiple years. We will adopt this model as a reference and in what follows we briefly discuss the theory and evidence on the determinants of average life evaluations across countries.

A large economic tradition assumed a positive relationship between *income* and well-being, arguing that the former allows individuals to satisfy their needs (Rojas, 2009) or, more generally, their wants according to their preferences (OECD, 2007). The capability approach pointed out that income enlarges the capacity of individuals to be and to do what they have reasons to value (Sen, 2000). In fact, it has been found that there is a strong positive relationship between cross-country levels of GDP per capita and average life evaluations (Deaton, 2008; Stevenson and Wolfers, 2008). However, evidence suggests that the relationship is larger among poorer countries and the sign of the relationship may actually reverse at around 30,000 USD (Proto and Rustichini, 2013).

Some evidence suggests that there are economic factors other than income underlying societal differences in average life evaluations. Regarding the effects of unemployment and inflation, evidence is mostly focused on developed countries and is not conclusive for Europe regarding unemployment (Dolan et al., 2008). Gándelman and Hernández-Murillo (2009), whose sample includes developing countries, find a significant association between unemployment and inflation rates and average life evaluations, although they do not control for any other economic factor. Bjørnskov et al. (2008), controlling for several personal characteristics and aggregate factors, find that the effects of national income and the unemployment and inflation rates are not significant. On the other hand, they find that the business climate and the openness of the economy are the only relevant aggregate economic factors.¹ This evidence is promising, although so far the only economic factor that is well-established in the literature is national income.

Some evidence suggests that income inequality has a negative effect on life evaluations (Alesina et al., 2004; Helliwell et al., 2016). Diener et al. (1995) argue that more individuals may be able to achieve their goals in countries where there is more equality and, moreover, it is likely that where inequalities are high, issues of equity and social justice arise. Knack and Keefer (1997) show that income inequality is associated with several social problems that negatively impact life evaluations, such as the bad functioning of institutions, antisocial behaviour, insecurity and low interpersonal trust. On the other hand, Alesina et al. (2004) showed that the relationship between income inequality and life evaluations depends on the perception of social mobility and there are significant differences across Europe and US. Similarly, other studies highlight the ambivalence of income inequality in terms of life evaluations depending on the opportunities open to individuals (Dolan et al., 2008; Bjørnskov et al., 2008). In general, evidence on the negative effect of income inequality on life evaluations is not conclusive (Graham and Felton, 2006; Bjørnskov et al., 2008; Helliwell et al., 2016).

The average *health* status of the population is a factor widely considered because of the importance of health for individuals (Dolan et al., 2008), although when life expectancy at birth is used as a proxy it hardly contributes to explaining average happiness once income per capita is taken into account (Stutzer and Frey, 2012). Nevertheless, adjusting life expectancy at birth by the quality of life (healthy life expectancy at birth), it results to be an important factor (Layard et al., 2012; Helliwell and Wang, 2013). Bjørnskov et al. (2008) find that the infant mortality rate is significantly associated with the life satisfaction of individuals. Ideally, cross-country analyses would have to consider mental health measures (Clark et al., 2017), although there are important data constraints in this regard.

Social support refers to a range of different kinds of help and advantages facilitated by people's social ties (Scrivens and Smith, 2013). Moreover, it is argued that close interpersonal relations are important not only in terms of emotional and material support but more broadly in terms of sharing daily life (Rojas, 2018). In this regard, genuine – non-instrumental– interpersonal relationships give rise to intrinsically valued 'relational goods' (Bruni, 2008). The *social support and sociality* factor is extensively considered in both within and cross-country analyses (Blanchflower and Oswald, 2011; Layard et al., 2012; Clark et al., 2017). The condition of basic psychological need of positive social relationships that provide support and meaning is currently one of the most widely accepted and influential thesis in this literature (Diener and Seligman, 2004; Helliwell et al., 2017). However, there are some differences across countries. For instance, the effect of loneliness depends on cultural factors (Schumaker et al., 1993).

Freedom – the ability to pursue one's personal goals and desires– is strongly associated with average life evaluations at the cross-country level (Inglehart et al., 2008; Layard et al., 2012; Helliwell and Wang, 2013). Inglehart et al. (2008) find that among the main drivers of self-perceived freedom are income growth, social tolerance (towards sexual orientation and gender equality) and democratisation. Human rights and gender equality are associated with average life evaluations (Diener et al., 1995). In contrast, the effect of democracy has fail to be significant in studies that include both developed and developing countries, where stability and the quality of public services seem to be more important (Bjørnskov et al., 2008; Helliwell et al., 2017). In this regard, the sense of freedom seems to be more important in richer countries (Inglehart et al., 2008). Interestingly, the results of Helliwell et al. (2017) suggest that the impact of freedom on average life evaluations is channelled through positive affective experiences.

Altruism is an attitude that entails a cost (Layard et al., 2012), although evidence shows that giving support to other people might be at least as important as receiving social support in terms of coping with stressors, longevity and subjective well-being (Diener and Seligman, 2004; Helliwell et al., 2017). Altruistic behaviours, such as volunteering, may enhance subjective well-being either for intrinsic or extrinsic reasons, but evidence shows that the effect is larger when it is intrinsically motivated (Meier and Stutzer, 2008).² In fact, some evidence suggests that volunteering enhances subjective well-being only if it is carried out for other-regarding motivations, which is a particular kind of intrinsic motivation (Becchetti et al., 2017). Research at the cross-country level is scarce yet. Helliwell and Wang (2013) find that generosity (measured by an index of donors to charities that accounts for the income level of the country) is positive and significantly associated with average life evaluations.

The measure of corruption is a proxy for *social trust and good governance*, defined as the overall trustworthiness of public institutions and the quality of delivery of public

services (Helliwell et al., 2017). Social trust and good governance are usually interpreted as the outcome of collective social capital: the networks together with shared norms, values and understandings that facilitate cooperation and trust between all members of a community or the larger society (Scrivens and Smith, 2013). Knack and Keefer (1997), focusing on the economic effects of civic cooperation, note that these norms act as constraints on narrow self-interest increasing inter-personal trust and, consequently, reducing transaction costs and realising resources for more valuable purposes. The positive impact of civic norms and behaviours may likely go beyond the economic sphere as trust replaces suspicion and fear (Helliwell, 2003) and certain civic behaviours, such as politeness, constitute relational goods that are valuable in their own right.³ The extent of corruption is negatively correlated with average life evaluations (Layard et al., 2012; Helliwell and Wang, 2013) while interpersonal trust is positively correlated with them (Helliwell, 2003).

Beyond the absence of corruption, good governance also concerns the “reliability and responsiveness [of institutions] in their design and delivery of services” (Helliwell et al., 2017, p.34). For instance, the education and health systems may reflect the ability of a society to care for people. In this regard, strong welfare States and public spending have been found to enhance average happiness (Blanchflower and Oswald, 2011), at least for some important social groups (Bjørnskov et al., 2008). Confidence in public institutions is also associated with average life evaluations (Helliwell et al., 2017). One important aspect of the public life domain is social stability (Diener and Seligman, 2004). In this regard, according to Bjørnskov et al. (2008) the positive association between the bicameral political system and life evaluations may respond to the fact that such political system facilitates stability with its checks and balances.

Finally, there is an important factor left in the error term of the model that seemingly exerts an impact on average life evaluations: *culture*. Thus, individualistic societies – where the goals and desires of individuals have priority over those of the group– show higher levels of happiness, arguably because “individualists are likely to place more value on personal well-being and thus seek SWB [subjective well-being] to a greater extent” (Diener et al., 1995, p.853). Another cultural aspect that seems to influence average levels of happiness is the set of norms governing the experience and expression of emotions. In some cultures life satisfaction and pleasant affects are strongly desirable, whereas in other it is emphasised the relative appropriateness of unpleasant emotions. Cultures that view pleasant emotions as positive and desirable tend to be happier (e.g., Anglo-Saxon and Latin American countries), whereas cultures that perceive unpleasant emotions as normatively desirable experience lower subjective well-being (e.g., East Asian countries) (Diener and Suh, 1999).

2.2 *Interrelations among determinants of life satisfaction*

The determinants of life satisfaction are strongly interrelated. Firstly, evidence suggest that there may be a trade-off between income growth and other sources of life satisfaction, such as family life (marital breakdown rates) and social relationships (time spent with family members and friends, and social participation), which are key sources of social support and sociality (Blanchflower and Oswald, 2004; Bartolini et al., 2013; Bartolini and Sarracino, 2015). Moreover, economic growth may worsen the community and the larger society circumstances, reducing interpersonal trust and confidence in institutions (Bartolini et al., 2013; Bartolini and Sarracino, 2015; Mikucka et al., 2017).

Besides, it has been pointed out that there may be a trade-off between social ties and freedom to make life choices (OECD, 2001). As a result, greater freedom might come at the expense of available social support (Diener et al., 1995; Helliwell et al., 2017). In this regard, Alesina and Giuliano (2010) show that the strength of family ties —arguably one of the main sources of social support and sociality – is negatively associated with women's labour force participation – an indicator of gender equality and thus individual freedom. On the other hand, friendship ties are another source of social support, arguably more favourable for individual freedom (Pugno and Verme, 2012).

Regarding trust, it has been linked to societies with greater individual freedom. Thus, it has been found that economic freedom –as measured by an indicator of the legal structure and security of property rights – enhances social trust (Berggren and Jordahl, 2006).⁴ Furthermore, it has been also consistently found that collectivism is negatively associated with social trust (Macy and Sato, 2002; van Hoorn, 2015). This relationship is not obvious as we may expect societies with stronger group obligations – and therefore less freedom to make life choices– to be more cooperative and trustworthy. Macy and Sato (2002) suggest that interpersonal trust does not only depend on the strength of group obligations but also on the level of social and spatial mobility that requires individuals to learn how to interact effectively with strangers. According to these authors, too much mobility may diminish interpersonal trust, but until a certain threshold it may actually enhance it. However, there may be differences among collectivistic cultures. In this regard, Delhey and Newton (2005) point out that Confucianism (a canonical collectivistic culture) may enhance interpersonal trust; hence they classify China as a high-trust society, and Japan and South Korea as medium-trust societies.

On the other hand, trust is more clearly associated with altruism. Falk et al. (2018) show that countries with higher levels of social trust tend to show also higher levels of altruism, and that both aspects “describe positive behavioural dispositions toward others” (p.1665).

Finally, culture also interacts with the rest of variables. For instance, the individualism-collectivism spectrum, which reflects the weight of personal vs. group goals and procedures, may underlie the relationship between freedom and social support. In this regard, it is argued that in individualistic cultures there is more personal freedom, whereas in non-individualistic cultures there might be greater feelings of social support (Diener et al., 1995). Note, however, that culture also moderates the impact of loneliness (lack of sociality) on life satisfaction (Schumaker et al., 1993). Individualism is also associated with the income level (Diener and Seligman, 2004). In this regard, Falk et al. (2018) show that individualism and weak family ties are associated with the prevalence of a particular time preference: patience, that enhance economic growth (Falk et al., 2018).

2.3 Country groupings by life satisfaction patterns

Previous literature has identified several groups of countries showing distinct life satisfaction patterns: (1) Latin America and the Caribbean, (2) East Asia, (3) ex-communists, (4) North America, Australia and New Zealand and (5) Scandinavia.

Countries from Latin America and the Caribbean show, on average, higher levels of happiness than those their life circumstances would predict (Inglehart et al., 2008;

Helliwell and Wang, 2013). It is argued that in these countries warm social relationships are particularly important for happiness and that they are indeed stronger than in other regions (Beytía, 2016; Rojas, 2018). Moreover, Inglehart et al. (2008) found that countries from Latin America and the Caribbean show on average higher levels of self-perceived freedom than countries from other regions. Diener and Suh (1999) point out that there is a tendency in these countries to view pleasant emotions and satisfaction with life as desirable, and unpleasant emotions as relatively inappropriate.

Ex-communist countries have been considered a group in terms of happiness (Inglehart et al., 2008; Bjørnskov et al., 2008), albeit this group may be vanishing nowadays (Guriev and Melnikov, 2018). These countries have shown, on average, lower levels of happiness than those their life circumstances would have predicted. Bjørnskov et al. (2008) argue that the collapse of their political and economic systems sunk these countries into a long-lasting period of instability that may not be fully captured by standard indicators. Importantly, Helliwell (2003) distinguishes between countries of Central and Eastern Europe and countries of the former Soviet Union since the former seemed to be converging faster in terms of life satisfaction with the rest of countries. Moreover, Guriev and Melnikov (2018) find that former Soviet republics from Central Asia are performing better in terms of happiness than other members of the Commonwealth of Independent States.

Another well-established group comprises the countries of East and Southeast Asia with a strong Confucian culture: China, Hong Kong, Vietnam, Singapore, Korea and Japan (Ng, 2002; Weiming, 2019). These countries present, on average, levels of happiness below those their life circumstances would predict (Helliwell and Wang, 2013). It is argued that in these countries there is relatively more acceptance of unpleasant emotions and relatively less acceptance of pleasant ones (Diener and Suh, 1999). Moreover, in China the ideal level of life satisfaction is neutrality –neither satisfied nor dissatisfied–, whereas in other countries respondents view the ideal as strong satisfaction with life (Diener and Suh, 1999). Ng (2002) points out that in countries from East Asia there may be excessive competitiveness and conformity.

Finally, countries from both Scandinavia (Helliwell, 2003) and NAANZ (North America, Australia and New Zealand) (Helliwell and Wang, 2013) show, on average, higher life evaluations than those their life circumstances would predict. Scandinavian countries stand out because they show better life circumstances than other OECD countries. Regarding NAANZ, the reason may partially lie in the existence of strong norms encouraging the experience and expression of positive emotions and satisfaction with life (Diener and Suh, 1999).

3 Methods

Previous discussion encourages us to further studying the joint distribution of life satisfaction and its determinants across countries. Cluster analysis is a set of numerical techniques that are suitable for classifying a sample of heterogeneous countries in a limited number of groups, each of which is internally homogeneous in terms of the similarities between the countries that comprise it on a range of indicators that partially captures the multidimensional nature of life satisfaction.

3.1 *Variables and data*

In this paper we use data from the Integrated Values Survey 1981–2014 (IVS), which merges data from the World Values Survey (WVS, 2015) and the European Values Study (EVS, 2015). The latter have collected data on life satisfaction and most of its determinants since 1981 for nationally representative samples of more than a hundred countries around the world. A traditional caveat with the IVS family of surveys was that low and middle income countries were under-represented. However, in the last rounds of the World Values Survey this limitation has been mitigated.

Cluster analysis entails a number of somehow discretionary decisions. The first concerns the number of clustering variables to be used in the analysis. In terms of a reasonable relationship between the sample size and the number of clustering indicators, we apply Formann's (1984) rule that recommends a minimum sample size of 2^n , where n equals the number of clustering variables. In our case, with a sample of 103 countries, the maximum n is equal to six. We may focus on life satisfaction and the five variables that have been consistently found as important determinants of the distribution of average life evaluations across countries: income, health, social support and sociality, freedom, and social trust and good governance.

Table 1 shows the different variables considered in our analysis, the indicators and the method of construction. All indicators but the GDP per capita are based on individual level data collected along several years in different countries. We aggregate individual-level data to get country-level data using the weights provided by the survey. To make them more reliable and cancel out transitory effects of economic fluctuations, which occurrence varies across countries, we compute averages by country using the IVS rounds carried out between 1994 and 2014.⁵ The value of income corresponds with the average GDP per capita of the years in which each country participated in the survey weighted by the corresponding sample size.

Regarding life satisfaction, we use the average response to the question:

“All things considered, how satisfied are you with your life as a whole these days? Using this card on which 1 means you are ‘completely dissatisfied’ and 10 means you are ‘completely satisfied’ where would you put your satisfaction with your life as a whole?”

The income variable is expressed in logarithmic terms because, despite certain caveats (Proto and Rustichini, 2013), the relationship between the national levels of income and life evaluations seemingly responds to this functional form (Deaton, 2008). Moreover, in cluster analysis highly skewed variables, as it is the case of GDP per capita in a worldwide sample, are usually transformed by taking the natural logarithm (e.g., Onda et al., 2014). In the next section we comment how this transformation affects our results.

Regarding health, we use a 5-point scale measure of self-reported health, which responses range from 1 (‘very good’) to 5 (‘very poor’). We reverse the response scale for the ease of interpretation. This measure is widely supported by the specialised literature as a measure of health within countries because it has been found to be a good predictor of morbidity and mortality (Cislaghi and Cislaghi, 2019), and also in cross-country health research as a more holistic measure of mental and physical health (OECD, 2019), although some caveats have been raised regarding the comparability of the responses across countries, especially across developed and developing countries (Kuhn et al., 2006).

Table 1 Classification variables

<i>Classification variables</i>	<i>Indicators [range]</i>	<i>Sources</i>	<i>Methods of construction</i>	<i>Range</i>
I. Life satisfaction	1.1. Self-reported life satisfaction	WVS (2015) and EVS (2015)	Weighted average by country	1–10
II. Income	2.1. ln GDP per capita PPP (constant 2011 int. \$)	World Bank (2018)	Natural logarithm of the weighted average by country	
III. Health	3.1. Self-reported health	WVS (2015) and EVS (2015)	Weighted average by country	1–5
IV. Social ties	4.1. Family is considered very important [0–1]	WVS (2015) and EVS (2015)	Sum of both indicators (each computed as a weighted average by country)	0–2
	4.2. Friends are considered very important [0–1]	WVS (2015) and EVS (2015)		
V. Freedom	6.1. Self-reported freedom of choice	WVS (2015) and EVS (2015)	Weighted average by country	1–10
VI. Social trust	5.1. Interpersonal trust [0–1]	WVS (2015) and EVS (2015)	Sum of the three indicators (each computed as a weighted average by country)	0–3
	5.2. Confidence in police [0–1]	WVS (2015) and EVS (2015)		
	5.3. Confidence in Parliament [0–1]	WVS (2015) and EVS (2015)		

Source: Authors

Social support and sociality are measured by means of an index of social ties. The index is constructed using two questions on the importance of family and friends for the respondent. The responses range from 1 (very important) to 4 (none at all important). We transform these 4-point scale variables into two binary variables indicating whether the specific source of social support is very important for the respondent. After computing the country averages of these two items, the index of social support is calculated as the sum of the resulting two proportions. The importance of the family has been previously used in measures of family ties (Alesina and Giuliano, 2010). Helliwell et al. (2017) point out the need for measures of social support and sociality assessing the contributions of both family and friends. In this regard, Pugno and Verme (2012) show the importance of including both sources of social support and sociality.

Freedom is measured as the average response to the question:

“Some people feel they have completely free choice and control over their lives, while other people feel that what they do has no real effect on what happens to them. Please use this scale where 1 means ‘no choice at all’ and 10 means ‘a great deal of choice’ to indicate how much freedom of choice and control you feel you have over the way your life turns out.”

Following Helliwell et al. (2017), social trust and good governance are measured by means of an index of interpersonal trust and confidence in two key public institutions: police and parliament. The IVS measures generalised trust by means of a binary variable indicating whether, according to the respondent, most people can be trusted, whereas the confidence in the public institutions is measured by means of 4-point scale variables which responses range from 1 ('great deal of confidence') to 4 ('none at all'). We transform the latter into binary variables taking value one whenever the response was either "great deal of confidence" or "quite a lot of confidence" to resemble the response scale of generalised trust. After computing the country averages of those three items, the index of collective social capital is calculated as the sum of the resulting three proportions.

All in all, we have complete data for 103 countries that comprises 88% of the world population and covers all geographical and cultural regions. Table A1 in Appendix A shows the complete dataset and Table A2 shows the descriptive statistics of the six variables.

Importantly, one must examine the variables for substantial collinearity before the clustering process.⁶ We do not find evidence indicating problematic correlations between pairs of variables, although *freedom* and *life satisfaction* are the pair of variables with a higher correlation coefficient (see correlation matrix in Table A3).

3.2 Clustering method

We conduct a hierarchical cluster analysis using the Ward's (1963) method,⁷ computing the squared Euclidean distances between each element,⁸ and standardising the variables to correct differences in scale according to the 'range -1 to 1'.⁹ Regarding the number of country groups – that is, the number of clusters to retain from the data –, we use two different tools: the dendrogram and the agglomeration schedule. All computations are performed with R version 3.6.1 (R Core Team, 2018), using the packages *BBmisc* (Bischl et al., 2017) for data standardisation and *factoextra* for cluster analysis (Kassambara and Mundt, 2019).

Figure 1 shows the resulting dendrogram. The nodes represent clusters and the lengths of the stems (heights) represent the distances at which clusters are joined (Everitt et al., 2011). This graph provides guidance regarding the number of groups to retain, suggesting that a five-cluster solution is appropriate.

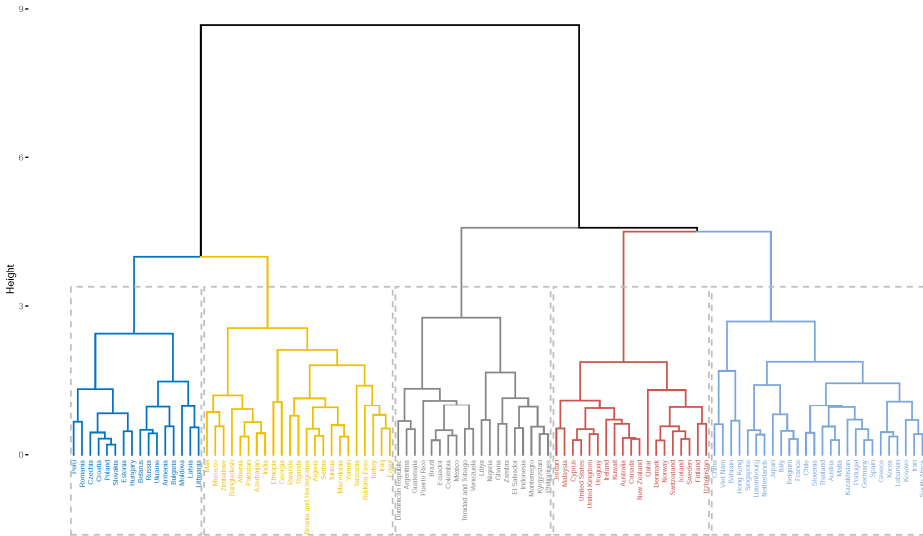
The agglomeration schedule displays the clusters combined at each stage and the distances at which clusters merge. By plotting those distances against the number of clusters we can identify a break or 'elbow', that is, where an additional combination of two clusters occurs at a greatly increased distance. The number of clusters prior to the merger is the most probable solution. The scree plot shows a distinct break due to the increase in distance when switching from a five to a four-cluster solution (Figure B1 in Appendix B).

Therefore, the two procedures (the dendrogram and the distances scree plot) suggest that a well-founded number of clusters is five.¹⁰

Before comparing the characteristics of these five clusters, it is worthwhile to distinguish which variables are more influential in discriminating between countries. According to a one-way ANOVA analysis (Table A4 in Appendix A), the variables with the greatest discriminating power are *freedom* and *life satisfaction*, followed by *health*.

By contrast, the variables with lowest relative importance in the classification are *social ties* and *income*.

Figure 1 Dendrogram: five-cluster solution (see online version for colours)



4 Results

4.1 Main features of the life satisfaction taxonomy

Figure 1 shows the dendrogram with the chosen one-level partition distinguished by colour. The first cluster (*C1*) includes 16 countries; the second (*C2*) is composed of 23 countries; the third (*C3*) includes 19; the fourth (*C4*) 19; and the fifth (*C5*) includes 26 countries.¹¹

The life satisfaction clusters are scattered across geographical regions, albeit we may distinguish some clear patterns (Figure 2): there is a concentration of *C1* countries in Central and Eastern Europe and the European side of the Commonwealth of Independent States; *C2* includes many countries from the Middle East, North Africa, and South Asia; there is a concentration of countries from Latin America and the Caribbean and the Southeast Asian islands in *C3*; *C4* includes countries from the Anglosphere and Scandinavia; and there is a concentration of countries in *C5* in Western (continental) Europe and Central and East Asia.

A more precise interpretation of the characteristics of the five clusters involves examining the cluster centroids (that is, the variables' average values of all countries in a given cluster). This procedure enables us to compare the average characteristics of each group of countries. Figure 3 graphically displays the relative value of the cluster centroids in terms of the maximum and minimum values of the different clustering variables.¹²

Figure 2 Geographic distribution of the life satisfaction country clusters (see online version for colours)

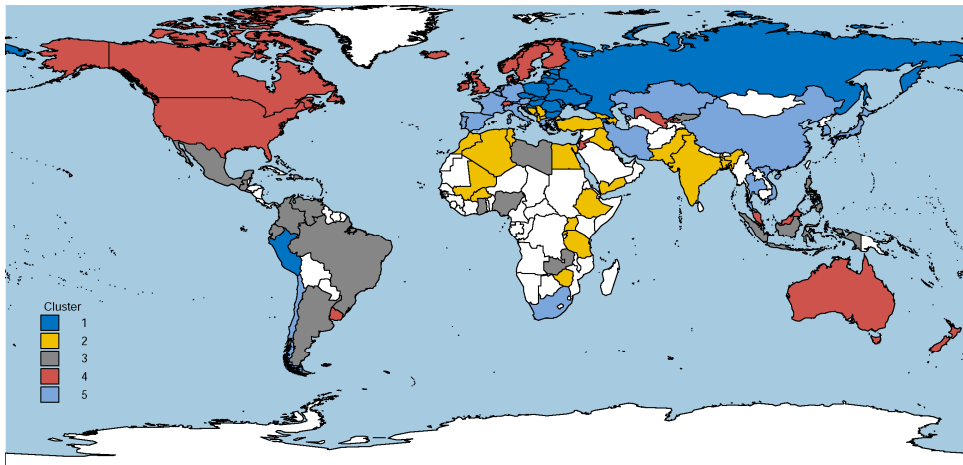


Figure 3 Variables' relative values across clusters (see online version for colours)



Cluster 1 consists of *relatively dissatisfied countries with low levels of freedom and social trust, and weak social ties*. In particular, these countries have the lowest levels of health, social ties and, along with C3, social trust. They also have the second poorest freedom and life satisfaction indicators. However, the income level is not as low as in C2 and C3.

Cluster 2 consists of *relatively dissatisfied countries characterised by combining low levels of income and freedom with strong social ties*. In particular, these countries have the lowest levels of income, freedom and life satisfaction. On the other hand, they fare better in terms of health, social trust and, especially, social ties.

Cluster 3 is composed of highly satisfied countries characterised by combining high levels of freedom and moderately strong social ties with low levels of social trust. Thus, these countries rank at the top in terms of the indicator of freedom; second in terms of indicators of health and life satisfaction; and third in terms of the indicator of social ties. On the other hand, these countries have on average very low levels of income and social trust.

Cluster 4 consists of *highly satisfied countries with high levels in all the life satisfaction determinants*. In fact, these countries rank first in terms of all the clustering indicators.

Finally, cluster 5 consists of moderately satisfied countries with high levels of income and social trust but relatively low levels of freedom and weak social ties. In particular, these countries rank second in terms of indicators of income and social trust. Moreover, they have the third highest levels of health and life satisfaction. They also rank third in terms of the indicator of freedom, although in this case the gap with respect to C4 and C3 is larger. On the other hand, these countries have the second lowest indicator of social ties.

4.2 *Comparison with other external classifications*

As noted above, previous literature on happiness has also distinguished five groups of countries: Latin America and the Caribbean; Confucian Asia; ex-communists; North America, Australia and New Zealand; and Scandinavia. Moreover, the World Happiness Report relies on a complete classification of countries that, although is not directly based on subjective well-being patterns but on geographical and cultural criteria, has the appealing of including generally more compact country groups than other classifications –except for the case of Scandinavia, which is included in Western Europe.

Adjusting the World Happiness Report classification, we consider an external classification constituted by 11 country groups:

- i Western Europe.
- ii Scandinavia.
- iii Central and Eastern Europe.
- iv The Commonwealth of Independent States (CIS).
- v South Asia.
- vi Southeast Asia.
- vii Confucian Asia.
- viii Latin America and the Caribbean (LAC).
- ix North America, Australia and New Zealand (NAANZ).
- x The Middle East and North Africa (MENA).
- xi Sub-Saharan Africa (SSA).

We now compare our proposed taxonomy (derived from a cluster analysis) with this external classification.

There is an external validity measure that assesses how closely a taxonomy reflects an external classification: the F-measure, which evaluates whether a hierarchical clustering contains, for each externally derived group, at least one cluster that is relatively pure – i.e., includes few countries of other groups– and includes most of the countries of that group (Tan et al., 2006). The overall F-measure is a weighted average of the F-measures associated with the different country groups. It takes value 1 when all countries of a given group are included in a cluster and no country of a different group is included in the same cluster. The F-measure tends to 0 when the fraction of a cluster that consists of countries of a given group or the fraction of countries of that group in the cluster tends to zero. We are not as much interested in the overall F-measure as in the F-measures associated with the different country groups, especially those that have been established on the basis of subjective well-being criteria.

Table 2 top panel shows that our taxonomy presents, in general, large F-measures. In particular, the F-measures associated with NAANZ, LAC, Scandinavia and South Asia, are greater than 0.75. The F-measures associated with Central and Eastern Europe, Confucian Asia, Western Europe and CIS are greater than 0.55.

The F-measures evaluate the overall taxonomy. Regarding the final five-cluster solution, it is worth assessing whether for each externally derived group there is any cluster with a high *recall* (i.e., it includes most of the objects of that group).¹³ As compared to the F-measures, *recall* does not penalise the possible lack of purity. The bottom panel of Table 2 shows that all countries from Scandinavia and NAANZ end up in *C4*. Similarly, all countries from South Asia end up in *C2*. Note that countries from these three groups are grouped together in the early stages of the hierarchical clustering procedure as shown by their associated F-measures.

Table 2 External validity measures for country clusters

	Western Europe	Scandinav.	Cent- East Europe	CIS	South Asia	Southeast Asia	Confucian Asia	LAC	NAANZ	MENA	SSA
<i>F-measure</i> ($F = 0.6$)	0.595	0.833	0.606	0.556	0.75	0.444	0.6	0.833	0.857	0.444	0.412
<i>Recall</i>											
<i>C1</i>	0	0	0.588	0.500	0	0	0	0.071	0	0	0
<i>C2</i>	0	0	0.235	0.200	1	0	0	0	0	0.500	0.636
<i>C3</i>	0	0	0.059	0.100	0	0.500	0	0.786	0	0.071	0.273
<i>C4</i>	0.267	1	0	0.100	0	0.250	0	0.071	1	0.214	0
<i>C5</i>	0.733	0	0.118	0.100	0	0.250	1	0.071	0	0.214	0.091

Data: WVS (2015), EVS (2015), and World Bank (2018). Notes: CIS: Commonwealth of Independent States; LAC: Latin America and the Caribbean; NAANZ: North America, Australia and New Zealand; MENA: Middle East and North Africa; SSA: Sub-Saharan Africa. External country classification based on Helliwell and Wang (2013), Helliwell (2003) regarding Scandinavia, and Ng (2002) and Weiming (2019) regarding Confucian Asia.

Source: Authors

Interestingly, all countries from Confucian Asia also end up in *C5*. In this case, the associated F-measure indicates that the fusion of a significant fraction of those countries occurs in the last stages of the clustering procedure. Figure 1 shows that China, Hong Kong and Vietnam are grouped together very early, whereas Singapore, Korea and Japan follow another path along with countries from other regions and only merge with the rest of Confucian countries at the end of the clustering procedure. Therefore, we can distinguish two subclusters. Table A1 shows that China, Hong Kong and Vietnam present higher levels of social trust and much weaker social ties than Singapore, Korea and Japan.

Most countries from LAC are grouped together in the early stages of the clustering procedure. However, there are three countries –Peru, Uruguay and Chile– that are not grouped together with the rest of Latin American countries. Peru, with low levels of both social ties and social trust, is included in *C1*. Uruguay, with higher levels of social trust, is included in *C4*. And Chile, which also has a higher level of social trust, is included in *C5*.¹⁴

Regarding the two ex-communist groups of countries –Central and Eastern Europe and CIS–, there are three countries from Eastern Europe (Latvia, Lithuania and Bulgaria) that are included in the CIS subcluster. Country level data displayed in Table A1 shows that countries from Central and Eastern Europe outperforms the CIS countries in most of the domains. Interestingly, countries from both subclusters present similar values on both social ties and social trust.

Importantly, ex-communist countries from the Balkans, Caucasus and Central Asia are scattered over the rest of the clusters. Thus, four Balkan countries (Albania, Bosnia and Herzegovina, Serbia and Macedonia) and two Caucasian countries (Azerbaijan and Georgia) are included in *C2*. One Balkan country (Montenegro) and one Central Asia former Soviet republic (Kyrgyzstan) are included in *C3*. Another country from Central Asia (Uzbekistan) is included in *C4*. And two Balkan countries (Slovenia and Kosovo) and one country from Central Asia (Kazakhstan) are part of *C5*.

Continental Western European countries concentrate in *C4* and, especially, *C5*, where more than 70% of them are included. Regarding Southeast Asia, Indonesia and Philippines are included in *C3*; Malaysia in *C4*; and Thailand in *C5*. Finally, over half of the countries from both MENA and Sub-Saharan Africa are included in *C2*. An important fraction of countries from MENA spread over *C4* and *C5*; whereas a relevant fraction of countries from Sub-Saharan Africa is included in *C3*.¹⁵

5 Conclusions

The average level of life satisfaction is an increasingly accepted indicator of well-being. Current interest in subjective well-being indicators is partly due to the evidence that economic growth is not necessarily linked to greater average happiness but it depends on other factors such as people's values, social relationships and the community circumstances. In general, the determinants of life satisfaction are common across countries, although there are some heterogeneities and the determinants are jointly determined along time. This encourages us to further studying the joint distribution of life satisfaction and its determinants across countries.

In particular we have conducted a hierarchical cluster analysis of 103 countries using the Ward's method. This analysis allows us classifying a sample of heterogeneous

countries in five groups, each of which is internally homogeneous in terms of the similarities in the joint distribution of six relevant variables: average levels of life satisfaction, income, health, social ties, freedom and social trust. The five groups of countries may be mainly characterised as follows:

Relatively dissatisfied countries with low levels of self-perceived freedom, social ties and social trust (C1): group that consists of the European ex-communist countries (except those from the Balkans) plus Peru.

Relatively dissatisfied countries with low levels of income and freedom but strong social ties (C2): group that consists of countries from North and East Africa, the Balkans, the Middle East and South Asia.

Highly satisfied countries that combine high levels of self-perceived freedom and moderately strong social ties with low levels of social trust (C3): this group includes Latin American countries but also countries such as Montenegro, Kyrgyzstan, Indonesia and Ghana.

Highly satisfied countries with high levels in all the variables (C4): group that mainly includes countries from Scandinavia and the Anglosphere (UK, Ireland, US, Canada, Australia and New Zealand), but also countries such as Cyprus, Uzbekistan, Kuwait, Malaysia and Uruguay.

Moderately satisfied countries with high levels of income and social trust but relatively low levels of freedom and weak social ties (C5): group that includes countries from Western (continental) Europe and East Asia and also countries such as Kazakhstan, Iran, Chile, and South Africa.

Regarding the main country groups identified by previous research, the taxonomy suggests that the group of ex-communist countries still constitutes a life satisfaction cluster, albeit it only includes countries from Central and Eastern Europe –not those from the Balkans, Caucasus and Central Asia. In general, our taxonomy seems to reflect the changes taking place in this group. The classification confirms that countries from Latin America and the Caribbean constitute a distinguishable cluster, although it suggests that it is less compact than usually assumed. Thus Peru, Uruguay and Chile are not classified with the rest of LAC countries. Finally, East Asian Confucian countries end up in the same cluster albeit we have pointed out that Singapore, Korea and Japan are in fact closer to countries from other regions than to China, Honk Kong and Vietnam, which show weaker social ties and higher levels of social trust than the former.

We think that the proposed life satisfaction taxonomy makes two major contributions. First, to the best of our knowledge it provides the first systematic classification of countries based on life satisfaction and its main determinants. Interestingly, it is in general consistent with previous partial findings, although it shows some noteworthy discrepancies with them. And second, the taxonomy constitutes an appealing multidimensional approach to well-being that can be used to assess countries based on the insights provided by happiness studies.

Classifications of this kind may be useful for policy makers suggesting which are the main strengths and weaknesses of a country, thus helping them to design better integrated policies.

However, the classification to be fully fruitful should be further developed as it may be considered only an initial piece of research. Regarding the method, further research is needed using alternative data sources and indicators (possibly using clustering methods well suited for dealing with high-dimensional data), and future research would benefit

from refinements in our understanding and measurement of the phenomena of interest for a large sample of countries.

Moreover, future research should focus on the correlates of the classification. On the one hand, we may have to look for differences between clusters in terms of the distribution of life satisfaction within countries, the evolution of such distribution and its determinants. Besides, it may be interesting to check whether the consequences of changes in average life satisfaction or any of its determinants vary across clusters. Thus, it is well-established that the inter-relations among the different factors affecting subjective well-being are very strong and, in many cases, there may be multiplying effects (Stiglitz et al., 2018). Recognising those inter-relations is therefore very important and cluster analysis may play an important role. In this regard, we note that Ireland and Iceland, which have been found strongly resilient to severe economic upheavals (Helliwell et al., 2015), are both included in C4, whereas Greece, which resiliency was found weaker in the same study, is included in C5. More evidence in this regard – along with a convincing explaining theory – is needed to derive further implications from the classification.

Besides, the taxonomy provides a snapshot of the life satisfaction phenomena, but we should also contemplate possible structural changes in the joint distribution of average life satisfaction and its determinants over time. In this regard, determinants of life satisfaction, though highly persistent, are susceptible of changes. Cluster analysis can be carried out in different time periods to compare the groups found in each period and to analyse the dynamics of each country in comparative terms.

Acknowledgements

Authors want to acknowledge financial support provided by the University of Cantabria through the program: “Convocatoria Proyectos Emergentes UC – 2016”. We want also to thank two anonymous referees for their comments on an earlier version of the paper. All remaining errors are our own responsibility.

References

- Alesina, A. and Giuliano, P. (2010) ‘The power of the family’, *Journal of Economic Growth*, Vol. 15, No. 2, pp.93–125.
- Alesina, A., Di Tella, R. and MacCulloch, R. (2004) ‘Inequality and happiness: are Europeans and Americans different?’, *Journal of Public Economics*, Vol. 88, No. 9, pp.2009–2042.
- Appleton, S. and Song, L. (2008) ‘Life satisfaction in urban China: components and determinants’, *World Development*, Vol. 36, No. 11, pp.2325–2340.
- Bartolini, S. and Sarracino, F. (2015) ‘The dark side of Chinese growth: declining social capital and well-being in times of economic boom’, *World Development*, Vol. 74, No. C, pp.333–351.
- Bartolini, S., Bilancini, E. and Pugno, M. (2013) ‘Did the decline in social connections depress Americans’ happiness?’, *Social Indicators Research*, Vol. 110, No. 3, pp.1033–1059.
- Becchetti, L., Corrado, L. and Conzo, P. (2017) ‘Sociability, altruism and well-being’, *Cambridge Journal of Economics*, Vol. 41, No. 2, pp.441–486.
- Berggren, N. and Jordahl, H. (2006) ‘Free to trust: economic freedom and social capital’, *Kyklos*, Vol. 59, No. 2, pp.141–169.

- Beytia, P. (2016) 'The singularity of Latin American patterns of happiness', in Rojas, M. (Ed.): *Handbook of Happiness Research in Latin America*, Springer, pp.17–30.
- Bischl, B., Lang, M., Bossek, J., Horn, D., Richter, J. and Surmann, D. (2017) *BBmisc: Miscellaneous Helper Functions for B. Bischl* (R package version 1.11).
- Bjørnskov, C. (2007) 'Determinants of generalized trust: a cross-country comparison', *Public Choice*, Vol. 130, No. 1, pp.1–21.
- Bjørnskov, C., Dreher, A. and Fischer, J. (2008) 'Cross-country determinants of life satisfaction: exploring different determinants across groups in society', *Social Choice and Welfare*, Vol. 30, No. 1, pp.119–173.
- Blanchflower, D. and Oswald, A. (2004) 'Well-being over time in Britain and the USA', *Journal of Public Economics*, Vol. 88, Nos. 7–8, pp.1359–1386.
- Blanchflower, D. and Oswald, A. (2008) 'Hypertension and happiness across nations', *Journal of Health Economics*, Vol. 27, No. 2, pp.218–233.
- Blanchflower, D. and Oswald, A. (2011) 'International happiness: a new view on the measure of performance', *Academy of Management Perspectives*, Vol. 25, No. 1, pp.6–22.
- Bloom, L.B. (2020) 'Ranked: The 20 happiest countries in the world', *Forbes*, Retrieved 26 October, 2020 from <https://www.forbes.com/sites/laurabegleybloom/2020/03/20/ranked-20-happiest-countries-2020/>
- Bruni, L. (2008) 'Back to aristotle? Happiness, eudaimonia, and relational goods', in Bruni, L., Comim, F. and Pugno, M. (Eds.): *Capabilities and Happiness*, Oxford University Press, Oxford.
- Caliński, T. and Harabasz, J. (1974) 'A dendrite method for cluster analysis', *Communications in Statistics*, Vol. 3, No. 1, pp.1–27.
- Cislaghi, B. and Cislaghi, C. (2019) 'Self-rated health as a valid indicator for health-equity analyses: evidence from the Italian health interview survey', *BMC Public Health*, Vol. 19, No. 1, p.533.
- Clark, A., Fleche, S., Layard, R., Powdthavee, N. and Ward, G. (2017) 'The key determinants of happiness and misery', in Helliwell, J., Layard, R. and Sachs, J. (Eds.): *World Happiness Report 2017*, Sustainable Development Solutions Network, pp.122–143.
- Deaton, A. (2008) 'Income, health, and well-being around the world: evidence from the gallup world poll', *Journal of Economic Perspectives*, Vol. 22, No. 2, pp.53–72.
- Delhey, J. and Newton, K. (2005) 'Predicting cross-national levels of social trust: global pattern or nordic exceptionalism?', *European Sociological Review*, Vol. 21, No. 4, pp.311–327.
- Di Tella, R., MacCulloch, R.J. and Oswald, A.J. (2003) 'The macroeconomics of happiness', *The Review of Economics and Statistics*, Vol. 85, No. 4, pp.809–827.
- Diener, E. and Seligman, M. (2004) 'Beyond money: toward an economy of well-being', *Psychological Science in the Public Interest*, Vol. 5, No. 1, pp.1–31.
- Diener, E. and Suh, E. (1999) 'National differences in subjective well-being', in Kahneman, D., Diener, E. and Schwarz, N. (Eds.): *Well-Being: The Foundations of Hedonic Psychology* Russell Sage Foundation, pp.434–450.
- Diener, E., Diener, M. and Diener, C. (1995) 'Factors predicting the subjective well-being of nations', *Journal of Personality and Social Psychology*, Vol. 69, No. 5, pp.851–864.
- Dolan, P., Peasgood, T. and White, M. (2008) 'Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being', *Journal of Economic Psychology*, Vol. 29, No. 1, pp.94–122.
- Everitt, B., Landau, S., Leese, M. and Stahl, D. (2011) *Cluster Analysis*, 5th ed., Wiley, Chichester, West Sussex, UK.
- EVS (2015) *European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008) (ZA4804 Data File Version 3.0.0 (2015-10-30))*, GESIS Data Archive, doi:10.4232/1.12253.

- Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D. and Sunde, U. (2018) 'Global evidence on economic preferences', *The Quarterly Journal of Economics*, Vol. 133, No. 4, pp.1645–1692.
- Formann, A. (1984) *Latent Class Analyse Einführung in Die Theorie Und Anwendung [Latent Class Analysis—Introduction to Theory and Application]*, Beltz.
- Gándelman, N. and Hernández-Murillo, R. (2009) 'The impact of inflation and unemployment on subjective personal and country evaluations', *Federal Reserve Bank of St. Louis Review*, Vol. 91, No. 3, pp.107–126.
- Graham, C. and Chattopadhyay, S. (2009) 'Well-being and public attitudes in Afghanistan', *World Economy*, Vol. 10, No. 3, pp.105–146.
- Graham, C. and Felton, A. (2006) 'Inequality and happiness: insights from Latin America', *The Journal of Economic Inequality*, Vol. 4, No. 1, pp.107–122.
- Graham, C. and Pettinato, S. (2001) 'Happiness, markets, and democracy: Latin America in comparative perspective', *Journal of Happiness Studies*, Vol. 2, No. 3, pp.237–268.
- Guriev, S. and Melnikov, N. (2018) 'Happiness convergence in transition countries', *Journal of Comparative Economics*, Vol. 46, No. 3, pp.683–707.
- Helliwell, J. (2003) 'How's life? Combining individual and national variables to explain subjective well-being', *Economic Modelling*, Vol. 20, No. 2, pp.331–360.
- Helliwell, J. and Wang, S. (2013) 'World happiness: trends, explanations and distribution', in Helliwell, J., Layard, R. and Sachs, J. (Eds.): *World Happiness Report 2013*, UN Sustainable Development Solutions Network, pp.8–37.
- Helliwell, J., Huang, H. and Wang, S. (2015) 'The geography of world happiness', in Helliwell, J., Layard, R. and Sachs, J. (Eds.): *World Happiness Report 2015*, Sustainable Development Solutions Network, pp.12–41.
- Helliwell, J., Huang, H. and Wang, S. (2016) 'The distribution of world happiness', in Helliwell, J., Layard, R. and Sachs, J. (Eds.): *World Happiness Report 2016 Update (Volume I)*, Sustainable Development Solutions Network, pp.8–48.
- Helliwell, J., Huang, H. and Wang, S. (2017) 'The social foundations of world happiness', in Helliwell, J., Layard, R. and Sachs, J. (Eds.): *World Happiness Report 2017*, Sustainable Development Solutions Network, pp.8–47.
- Helliwell, J., Huang, H. and Wang, S. (2019) 'Changing world happiness', in Helliwell, J., Layard, R. and Sachs, J. (Eds.): *World Happiness Report 2019*, Sustainable Development Solutions Network, pp.11–46.
- Inglehart, R., Foa, R., Peterson, C. and Welzel, C. (2008) 'Development, freedom, and rising happiness: a global perspective (1981–2007)', *Perspectives on Psychological Science*, Vol. 3, No. 4, pp.264–285.
- Kassambara, A. and Mundt, F. (2019) *Factoextra: Extract and Visualize the Results of Multivariate Data Analyses* (R package version 1.0.6).
- Kaufman, L. and Rousseeuw, P.J. (1990) *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley, New York.
- Knack, S. and Keefer, P. (1997) 'Does social capital have an economic payoff? A cross-country investigation', *The Quarterly Journal of Economics*, Vol. 112, No. 4, pp.1251–1288.
- Kuhn, R., Rahman, O. and Menken, J. (2006) 'Survey measures of health: how well do self-reported and observed indicators measure health and predict mortality?', in Cohen, B. and Menken, J. (Eds.): *Aging in Sub-saharan Africa: Recommendation for Furthering Research*, National Academies Press, Washington, DC.
- Layard, R., Clark, A. and Senik, C. (2012) 'The causes of happiness and misery', in Helliwell, J., Layard, R. and Sachs, J. (Eds.): *World Happiness Report 2012*, United Nations, pp.58–89.
- Macy, M. and Sato, Y. (2002) 'Trust, cooperation, and market formation in the U.S. and Japan', *Proceedings of the National Academy of Sciences*, Vol. 99, Suppl 3, pp.7214–7220.

- Meier, S. and Stutzer, A. (2008) 'Is volunteering rewarding in itself?', *Economica*, Vol. 75, No. 297, pp.39–59.
- Mikucka, M., Sarracino, F. and Dubrow, J. (2017) 'When does economic growth improve life satisfaction? Multilevel analysis of the roles of social trust and income inequality in 46 countries, 1981–2012', *World Development*, Vol. 93, pp.447–459.
- Mooi, E. and Sarstedt, M. (2011) 'Cluster analysis', in Mooi, E. and Sarstedt, M. (Eds.): *A Concise Guide to Market Research*, Springer, Berlin Heidelberg, pp.237–284.
- Ng, Y-K. (2002) 'East-Asian happiness gap', *Pacific Economic Review*, Vol. 7, No. 1, pp.51–63.
- OECD (Ed.) (2001) *The Well-Being of Nations: The Role of Human and Social Capital*, OECD.
- OECD (2007) *Society at a Glance 2006: OECD Social Indicators*, OECD Publishing, Paris.
- OECD (2019) *Health at a Glance 2019: OECD Indicators*, OECD Publishing, Paris.
- Onda, K., Crocker, J., Kayser, G. and Bartram, J. (2014) 'Country clustering applied to the water and sanitation sector: a new tool with potential applications in research and policy', *International Journal of Hygiene and Environmental Health*, Vol. 217, No. 2, pp.379–385.
- Proto, E. and Rustichini, A. (2013) 'A reassessment of the relationship between GDP and life satisfaction', *PLOS ONE*, Vol. 8, No. 11, e79358.
- Pugno, M. and Verme, P. (2012) *Life Satisfaction, Social Capital and Bonding-Bridging Nexus* (Policy Research Working Paper No. 5945), The World Bank.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing (<https://www.R-project.org>)
- Reyes-García, V., Babigumira, R., Pyhälä, A., Wunder, S., Zorondo-Rodríguez, F. and Angelsen, A. (2016) 'Subjective wellbeing and income: empirical patterns in the rural developing world', *Journal of Happiness Studies*, Vol. 17, No. 2, pp.773–791.
- Rojas, M. (2009) 'Economía de la felicidad: Hallazgos relevantes respecto al ingreso y el bienestar', *El Trimestre Económico*, Vol. 76, No. 3, pp.537–573.
- Rojas, M. (2018) 'Happiness in Latin America has social foundations', in Helliwell, J., Layard, R. and Sachs, J. (Eds.): *World Happiness Report 2018*, Sustainable Development Solutions Network, pp.115–146.
- Schumaker, J.F., Shea, J.D., Monfries, M.M. and Groth-Marnat, G. (1993) 'Loneliness and life satisfaction in Japan and Australia', *The Journal of Psychology*, Vol. 127, No. 1, pp.65–71.
- Scrivens, K. and Smith, C. (2013) *Four Interpretations of Social Capital: An Agenda for Measurement* (OECD Statistics Working Paper No. 2013/6), OECD Publishing.
- Selim, S. (2008) 'Life satisfaction and happiness in Turkey', *Social Indicators Research*, Vol. 88, No. 3, pp.531–562.
- Sen, A. (2000) *Development as Freedom*, Anchor Books, New York.
- Shields, M.A., Wheatley Price, S. and Wooden, M. (2009) 'Life satisfaction and the economic and social characteristics of neighbourhoods', *Journal of Population Economics*, Vol. 22, No. 2, pp.421–443.
- Silver, H.C., Caudill, S.B. and Mixon, F.G. (2017) 'Human capital and life satisfaction in economic transition', *Economics of Transition and Institutional Change*, Vol. 25, No. 2, pp.165–184.
- Smith, K. (2003) 'Individual welfare in the soviet union', *Social Indicators Research*, Vol. 64, No. 1, pp.75–105.
- Stevenson, B. and Wolfers, J. (2008) 'Economic growth and subjective well-being: reassessing the Easterlin paradox', *Brookings Papers on Economic Activity*, Vol. 2008, No. 1, pp.1–87.
- Stiglitz, J., Fitoussi, J-P. and Durand, M. (2018) *Beyond GDP: Measuring What Counts for Economic and Social Performance*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264307292-en>
- Strauss, T. and von Maltitz, M.J. (2017) 'Generalising ward's method for use with manhattan distances', *PLOS ONE*, Vol. 12, No. 1, e0168288.

- Stutzer, A. and Frey, B. (2012) *Recent Developments in the Economics of Happiness: A Selective Overview* (IZA Discussion Paper No. 7078), Institute for the Study of Labor.
- Tan, P-N., Steinbach, M. and Kumar, V. (2006) *Introduction to Data Mining*, 1st ed. Pearson Addison Wesley.
- Tibshirani, R., Walther, G. and Hastie, T. (2001) ‘Estimating the number of clusters in a data set via the gap statistic’, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 63, No. 2, pp.411–423.
- Tiefenbach, T. and Kohlbacher, F. (2015) ‘Happiness in Japan in times of upheaval: empirical evidence from the national survey on lifestyle preferences’, *Journal of Happiness Studies*, Vol. 16, No. 2, pp.333–366.
- van Hoorn, A. (2015) ‘Individualist–collectivist culture and trust radius: a multilevel approach’, *Journal of Cross-Cultural Psychology*, Vol. 46, No. 2, pp.269–276.
- Ward, J. (1963) ‘Hierarchical grouping to optimize an objective function’, *Journal of the American Statistical Association*, Vol. 58, No. 301, pp.236–244.
- Weiming, T. (2019) ‘Confucianism’, *Encyclopædia Britannica*, Encyclopædia Britannica, Inc., <https://www.britannica.com/topic/Confucianism>
- World Bank (2018) *World Development Indicators*, The World Bank, Washington, DC.
- WVS (2015) *World Value Survey 1981-2014 Longitudinal Aggregate v.20150418, 2015*, World Values Survey Association (www.worldvaluessurvey.org), Aggregate File Producer: JDSystems.

Notes

¹The national income and the unemployment rate seem to be relevant only for some specific social groups.

²An intrinsic reward is a direct result of the activity or its outcome, whereas an extrinsic reward is an external benefit with respect to which the activity is instrumental.

³Seemingly considering this civic dimension of people’s social life, Clark et al. (2017) point out that “each of us [...] has a marked impact on the happiness of other people” (p.132).

⁴However, Bjørnskov (2007) does not find a significant effect using a larger sample and similar controls.

⁵Using 20-year averages implies a strong assumption of stability. In fact there is evidence of strong persistence in most of the selected variables. Bjørnskov et al. (2008) argue that the levels of life satisfaction within a country are relatively invariant over time as compared to their variability across countries. Regarding the life satisfaction determinants, Alesina and Giuliano (2010) attribute to second-generation immigrants in the US contemporaneous beliefs about the family in the country of origin of their parents arguing that cultural traits, such as those concerning family ties, are quite stable – and moreover partially inherited from previous generations. Knack and Keefer (1997) showed that the level of interpersonal trust is a highly stable social feature. Rojas (2018), implicitly assuming their stability, characterises subjective well-being and its social foundations in Latin America as compared to some Anglo-Saxon and Western European countries estimating average values of different indicators using data from all waves of the World Values Survey (1981–2014) and the Gallup World Poll (2006–2016). On the other hand, the World Happiness Report has documented significant changes in average life evaluations in recent years in several countries. Part of those changes is likely due to temporal economic, political and social stresses/easings. However, long term variations are also very likely as life circumstances may consistently worsen/improve – for instance, most countries from Central and Eastern Europe may be thought to be immersed in a long term improvement of their life conditions (Helliwell, 2003; Guriev and Melnikov, 2018). In this paper we are interested in identifying the current association structure between countries and left possible dynamics for further research. We may be safe as long as possible dynamics along the 20 years period it is not as high as to produce noise that obscures the identification of relevant clusters.

⁶If highly correlated variables are used for cluster analysis, specific aspects covered by these variables will be overrepresented in the outcome. Everitt et al. (2011) and Mooi and Sarstedt (2011) argue that absolute correlations above 0.9 are problematic.

⁷Relying on standardised data according to the 'range -1 to 1' and the Euclidean distance metric, the average linkage method results, as expected, in a higher cophenetic correlation (0.636) than the complete linkage and the Ward's methods (0.532 and 0.558 respectively). However, the average linkage method results in very unbalanced clusters that performs poorly in terms of the variance ratio criterion. According to several internal (variance ratio, silhouette coefficient and Dunn index) and external validity measures (see the discussion of the taxonomy in relation to an external classification in the next section) the Ward's method outperforms the complete linkage method. Results are available upon request.

⁸ Ward's method is associated with the Euclidean distance metric, although it has been shown that the Ward's algorithm can also be used with Manhattan distances, which has been called the least absolute error version of Ward's Method (Strauss and von Maltitz 2017). The least absolute error version of the method may be interesting when dealing with outliers. In fact, according to several external validity measures this version of the method is preferred when using GDP per capita as clustering variable instead of its logarithm, and the simple z standardisation to avoid several heterogeneous countries being merged together due to their very high levels of national income. Under the logarithmic transformation and the standardisation according to the 'range -1 to 1', the Euclidean distance metric is preferred. Results are available upon request.

⁹The 'range -1 to 1' is deemed to be preferable than other methods 'in most situations' (Mooi and Sarstedt 2011, p.247). We find that the clustering associated with the 'range -1 to 1' outperforms the one associated with the simple z standardisation in terms of the external validity measures used below. Results are available upon request.

¹⁰Both the dendrogram and the distances scree plot focus on the increase in the total within-cluster variation that occurs when two clusters are merged. Other methods that deal with the issue of the optimal number of groups to retain from the data considering the two aspects that characterise a good clustering: compactness (the aspect measured by the total within-cluster variation) and separation, suggest a different number of clusters. In particular, the variance ratio criterium (Caliński and Harabasz, 1974), the average silhouette method (Kaufman and Rousseeuw, 1990), and the gap statistic method (Tibshirani et al., 2001) suggest retaining two groups from the data (results available upon request). In this solution our proposed clusters 1 and 2 are joined in a single group that is clearly distinguishable from a second group that includes our proposed clusters 3 to 5. Despite, the two-cluster solution seems to be technically the optimal solution, evidence suggesting a five-cluster solution is not negligible, as shown by the visual methods relying on the distance at which clusters merge, and this second option is much more interesting from a policy point of view given that, as we would see, it depicts more specific characteristics that may have policy implications. Future work may try to validate the resulting classification using external criteria.

¹¹Table A1 includes the resulting clustering.

¹²Table A5 shows the absolute average value of the clustering variables by cluster and their standard deviations.

¹³It may be considered a good quality of the final five-cluster solution to group together most countries from a given geo-cultural region, especially in those cases in which the geo-cultural region has been shown to present special subjective well-being patterns.

¹⁴Data at the country level can be found in Table A1.

¹⁵Regarding the robustness of the classification to the clustering method, only 12 countries change clusters when we use K-means as clustering method. This number rises to 26 when we use the K-medoid clustering technique. However, overall results of the robustness analysis suggest that the main patterns found by the cluster analysis are robust (results available upon request).

Appendix A**Table A1** Dataset

<i>Country</i>	<i>Cluster</i>	<i>Life satisfaction</i>	<i>Income</i>	<i>Health</i>	<i>Social ties</i>	<i>Freedom</i>	<i>Social trust</i>
Armenia	1	4.98	5298.96	3.40	1.38	6.12	0.91
Belarus	1	5.21	11131.00	3.15	1.18	5.74	1.27
Bulgaria	1	5.31	11634.33	3.56	1.26	5.83	0.93
Croatia	1	6.60	17409.05	3.55	1.19	6.81	0.94
Czechia	1	6.96	23564.61	3.69	1.18	6.71	0.83
Estonia	1	6.06	20442.92	3.47	1.13	6.27	1.27
Hungary	1	5.99	20681.57	3.51	1.29	6.22	0.99
Latvia	1	5.59	14286.53	3.32	0.94	5.90	0.89
Lithuania	1	5.64	16480.51	3.41	0.86	6.51	0.75
Moldova	1	5.27	3351.60	3.21	1.06	6.49	0.85
Peru	1	6.75	7630.85	3.57	1.07	7.17	0.39
Poland	1	6.84	19166.82	3.54	1.24	6.52	0.98
Romania	1	5.93	16920.44	3.58	1.13	7.22	0.79
Russia	1	5.52	15679.31	3.22	1.16	6.04	0.87
Slovakia	1	6.51	19243.45	3.58	1.28	6.56	1.06
Ukraine	1	5.04	9925.70	3.16	1.25	5.78	0.88
Albania	2	5.54	7000.75	3.85	1.15	5.65	1.28
Algeria	2	5.97	11999.96	3.63	1.37	6.66	1.06
Azerbaijan	2	5.80	7660.26	3.68	1.25	6.17	1.37
Bangladesh	2	6.09	1596.75	3.55	1.20	5.98	1.54
Bosnia and Herzegovina	2	6.50	8395.58	3.79	1.40	6.62	1.04
Burkina Faso	2	5.56	50598.95	3.95	1.51	5.71	1.06
Egypt	2	5.42	9267.53	3.72	1.40	5.90	1.54
Ethiopia	2	4.94	856.99	3.79	1.71	6.14	0.80
Georgia	2	5.08	5636.71	3.39	1.67	6.16	1.00
India	2	5.66	3234.29	3.79	1.23	6.11	1.45
Iraq	2	5.03	11895.97	3.81	1.54	5.82	1.22
Macedonia	2	6.00	8288.27	3.89	1.57	6.50	0.90
Mali	2	6.07	1784.88	3.84	1.51	6.11	1.42
Morocco	2	5.68	6252.38	4.00	1.35	5.86	1.13
Pakistan	2	5.97	4791.52	3.87	1.19	5.85	1.09
Rwanda	2	5.74	22525.64	3.67	1.39	6.70	1.56
Serbia	2	6.05	11693.47	3.53	1.40	6.40	0.81
Tanzania	2	3.86	37159.80	3.80	1.24	5.81	1.54
Tunisia	2	5.58	31012.72	3.91	1.46	6.64	0.85
Turkey	2	6.32	13367.10	3.72	1.67	5.90	1.37

Table A1 Dataset (continued)

<i>Country</i>	<i>Cluster</i>	<i>Life satisfaction</i>	<i>Income</i>	<i>Health</i>	<i>Social ties</i>	<i>Freedom</i>	<i>Social trust</i>
Uganda	2	5.62	16783.44	3.91	1.66	6.78	1.40
Yemen	2	5.89	14981.63	3.89	1.46	6.40	0.84
Zimbabwe	2	5.06	2479.19	4.08	1.34	6.20	1.18
Argentina	3	7.37	16197.33	3.90	1.42	7.45	0.57
Brazil	3	7.75	14002.51	3.96	1.22	7.71	0.76
Colombia	3	8.32	9264.29	3.97	1.15	8.00	0.84
Dominican Republic	3	7.13	6787.85	3.91	1.21	7.37	0.52
Ecuador	3	7.92	10665.46	3.96	1.31	7.86	0.84
El Salvador	3	7.50	5634.01	3.83	1.62	7.50	0.95
Ghana	3	6.12	3165.79	4.23	1.39	7.19	1.23
Guatemala	3	7.93	6213.12	3.80	1.42	7.44	0.52
Indonesia	3	6.92	6709.76	3.88	1.55	7.35	1.35
Kyrgyzstan	3	6.77	2654.14	3.81	1.30	7.25	1.18
Libya	3	7.25	16371.90	4.34	1.60	7.30	0.84
Mexico	3	8.07	11844.33	3.78	1.23	8.08	0.83
Montenegro	3	7.19	5937.50	3.67	1.42	7.54	1.09
Nigeria	3	6.59	4146.57	4.32	1.58	7.06	0.88
Philippines	3	6.95	5121.92	3.67	1.33	7.10	1.27
Puerto Rico	3	8.25	68146.44	3.93	1.24	8.28	0.98
Trinidad and Tobago	3	7.40	22263.27	4.09	1.32	8.05	0.54
Venezuela	3	7.12	11187.37	4.05	1.45	8.14	0.79
Zambia	3	6.06	3766.81	3.89	1.40	7.21	0.98
Australia	4	7.37	36720.91	4.05	1.53	7.63	1.57
Canada	4	7.78	39417.54	4.20	1.55	7.63	1.61
Cyprus	4	7.27	34013.02	4.12	1.50	7.46	1.20
Denmark	4	8.31	43876.20	4.21	1.46	7.56	2.25
Finland	4	7.80	35498.12	3.79	1.37	7.54	1.92
Iceland	4	8.03	36353.98	4.12	1.50	7.75	1.90
Ireland	4	7.98	40914.67	4.35	1.58	7.26	1.54
Jordan	4	6.45	8526.36	4.16	1.50	7.39	1.63
Kuwait	4	7.21	70832.37	4.26	1.49	7.97	1.50
Malaysia	4	7.01	21084.65	4.20	1.42	7.42	1.52
New Zealand	4	7.75	23668.62	4.13	1.51	7.85	1.59
Norway	4	7.90	44182.00	4.13	1.52	7.47	2.23
Qatar	4	8.00	9346.04	4.38	1.72	7.93	2.04

Table A1 Dataset (continued)

<i>Country</i>	<i>Cluster</i>	<i>Life satisfaction</i>	<i>Income</i>	<i>Health</i>	<i>Social ties</i>	<i>Freedom</i>	<i>Social trust</i>
Sweden	4	7.66	36657.29	4.06	1.57	7.56	1.98
Switzerland	4	8.05	51345.01	4.15	1.44	7.32	1.82
UK	4	7.52	30445.50	3.97	1.54	7.18	1.34
US	4	7.49	41335.12	4.12	1.55	7.72	1.33
Uruguay	4	7.37	21062.06	3.97	1.44	7.51	1.20
Uzbekistan	4	7.89	13485.55	3.95	1.50	7.80	1.94
Austria	5	7.79	41040.69	3.94	1.34	7.24	1.43
Bahrain	5	6.79	43837.34	3.92	0.86	6.88	1.78
Belgium	5	7.61	37907.77	3.94	1.35	6.48	1.35
Chile	5	7.11	16268.27	3.78	1.21	7.17	1.06
China	5	6.77	7063.27	3.88	1.13	7.10	2.22
France	5	6.98	36001.71	3.92	1.41	6.44	1.37
Germany	5	7.22	38649.87	3.80	1.27	6.94	1.44
Greece	5	6.81	28472.33	4.12	1.29	6.89	0.93
Hong Kong	5	6.60	45722.87	3.66	1.06	6.57	1.75
Iran	5	6.40	15340.58	3.89	1.24	6.86	1.44
Italy	5	7.10	35927.89	3.81	1.30	6.07	1.38
Japan	5	6.78	34714.97	3.58	1.39	5.87	1.31
Kazakhstan	5	7.23	21276.93	3.70	1.41	7.03	1.56
Korea	5	6.36	23704.45	3.93	1.37	6.79	1.05
Kosovo	5	6.90	7530.63	3.80	1.16	6.56	1.47
Lebanon	5	6.50	14402.30	3.98	1.31	6.97	0.87
Luxembourg	5	7.89	87124.97	4.00	1.39	6.95	1.70
Malta	5	8.03	26341.28	3.89	1.32	7.45	1.50
Netherlands	5	7.76	38874.85	3.93	1.41	6.75	1.68
Portugal	5	6.89	26639.51	3.85	1.15	6.53	1.32
Singapore	5	7.02	66463.31	4.06	1.40	6.89	1.84
Slovenia	5	7.20	25861.81	3.62	1.31	7.36	0.95
South Africa	5	6.25	10715.72	4.09	1.29	6.96	1.36
Spain	5	7.03	29492.61	3.91	1.31	6.76	1.36
Thailand	5	7.37	37371.09	3.99	1.20	7.19	1.22
Viet Nam	5	6.86	3241.92	3.63	1.04	7.24	2.38

See Table 1 for definitions of the variables.

Table A2 Descriptive statistics of the dataset

	<i>Minimum</i>	<i>Maximum</i>	<i>Mean</i>	<i>Std. deviation</i>
Life satisfaction	3.857	8.319	6.702	0.958
Income	856.995	87124.971	21271.318	17014.235
Health	3.149	4.377	3.844	0.262
Social ties	0.859	1.719	1.35	0.175
Freedom	5.649	8.284	6.89	0.677
Social trust	0.387	2.378	1.251	0.412

See Table 1 for definitions of the variables.

Table A3 Correlation matrix

	<i>Life satisfaction</i>	<i>Income</i>	<i>Health</i>	<i>Social ties</i>	<i>Freedom</i>	<i>Social trust</i>
Life satisfaction	1					
Income	0.444	1				
Health	0.541	0.267	1			
Social ties	0.170	0.026	0.506	1		
Freedom	0.803	0.249	0.564	0.212	1	
Social trust	0.276	0.286	0.321	0.150	0.114	1

Note: see Table 1 for definitions of the variables.

Table A4 ANOVA output of the life satisfaction clusters

		<i>Sum of squares</i>	<i>Df.</i>	<i>Mean square</i>	<i>F value</i>	<i>Pr(>F)</i>
Life satisfaction	Between	63.15	4	15.79	50.81	<0.01
	Within	30.45	98	0.31		
Income	Between	30.74	4	7.69	13.21	<0.01
	Within	57.00	98	0.58		
Health	Between	4.49	4	1.12	43.61	<0.01
	Within	2.52	98	0.03		
Social ties	Between	1.36	4	0.34	18.91	<0.01
	Within	1.77	98	0.02		
Freedom	Between	33.9	4	8.48	64.46	<0.01
	Within	12.89	98	0.13		
Social trust	Between	9.03	4	2.26	26.62	<0.01
	Within	8.31	98	0.08		

See Table 1 for definitions of the variables.

Table A5 Means (Std. dev.) for classification variables by country clusters

Cluster	Life satisfaction		Income		Health		Social ties		Freedom		Social trust	
	Mean	Std. devs.	Mean	Std. devs.	Mean	Std. devs.	Mean	Std. devs.	Mean	Std. devs.	Mean	Std. devs.
1	5.886	0.670	14,552.98	5,866.45	3.432	0.173	1.162	0.132	6.368	0.464	0.912	0.203
2	5.627	0.563	12,576.69	12,301.22	3.785	0.160	1.421	0.166	6.176	0.348	1.193	0.259
3	7.294	0.651	12,109.49	14,560.79	3.947	0.189	1.378	0.138	7.572	0.385	0.894	0.254
4	7.623	0.441	33,619.21	15,166.27	4.122	0.140	1.509	0.074	7.576	0.220	1.689	0.319
5	7.049	0.474	30,768.81	18,360.39	3.870	0.143	1.266	0.135	6.844	0.371	1.450	0.359

See Table 1 for definitions of the variables.

Appendix B

Figure B1 Scree plot: distances against number of clusters

