THE EFFECT OF DISADVANTAGES ON LIFE SATISFACTION

Adolfo Morrone[1]

Istat, Italian National Institute of Statistics, Roma, Italy

Alfonso Piscitelli

Department of Political Science, University of Naples "Federico II", Naples, Italy, alfonso.piscitelli@unina.it

***Abstract**. Subjective well-being is an important aspect of people's quality of life and life satisfaction is one of the key variables to measure it. However for policies to be effective in maximising subjective well-being it is of paramount importance to study the factors that influence it. This paper use microdata from the 2013 wave of Multipurpose survey on Everyday life aspects which is a large scale cross sectional survey carried out by Istat annually since 1993 to study the effect of disadvantages on people's life satisfaction. Trees can be used for exploratory analysis in order to understand the relationships between the target variable and the predictors. We use a regression tree technique to show the effect of disadvantages on life satisfaction. The technique shows that in general the higher the level of disadvantages, the lower the life satisfaction but different combination of disadvantages have different effects on life satisfaction. The regression tree used allows to measure the importance of each factor.*

***Keywords**: Life satisfaction, Regression tree, Multiple disadvantages, Well-being, BES*

1. INTRODUCTION

"Measuring national well-being and societal progress in Italy is one of the challenges that the Italian National Institute of Statistics (Istat) is called to face. In recent years, Istat's attention towards this issue has taken the form of a number of activities aimed at strengthening the ability of official statistics to measure specific dimensions of well-being. Such initiatives include objective and subjective measures of individual well-being, the strengthening of environmental measures and accounts, and the adaptation of macroeconomic aggregates to provide distributional information and to overcome GDP limitations in general" (Giovannini and Rondinella, 2012).

---

[1] Corresponding author: Adolfo Morrone, e-mail: adolfo.morrone@istat.it

The most important of these initiatives is the BES (Benessere Equo Sostenibile) project that has been launched in 2010. The BES project has the ambition to measure "Equitable and Sustainable Well-being" and it is based on a framework that allows to analyse the levels, time trends and distribution of the various dimensions of well-being in Italy, in order to identify the strengths and weaknesses, as well as particular territorial imbalances or advantaged/disadvantaged social groups, considering them from an inter-generational point of view (sustainability) (Cnel and Istat, 2012).

The BES has published two reports (Istat, 2013; 2014) based on 129 statistical indicators to describe the average level and distribution of outcomes across twelve distinct dimensions that are thought to be of general importance to the well-being of people in Italy.

The twelve dimensions of well-being used in BES are broadly consistent with those put forward by the report of the Sen/Stiglitz/Fitoussi Commission (Stiglitz et al., 2009) and by other similar attempts to monitor well-being in individual countries.

One of the twelve dimensions of well-being considered in the BES framework is subjective well-being. Recent years have seen an explosion in the literature on the causes and correlates of subjective well-being, brought about by the increasing availability of data and evidence showing that self-reports of life satisfaction and current feelings are valid and consistent measures of people's sense of well-being.

It is widely acquired in literature that collecting information on subjective aspects is of high information and analytical value (OECD, 2013). Perceptions and evaluations affect the way people face life and take advantage of opportunities. Subjective indicators are useful complement to strictly objective indicators, because they allow evaluating the possible differences between what people report and what do objective indicators capture. Taking into account such indicators allows having a more detailed and complete overview, which is extremely useful when describing well-being.

Considering subjective well-being indicators when assessing individual and society's well being is important per se, because these indicators provide additional information relative to information on more objective dimensions. In addition, indicators of subjective well-being allow for better understanding of the relationship between subjective and objective well-being, and what life circumstances determine people's sense of well-being (Boarini et al., 2012, p. 6).

There is increasing interest in the measurement and use of subjective well-being for policy purposes (Dolan et al., 2011). However for policies to be effective in maximising subjective well-being it is of paramount importance to study the factors that influence it.

Using life satisfaction (LS) as a target variable and as key component of the overall well-being of the population this paper is built on the existing literature to bring new evidence on two aspects. First, using the BES framework to identify different objective and subjective disadvantages (e.g. being unemployed), it intends to assess the relative importance of them to LS. Second the paper intends to measure how different combinations of disadvantages can have a different impact on LS (e.g. being unemployed and in bad health can have a different impact than being poor and without friends). The second aspect is particularly important since "the consequences for quality of life of having multiple disadvantages far exceed the sum of their individual effects" (Stiglitz et al., 2009). This aspect is generally underestimated since "these cumulative effects require information on the joint distribution of the most salient features of quality of life" (Stiglitz et al., 2009).

The paper is organised as follows: in Section 2 we describe the experimental setting and the primary results about subjective well-being, and in Section 3 we introduce the statistical method used in our analysis. Section 4 illustrates the impact of multiple disadvantages on LS. The paper ends with summing up the results and describing possible future developments.

2. THE VARIABLES OF STUDY

To assess the effect of multiple disadvantages on LS we need to measure the different variables from a unique source. In this study we will use the Multipurpose survey on Everyday life aspects which is a large scale cross sectional survey carried out by Istat annually since 1993. The study is carried out on the 2013 wave, which encompasses several innovations introduced to improve the measurement of quality of life in the BES framework.

The survey is based on a face-to-face questionnaire, designed to collect questions that require the help of an interviewer and family questions, and a self-administered questionnaire designed to collect questions that can be answered directly by the individuals. This survey is currently used to monitor the state of Italian society on various social subjects. Since 2012 the survey has been improved to take into account aspects of quality of life not present in the past but crucial to measure well-being. In particular measures of life satisfaction, interpersonal trust and trust in institutions have been introduced to fill important gaps. Currently is ongoing a study to add an important section to measure the civicness of the country.

The two questionnaires ask respondents a broad set of questions on socio-economic background, civil engagement, and satisfaction of living standard among other domains. One important aspect of this survey is that it combines information on both subjective well-being, and on people's self-assessments of their objective determinants. The sample size of the survey is of around 50,000 respondents and is designed to be representative of the entire Italian population and to allow disaggregation of the main indicators at regional level and by socio-demographic characteristics.

The variables selected for analysis in this paper fall into four broad groups: (i) one measure of life satisfaction; (ii) demographic controls; (iii) individual well-being achievements that are proxies for the 11 domains of BES. The analysis will be focused on people aged 25-64 equal to a sample size of 25,382. This decision is due to the fact that factors effecting LS of children and young should be measured with a completely different set of variables (UNICEF, 2013) and the same applies for the elderly. In order to avoid this problem we decided to limit the study to a more homogenous population.

Life satisfaction, as suggested in the Oecd guidelines on measuring subjective well-being (OECD, 2013), is intended to capture the respondent's evaluative judgement of how their life is going while imposing a minimum level of respondent burden. It is envisaged as the primary measure of subjective well-being when a single measure is required (OECD, 2013). Life satisfaction is the result of a cognitive evaluation on the part of the individual rather than a description of current emotional state. One strength of measures of life evaluation is that they appear to tap the same underlying construct that people use when they pause and make a conscious decision about whether one course of action is preferable to another (Kahneman, 1999; Helliwell and Barrington-Leigh, 2010).

In our study, in accordance with the OECD and Eurostat guidelines, LS is measured using a 0-10 scale where 0 indicate, "not at all satisfied" and 10 "completely satisfied. In our data LS has a mean of 6.8 and the distribution is left skewed (skewness = -.946).

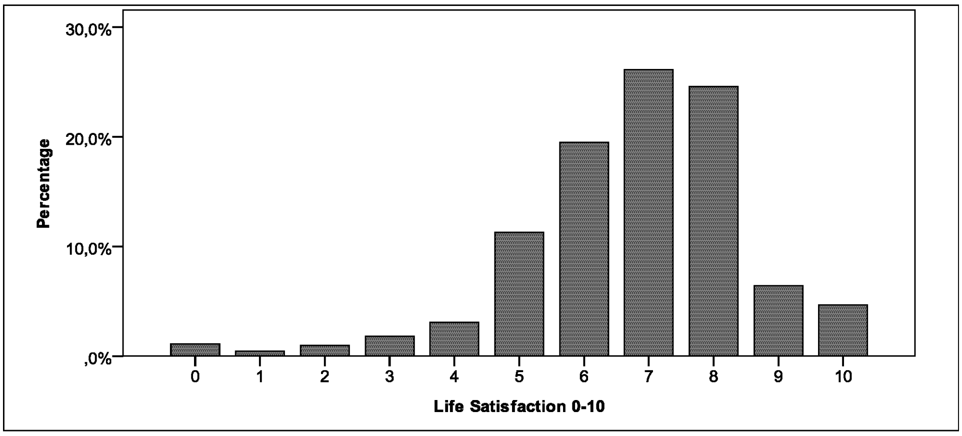

Figure 1. Distribution of life satisfaction variable

2.1 DATA COLLECTION

The variables that are used as determinants of LS are selected taking into account the BES framework. For each domain of the BES the most relevant candidate as determinants of LS are selected taking into account the wide literature on the topic (Boarini et al., 2012).

Since the aim of the paper is to study the effect of disadvantages on LS all variables are coded as dichotomous where the variable has value 1 if the responded is disadvantaged and 0 otherwise. For example the variable indicating perception of the health status is coded 1 for people with bad perceived health and 0 otherwise. This choice is pretty common in this kind of studies since is far easier to study the effect of a disadvantage on LS than studying the effect of a positive situation.

The variables considered and the motivations for our choices are briefly discussed in the following paragraph that draws most of the material from the extensive work of Boarini et al. (2012).

Health is considered in the literature one of the most important determinants of LS and there is an extensive literature on the relationship between health status and life satisfaction. Considering the BES framework and the variables available in our survey we selected three indicators. Two objective indicators, measuring the presence of chronical illnesses or limitation in daily activities give information on the state of health. Here the relationship goes both ways with high life satisfaction causing good health (Diener and Chan, 2011), but also with a strong causal relationship running from health to life satisfaction. For example, disability has a large and lasting causal impact on life satisfaction (Lucas, 2007). On the subjective side we use self-assessed health status that is known to have a large negative impact on life satisfaction (Dolan et al., 2008).

Education and training is measured considering educational attainment, computer skills and cultural participation that, in the BES framework, is considered a proxy of informal learning. On this domain the evidences from the literature are mixed with studies finding a strong correlation between measures of education and skills and life satisfaction across people (OECD, 2011) while other studies find that the relationship is weaker or null after considering income, health, and social trust (Helliwell, 2008).

The relationship between unemployment and life satisfaction has been extensively investigated. Unemployment is associated with a large negative impact on life satisfaction at the individual level. The size of the effect is considerably larger than that due to the associated fall in income, and persists when income is controlled for separately (Winkelman and Winkelman, 1998). Interestingly, it is being unemployed that has a negative impact on life satisfaction rather than having a job per se being associated with higher subjective well-being. Groups without a job, but that are not unemployed such as the retired, students, and full time parents, do not consistently report lower levels of life satisfaction (Blanchflower and Oswald, 2011). Research indicates that job satisfaction is significantly related to life satisfaction. However, in a nationally representative (US) longitudinal survey, Rode (2004) shows that core self-evaluations are significantly related to both job satisfaction and life satisfaction over time, and that the relationship between job satisfaction and life satisfaction is not significant after taking into account the effects of core self-evaluations and non-work satisfaction.

There is a general consensus with regard to the empirical relationship between income and life satisfaction on a cross-sectional basis at both the individual and cross-country level. Higher income is associated with a higher level of life satisfaction, but with diminishing returns as income increases. Sacks, Stevenson and Wolfers (2010), for example, find a constant relationship at both the individual and cross-country level, where a doubling of income is associated with 0.3 point increase in life satisfaction on a 0 to 10 scale. However for our study we lack a reliable income variable. For this reason we have to rely on proxies such as jobless households, people in households with low economic resources and people who do not own the house they live in.

Social connections, human contact and civic engagement are strongly associated with life satisfaction. In our study variables of satisfaction for family and friend relationship are used to measure the quality of relationships, which have a strong effect on LS (Helliwell, 2008). We also have variables of informal network support and general trust, which are also proved to be positively associated with life satisfaction (Helliwell and Wang, 2011). In this domain we also focus our attention on social and civic participation although there are mixed results on the strength of this relationship (see Frey and Stutzer, 2000; Dolan et al., 2008). Very closely correlated with social and civic engagement is governance, which is generally considered to be important to life satisfaction. This aspect is measured with a variable on trust in institution. A stronger predictor could be the level of corruption that has a strong negative correlation with average life satisfaction (Helliwell, 2008) but this variable is difficult to measure and is not available in our data.

Table 1. Summary of the Istat Multipurpose survey on Everyday life aspects variables used in the analysis

Domains	Objectives	Subjective
Health	Limitations in daily activities At least one chronical illness	Bad perceived health
Education and training	Max education level lower secondary Low computer skills Low cultural participation	
Work and life balance	People not employed	Low satisfaction for the job
Economic well-being	People in jobless households People in households with low economic resources People who do not own the house they live in	Low satisfaction for the economic situation
Social relationships	No social participation Meet friends never or very rarely Have people to count on	Low satisfaction for leisure time Low satisfaction for family relationships Low satisfaction for friends relationships Low generalised trust
Politics and Institutions	No civic and political participation	Low trust in institutions
Security	People noticing often elements of decline in the living place	Low sense of security walking at night alone
Landscape and cultural heritage		Landscape of the living place affected by evident signs of decline Low satisfaction for the environment of the living place
Environment	People living in areas with no parks	People perceiving high level of air pollution
R&D	Internet at least once a week	
Quality of services	People not connected to gas network People experiencing irregular distribution of water People living in areas with no waste sorting Very difficult to reach 3 or more essential services	

Living in an unsafe or deprived area is associated with a lower level of life satisfaction after controlling for one's own income (Dolan et al., 2008; Balestra and Sultan, 2012). In Multipurpose survey on Everyday life aspects there is a variable on the level of decline of the area where people live in and a variable on the sense of security (safe walking alone at night). We are interested in evaluating the effects of these variables in our general models to see if they have an additive effect with other disadvantages or if they are smoothed when controlled by other factors.

The last three domains encompass aspects that cannot be measured at individual level. Quality of services, environmental quality and research and development are geographic phenomenon, and we lack datasets where these information are linked with household/individual level data on life satisfaction. The only possibility is to use proxy variables on citizen's perceptions. On this point there are some evidences that subjective satisfaction with air pollution is correlated with actual air pollution (Johnstone and De Keulenaer, 2012) and that satisfaction on quality of health services is the best predictor of LS of people living in nursing homes (Ng et al., 2012).

LS was measured as ordinal variable ranging from 0 to 10. Although in a strict sense this is an ordinal scale, studies that have used the variable both as ordinal and as metric lead to approximately the same results (Diener et al., 1995) and therefore it can be treated as a cardinal one without significant biases (OECD, 2013; Ng, 1997). According to this idea the main international indicator used to compare the LS levels among countries is the average of the 0-10 scores (e.g. Eurostat, OECD, Istat and all the NSI that publish data on LS).

To analyse the impact of factors influencing LS is very common the use of regression-like techniques, but traditional parametric methods do not always yield satisfactory results. The computer intensive data-mining techniques based on recursion, averaging, and randomizations can uncover hidden structures in the data and suggest better predictive models.

In this work we think traditional parametric methods such as OLS regression are non-optimal for the analysis of the relationship between the LS and indicators measuring specific disadvantages. Therefore we try to detect the effect of multiple disadvantages on LS using Regression Tree (CART; Breiman et al., 1984). The main advantage of such technique is that, unlike classical regression techniques for which the relationship between the response and predictors is pre-specified (for example, straight line, quadratic) and the test is performed to prove or disprove the relationship, regression tree analysis assumes no such relationship. The aim of the paper is the description of the phenomenon using regression tree analysis to determine the correlates of LS, that is, the variables that best predict LS, producing a predictor ranking (also known as variable importance).

3. STATISTICAL METHOD

Regression-tree is a non-parametric procedure for predicting continuous dependent variable (denoted by $\boldsymbol{y}$) with categorical and/or continuous p predictor variables (each of them denoted by $\boldsymbol{x}$). The data can be hierarchically organized in a connected and oriented graph called tree and are partitioned into nodes on the basis of conditional binary predictor variables. Graphically, the tree is a representation of sets of linked nodes in which two nodes are connected by a specific relation (path). The starting-node is called root and the end-nodes “leaves” (or terminal nodes). It is primarily a method of constructing a set of decision rules on the predictor variables (Breiman et al., 1984; Verbyla 1987; Clark and Pregibon 1992).

Regression tree methods use a binary tree to recursively partition the predictor space into subsets in which the distribution of $\boldsymbol{y}$ is successively more homogenous (Chipman et al., 1998). Every tree structure must satisfy two properties: the shape property and the heap property. According to the former property each node has a fixed number t of child nodes (for $t = 2$ a binary tree is assumed). The heap property states that each node is greater than or equal to each of its children according to some comparison predicate which is fixed for the entire data structure. CART procedure produces a tree structure that satisfies both these properties. It derives conditional distribution of $\boldsymbol{y}$ given $\boldsymbol{x}$, where $\boldsymbol{X}$ represents a matrix of predictors ($\boldsymbol{X}=[x_1,x_2,x_3,\ldots,x_p]$). A decision tree $\boldsymbol{\Pi}$ with L terminal nodes (leaves) is used for communicating the classification/regression decision. A parameter' $\Phi=(\varphi_1, \varphi_2, \varphi_3,\ldots,\varphi_L)$ associates the parameter value $\Phi_i(i=1,2,3,\ldots,L)$ with the ith terminal node.

The partitioning procedure searches through all values of predictor variables (matrix of predictors) to find the variable $\boldsymbol{x}$ that provides best partition into child nodes. The best partition is the one that minimizes the weighted within-nodes variance. The distribution $f(y|\varphi_i)$ of $\boldsymbol{y}|\boldsymbol{x}$ represents the situation that $\boldsymbol{x}$ lies in the region corresponding to the ith terminal node. In other words, a suitable objective optimisation criterion (e.g. the within nodes sum of squares) is necessary to identify, at each partition step, which explanatory variable determines the best split of the parent group according to the optimisation criterion.

To each leave of the tree is assigned a response value (e.g. the average of $\boldsymbol{y}$). Each path of the tree gives the combination of predictors which is necessary to belong to a final leave. Trees can be used for exploratory analysis in order to understand the dependence relationships between the target variable and the predictors. The quality of the configuration can be evaluated in terms of mean squared error based on a learning sample, test sample or cross-validation. In the trade-off between bias and variance it is necessary to identify the proper tree model complexity. Large tree structures are inaccurate because of a large variance (too much sensitivity sample) whereas tree-models with too few leaves are inaccurate because of a large bias (not enough flexibility). CART suggests to grow the maximal expanded tree, and then it finds a sequence of nested pruned trees (cutting off at each step the weakest link in terms of a cost-complexity perspective) (Therneau and Atkinson, 1997). Finally CART selects the best-pruned tree on the basis of the cross-validation or test sample estimate of the mean squared error. It is worth nothing that through the paper we use the pruning technique not to make predictions, but merely to avoid interpreting a huge tree-based structure.

Regression tree analysis is effective in uncovering structure in data with hierarchical or non-additive variables. Because no a priori assumptions are made about the nature of the relationships among the response and predictor variables, regression tree analysis allows for the possibility of interactions and non-linearity among variables (Moore et al., 1991). Splitting rules of regression tree analysis enable mapping the predictors with the greatest influence on distributions, providing greater insight into the spatial influence of the predictors (Iverson and Prasad, 1998). Yet there are certain disadvantages with regression tree analysis compared to conventional regression modelling: (a) simple linear functions are highly approximated; (b) the output can be a discontinuous, depending on the threshold established in the regression trees; (c) the output can be unstable that is, small changes in data can produce highly divergent trees; and (d) when there are both continuous and categorical predictors, the former category of variables can dominate over the latter in generating splitting rules (Prasad et al., 2006).

4. ANALYSIS OF MULTIPLE DISADVANTAGES

Data analysis was made with IBM-SPSS 20.0. Decision tree was selected via 10-fold cross-validation with the popular *1 – SE* rule[2] in the tree pruning procedure. The cross validated mean squared error was equal 2.461 (0.81 with respect to the mean squared error in the root node). The terminal nodes are L=16 (Figure 2). Considering the big size of the complete exploratory tree we decided to prune it using the cross-validation to improve the interpretation of the results.

As known in the literature the main problem of CART-like methods is their instability (e.g. high variance and low bias). Nevertheless it should be taken into account that the main purpose of the paper is the description of the phenomenon and not to get the best predictive performance. For this reason we think the CART is the best tool for an immediate and simple interpretation. The use of Boosting, Random Forest, Bagging etc., allow a great improvement of predictive power but at the same time these methods determine the complete loss of information that allow to interpret the phenomenon. As Breiman says: “What one loses, with the trees, is a simple and interpretable structure. What one gains is increased accuracy” (Breiman, 1996).

The regression tree analysis produced a partition of the sample of individuals into groups following the interactions between the predictors and the dependent variable (LS). In this way it is possible to measure the effect of the different interactions on LS. A node is declared “terminal node” according to two criteria: 1) if the impurity reduction achievable through the division of the node is less than a predetermined threshold; 2) if the size of the node is less than a predetermined threshold.

In the interpretation of the tree it is possible to follow different paths of the hierarchical structure by identifying the different interactions between predictors

---

[2] *1-SE* rule choose the pruned tree producing estimate within one standard error of the minimum: $R^{ts}(T^{**}) \leq [R^{ts}(T^{*}) \pm SE(R^{ts}(T^{*})]$, where $R^{ts}(T^{**})$ indicates any cross validation error within the standard error of the minimum error $R^{ts}(T^{*})$

and the LS. The presence or absence of specific disadvantages determine in which node the unity will fall. Each terminal node will be labelled with the average value of the response variable. In this way it is possible to understand how the LS varies in correspondence to different interactions between predictors.

Table 2 shows the gain summary for terminal nodes of our regression tree. Out of 16 terminal nodes, in 4 the average value of LS is grater than the one calculated on the entire sample. Node twenty-three represents 28.7% of the sample and has the higher level of LS (7.59). People in this node are characterized by living in household with adequate economic resources, employed, satisfied of their work and of their leisure time.

Node twenty-nine represents 11.3% of the sample and it has a slightly lower LS average (7.16). In this case we find in this node people with low economic resources and low satisfaction about leisure time (probably connected to their financial possibilities) but they are satisfied with the job, satisfied of the relationships within the family and they have a good perceived health.

Table 2. Gain Summary for Nodes

Node	n	Percent	LS Mean
23	7,063	28.7	7.59
29	2,772	11.3	7.16
28	3,751	15.2	7.07
24	694	2.8	7.01
11	1,553	6.3	6.70
27	1,100	4.5	6.44
26	3,048	12.4	6.35
18	1,379	5.6	6.18
22	285	1.2	6.14
30	83	0.3	5.76
25	1,356	5.5	5.75
19	662	2.7	5.46
12	165	0.7	5.35
16	238	1.0	4.71
20	365	1.5	4.62
10	120	0.5	3.22

In this kind of technique, however, it is interesting not only the analysis of the terminal nodes but also the analysis of the path leading to the terminal nodes that gives information of the effect of specific disadvantages on LS. For example the average LS of people aged 25-64 years old in our sample is 6.8. The first split of the tree is determined by the lack of economic resources in the household. The individuals in node one who lack economic resources have LS of 6.3. In node two, where we find people that do not suffer of this disadvantage the LS is much higher (7.3). In this way we can estimate the net effect of the variable on LS to one point on average over a maximum of ten between having and not having enough economic resources (Figure 2).

In the same way we can estimate the effect of the family relationship on LS. Node one (LS=6.3) is split by low satisfaction in family relationship. The difference in LS between node three (LS=6.4) that group people satisfied with their family relationships and node four (LS=4.9) with people unsatisfied of their family relationships is a staggering 1.5 points.

Another important element that can be measured with this analysis is how the impact of a specific disadvantage on LS varies according to the disadvantages experienced by respondents. This aspect is of paramount importance and little appreciated in the literature on the subject.

We will make it clear with an example. Let consider leaves ten (LS=3.2) and sixteen (LS=4.7): these two nodes are produced by a very different path but the final split is determined by the same variable: bad perceived health conditions. In node ten we have people that lack economic resources, have low satisfaction of family relationship and a bad perceived health. The net effect of bad perceived health measured by comparison with node nine (LS=5.1) is of a staggering 1.9 points. On the contrary node sixteen groups people with low economic resources and low satisfaction for the leisure time but good family relationships. Not only the average LS in node sixteen (4.7) is higher than that of node ten (3.2) but also the loss of average LS caused by a bad perceived health is higher splitting from node four to node ten than in that between node fifteen and sixteen. In the first case the loss of average LS is of 1.7 points in the second case the loss is of 1.4 points.

In node ten the accumulation of three important disadvantages determines a very low level of LS while in node sixteen the presence of a key component (satisfaction for family relationships) is an indicator of resilience that allows people to maintain higher level of LS even when facing very important disadvantages.

The nodes on the right side of Figure 2 represent people with adequate economic resources, who generally have a higher level of LS. On the left side of Figure 2 the leaves representing people with low economic resources who have significant lower level of LS.

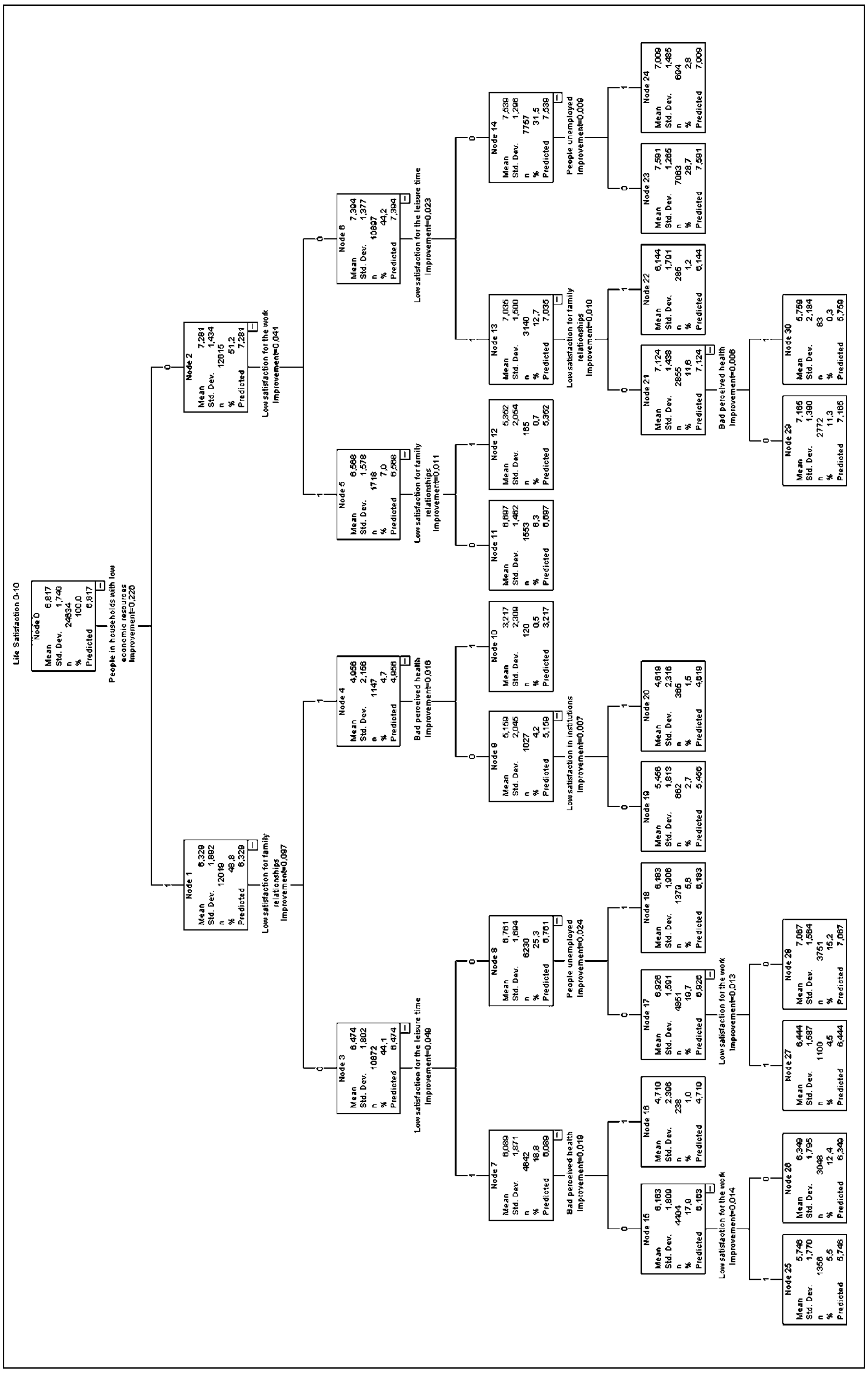

Figure 2. Regression Tree on Multipurpose survey on Everyday life aspects data set

However having low economic resources does not always determine low levels of LS. For example if we consider terminal node twenty-eight the LS is higher than the average (7.0). This node represent people with low economic resources but satisfied with their family relationships, leisure time and job. By consequence the LS is high even if they lack of economic resources.

On the contrary node twelve has a lower level of LS in comparison to the average (5.3). This node represents people with adequate economic resources but with low satisfaction for the work and low satisfaction for the family relationships. By consequence, the availability of economic resources is not enough to compensate poor for poor family relationships and low level of work satisfaction.

The model shows, also, how some variables can single out the effect of some disadvantages. If we consider node seven and node eight the net effect of having low satisfaction for leisure time is of 0.7 point. Considering node twenty-five which groups people that also have a low satisfaction for the work the LS drops of another 0.3 points with an additive effect due to the multiple disadvantages. On the contrary if we consider node twenty-six the LS is higher than in node seven (6.3) because being satisfied for its own work has a stronger positive effect on LS that smooth the effect of being unsatisfied for the leisure time.

Finally to have an overall assessment of the importance of each variable, in terms of its effect on LS, it is possible to use the CART approach, which allows to build a graph showing the importance of the variables in terms of their predictive power. Variable importance in CART is the contribution of each predictor in reducing the overall impurity measured using the generic measure of split. The greater the contribution of a predictor in reducing the overall impurity the greater its importance.

Figure 3 shows the ranking of variables according to their contribution in reducing the overall impurity. The force of each variable is evaluated in percentage with respect to the effect of the stronger variable. As shown in figure 3, the most important predictor are: “Low satisfaction for family relationships”; “Low satisfaction for leisure time”; “People in households with low economic resources”; “Low satisfaction for the job”; “Low satisfaction for friends relationship”; “Bad perceived health”; “People unemployed”. According to these results we can say that there are essentially three domains that strongly influence LS. Firstly it appears clearly that the “social relationships” are crucial for the LS. Having a low level of satisfaction for family relationship and a low satisfaction for leisure time or a low level of satisfaction for friends relationship rank as the 1st, 2rd and 5rd predictor by importance. This result is coherent with the literature and reinforces the idea that social connections and more generally the social environment where the individual lives are crucial for LS and for its well-being.

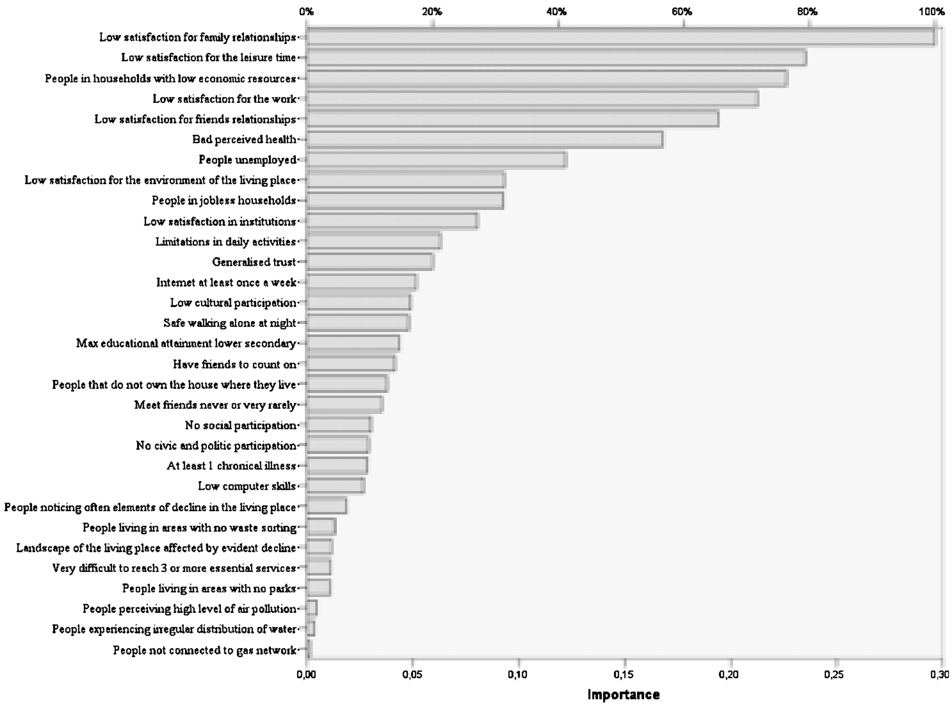

Figure 3. Normalized importance of predictors

Secondly, LS is strongly influenced by "economic well-being". People in household with low economic resources show a relevant decrease in LS. Thirdly work also stands out as an important factor of LS because work provides economic resources for the individuals and also for its social effects in terms of self-esteem, status and social contacts.

The other variables contribute in a less significant way (20% or less of the effect of the 1st variable) to reduce LS (Figure 3). However as shown in the analysis of the regression tree (Figure 2), even if a variable does not have a strong effect per se on LS it can have a strong cumulative effect when associated with other disadvantages.

5. CONCLUSIONS

In this paper we used regression tree analysis to evaluate the effect of disadvantages on LS, but also to identify the importance of the explanatory variables.

Our study allows to rank the disadvantages by their decreasing level of impact on LS giving thus an idea of the importance of each factor. Specifically, the analysis carried out through regression tree, has identified, “Low satisfaction for family relationships”; “People in households with low economic resources”; “Low satisfaction for friends relationships” and unemployment as major factors impacting on LS.

The results of the study show clearly that LS is achieved through a delicate balance of objective and subjective conditions which is coherent with the literature and the BES framework. Being LS a cognitive evaluation made by the individual about his/her life as a whole it is clearly effected by objective factors such as the economic resources and the working status of the individuals but also by subjective evaluations of specific aspects of his/her life. As the analysis points out the social and familiar relationships play a central role and are even stronger than objective factors.

For this reasons we criticize the analysis on LS impacts carried out considering only objective variables and we advocate the relevant role played by dissatisfaction of specific aspects of life as crucial in showing areas of strong vulnerabilities.

If we could draw some conclusions for policy makers out of this study it would be worth saying that policies aiming at having an impact on LS would be successful if they manage to improve the economic resources of the individuals (for example providing job opportunities) but policies could have a even stronger impact if they manage to promote a healthier social environment and trust among citizens.

REFERENCES

Balestra, C. and Sultan, J. (2013). *Home Sweet Home: The Determinants of Residential Satisfaction and Its Relation with Well-Being*. OECD Statistics Directorate Working Papers N.54, OECD Publishing, Paris.

Blanchflower, D.G. and Oswald, A.J. (2004). Money, sex and happiness: An empirical study. In *Scandinavian Journal of Economics*. 106: 393–415.

Boarini, R., Comola, M., Smith, C., Manchin, R. and de Keulenaer, F. (2012). *What Makes for a Better Life? The Determinants of Subjective Well-Being in OECD Countries – Evidence from the Gallup World Poll*. OECD Statistics Working Papers N.03, OECD Publishing, Paris.

Breiman, L. (1996). Bagging predictors. In *Machine Learning*. 24: 123–140.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984). *Classification and Regression Trees*. Wadsworth, Belmont, California.

Chipman, H.A., George, E.I. and McCulloch, R.E. (1998). Bayesian CART model search. In *Journal of the American Statistical Association*. 93: 935–960.

Clark, L.A. and Pregibon, D. (1991). Tree-Based Models. In J.M. Chambers and T.J. Hastie, editors, *Statistical Models in S.* Wadsworth and Brooks/Cole, Pacific Grove, California: 377–419.

Cnel and Istat (2010). *La misurazione del Benessere Equo e Sostenibile*. Istat, Roma.

Diener, E. and Chan, M.Y. (2011). Happy people live longer: Subjective well-being contributes to health and longevity. In *Applied Psychology: Health and Well-being*. 3: 1–43.

Diener, E., Diener, M. and Diener, C. (1995). Factors predicting the subjective well-being of nations. In *Journal of Personality and Social Psychology*. 69: 851–864.

Dolan, P., Layard, R. and Metcalfe, R. (2011). *Measuring Subjective Well-Being for Public Policy*. Office for National Statistics, London.

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. In *Journal of Economic Psychology*. 29: 94–122.

Frey, B.S. and Stutzer, A. (2000). Happiness, economy and institutions. In *The Economic Journal*. 110: 918–938.

Giovannini, E. Rondinella, T. (2012). Measuring Equitable and Sustainable Well-Being in Italy. In F. Maggino and G. Nuvolati, editors, *Quality of Life in Italy: Research and Reflections*, Social Indicator Research Series 48, Springer, Dordrecht-Heidelberg: 9–28.

Helliwell, J.F. (2008). *Life Satisfaction and the Quality of Development*. National Bureau of Economic Research, NBER Working Paper N.14507, Cambridge.

Helliwell, J.F. and Barrington-Leigh, C.P. (2010). *Measuring and Understanding Subjective Well-Being*. National Bureau of Economic Research, NBER Working Paper N.15887, Cambridge.

Helliwell, J.F. and Wang, S. (2011). *Weekends and Subjective Well-Being*. National Bureau of Economic Research, NBER Working Paper N.17180, Cambridge.

Istat, (2013). *Bes 2013. Il Benessere equo e sostenibile in Italia*. Istat, Roma.

Istat, (2014). *Bes 2014. Il Benessere equo e sostenibile in Italia*. Istat, Roma.

Iverson, L.R. and Prasad, A.M. (1998). Predicting abundance of 80 tree species following climate change in the eastern United States. In *Ecological Monographs*. 68: 465–485.

Kahneman, D., Diener, E. and Schwaarz, N. (Eds.) (1999). *Well-Being. The Foundations of Hedonic Psychology*. Russel Sage Foundation, New York.

Lucas, R. (2007). Long-term disability is associated with lasting changes in subjective well-being: Evidence from two nationally representative longitudinal studies. In *Journal of Personality and Social Psychology*. 92: 717–730.

Moore, D.M., Lees, B.G. and Davey, S.M. (1991). A new method for predicting vegetation distributions using decision tree analysis in a geographic information system. In *Environmental Management*. 15: 59–71.

Ng, Y.K. (1997). A case for happiness, cardinalism and interpersonal comparability. In *The Economic Journal*. 107: 1848–1858.

Ng, S.T., Tey, N.P., Yew, S.Y., Sia, B.K. and Long, B.S. (2012). Effects of quality of service and activities on life satisfaction of residents in nursing homes. In‘*Wulfenia Journal*. 19: 153–163.

OECD, (2011). *Education at a Glance: OECD Indicators*. OECD publishing, Paris.

OECD, (2013). *OECD Guidelines on Measuring Subjective Well-Being*. OECD publishing, Paris.

Prasad, A.M., Iverson, L.R. and Liaw, A. (2006). Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. In *Ecosystems*. 9: 181–199.

Rode, J.C. (2004). Job satisfaction and life satisfaction revisited: A longitudinal test of an integrated model. In *Human Relations.* 57: 1205–1230.

Sacks, W.D., Stevenson, B. and Wolfers, J. (2010). *Subjective Well-being, Income, Economic Development and Growth*. National Bureau of Economic Research, NBER Working Paper N.16441, Cambridge.

Silva, J., de Keulenaer F. and Johnstone, N. (2012). *Environmental Quality and Life Satisfaction: Evidence Based on Micro-Data*. OECD Statistics Working Papers N.44, OECD Publishing, Paris.

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

Therneau, T.M. and Atkinson, E.J. (1997). *An Introduction to Recursive Partitioning Using the RPART Routines*. Mayo Clinic, Section of Biostatistics, Department of Health Sciences Research Technical report N. 61. Rochester (MN). URL http://www.mayo.edu/research/documents/biostat-61pdf/doc-10026699?_ga=1.31854330.510244244.1445965172 Last access: 17/12/2014

UNICEF, (2013). *Child Well-Being in Rich Countries: A Comparative Overview*. Innocenti Report Card 11, UNICEF Office of Research, Florence.

Verbyla, D.L. (1987). Classification trees: A new discrimination tool. In *Canadian Journal of Forest Research*. 17: 1150–1152.

Winkelman, L. and Winkelman, R. (1998). Why are the unemployed so unhappy? Evidence from panel data? In *Economica*. 65: 1–15.