DOI: 10.47743/eufire-2024-1-10

FINANCIAL SECURITY, INDIVIDUAL AUTONOMY, AND DEMOCRATIC IDEALS: KEY LIFE SATISFACTION DRIVERS

DANIEL HOMOCIANU
Alexandru Ioan Cuza University of Iasi
Iasi, Romania
daniel.homocianu@uaic.ro

DINU AIRINEI
Alexandru Ioan Cuza University of Iasi
Iasi, Romania
adinu@uaic.ro

Abstract
This paper explores the most robust determinants of life satisfaction using empirical analysis of World Values Survey data, including the latest version 4.0, time series 1981-2022. Three significant ones emerged (a triad), namely contentment with household financial situation, freedom of choice and control, and democracy in terms of exposure to it and democratic values, all in a logit model with good accuracy of classification (AUC-ROC greater than 0.8) and a high maximum probability (more than 95%). Rigorous selection processes coupled with advanced methodologies such as variable selection, triangulation, cross-validations (random, and non-random - both on socio-economic variables and dataset versions), overfitting removal, collinearity and reverse causality checks, and different regressions contributed to the evaluation and validation of the final most robust model. Zlotnik and Abraira's prediction nomograms helped rank the predictors and stood behind the final model with three core components to estimate life satisfaction probabilities. The findings are a step forward in life satisfaction research and provide significant ideas for policymakers.
Keywords: *World Values Survey (WVS); life satisfaction; financial contentment; freedom of choice and control; democracy.*
JEL Classification: I31, C55, G51.

1. INTRODUCTION

Life satisfaction seems to be a complex and subjective concept that can vary so much from person to person and can depend on many factors such as personal values, relationships, health, financial situation and stability, and life experiences. While some people find happiness in money (Boyce *et al.,* 2010) and material possessions (Keng *et al.,* 2000; Sirgy *et al.,* 2013), others find it in spiritual or emotional fulfillment (Koohbanani, 2013). Life satisfaction does not seem to be under the governance of a single formula. Moreover, its attainment varies from person to person, and individuals must discern what brings them fulfillment and happiness and actively pursue those aspects in their personal lives. It could also mean there are general and specific patterns for this satisfaction. The primary focus in this paper is on the former.

The study of life satisfaction is not something new. Historically, this research line can be rooted in the 18th century (Veenhoven, 1996), associated with Enlightenment thought. From this point of view, the purpose of existence is life itself rather than serving the ruler or God. Therefore, self-improvement and happiness become central values in a society responsible for providing citizens with what is necessary for a good life. The same conviction manifested a century later under the form of the Utilitarian Creed that the best society is that which offers the greatest happiness for the highest number of people and inspired two centuries later large-scale attempts to carry out social reform and influenced the development of the welfare state (Pacek and Radcliff, 2008). The overall progress started with creative efforts of a better society, translated first into attempts to avoid ignorance, diseases, hunger, and poverty, as well as increasing the level of literacy and controlling diseases and epidemics and later into ways to ensure a good life for all, and a good material standard of living through monetary earnings, income security, and income equality. The latter has given rise to much social research on poverty and social inequality (Schröder, 2016; Roth *et al.,* 2017). Later, the term quality of life emerged in the context of the new themes related to the limits to economic growth and post-materialism.

In terms of differences between life satisfaction and happiness (San Martín *et al.,* 2010), the latter is often described as a more momentary and emotional state (Silbermann, 1985; Mangels, 2009; Sundriyal and Kumar, 2014), often influenced by external factors such as events, experiences, or possessions. It can be short-lived and fluctuate frequently. Moreover, life satisfaction is a more enduring and cognitive evaluation of a particular life, as a whole or overall happiness (Lu, 1999; Suldo *et al.,* 2006; Veenhoven, 2012; Zhang *et al.,* 2018). It comprises many factors, including individual overall sense of purpose, relationships, financial stability, health, etc. Life satisfaction tends to be a more stable and long-term assessment of happiness. It is worth noting that while there is often an overlap between happiness and life satisfaction, they are not the same thing and can exist independently of each other. One can feel happy at a particular moment but still have low life satisfaction or vice versa.

Other authors also indicate the role of socio-demographic and individual features (Argan *et al.,* 2018; Rajani *et al.,* 2019). They emphasize influences from this category above, such as age, gender (Becchetti and Conzo, 2022), psychological features, lifestyle, participation in leisure activity, and satisfaction related to spending free time or leisure satisfaction (Dirzyte *et al.,* 2022).

Life satisfaction is more liable to shifts in the aspiration level (Ng, 2022) when compared to happiness, thus reducing the comparability of the resulting indices. Moreover, life satisfaction is the evaluation of personal life, not simply the current level of happiness (Feldman, 2008).

According to some other authors (Gundelach and Kreiner, 2004; Li *et al.,* 2017), higher levels of freedom of choice and control are usually strong influences associated with life satisfaction. From other perspectives (Diener and Diener, 2009; Oishi *et al.,* 2009), those with higher levels of financial satisfaction are also more inclined to show higher levels of life satisfaction.

According to other scholars, those more inclined and exposed to democracy as an expression of the will of the people (Denton, 2015) and of the subjectivity of the society (Łużyński, 2019) or as a crucial way to realize human rights (Yin, 2022), are also more likely to be satisfied with their lives (Dorn *et al.,* 2007; Owen *et al.,* 2008).

Some socio-demographic features (Meléndez *et al.,* 2009) seem to be also significant influences associated with this type of satisfaction. For instance, some researchers (Kassenboehmer and Haisken-DeNew, 2012; Hellevik, 2017; Bartram, 2020; Bittmann, 2021; Kaiser *et al.,* 2022) invoke even an U-shape when it comes to the graphical representation of the influence of age on life satisfaction, with high levels of life satisfaction in young adulthood, a gradual decline in middle age with a minimum of being satisfied with life between 40 and 60 years of age, and then an increase in later life. Other scholars revealed significant correlations between personality, self-esteem, and life satisfaction (Halvorsen and Heyerdahl, 2006) or between optimism-related variables, goal orientation, and the same type of satisfaction (Supervía *et al.,* 2020).

Still, the main hypotheses of this paper are:

- H1-Freedom of choice and control (Verme, 2008; Sohier, 2018) are strongly related to well-being, happiness, and this type of satisfaction (Bartolini and Sarracino, 2015).
- H2-Financial choice and satisfaction are closely associated with well-being, the latter being considered more than happiness and life satisfaction (Ruggeri *et al.,* 2020). Therefore, the first two are also related to being satisfied with life.
- H3-Democratic values (Keane *et al.,* 2012) positively correlate with increased life satisfaction.

The article further describes the data and methodology used before presenting and discussing the main findings in a dedicated section.

2. DATA AND METHODS

This article started from one of the most comprehensive World Values Survey (WVS) datasets. The latter (version 4.0, WVS_TimeSeries_4_0.dta - https://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp) includes 1,045 variables and 450,869 raw observations. It served all selection rounds. Other three versions have been used just in the first selection round (Adaptive

Boosting in the Rattle visual library of R), namely: version 3.0 (WVS_TimeSeries_1981_2022_Stata_v3_0.dta, 1,041 variables and 440,055 observations, available online on the WVS site until the end of 2022), version 2.0 (WVS_TimeSeries_1981_2020_stata_v2_0.dta, 1,072 variables, and 432,482 records) and 1.6 (WVS_TimeSeries_stata_v1_6.dta, 1,045 variables, and 426,452 observations), the latter two still available on the WVS site. Their .csv (Comma-Separated Values data format) exports were preceded by designing, testing, and running a script sequence responsible for removing the DK/NA (Tsikriktsis, 2005) values, Don't Know or No Answer/No Opinion or Not Applicable/Not Asked - coded by WVS as negative ones, artificially increasing the scales, and not beneficial for selections. Figure 1 and https://tinyurl.com/ddz8239j of all variables and by a simple binary derivation, i.e. A170bin of the original variable to analyze A170, Satisfaction with your life. This applies considering the two symmetric halves of its original scale: 1-5 for 0, and 6-10 for 1, https://tinyurl.com/3h59wyhp. Moreover, the option to generate numerical values for labeled variables, instead of the text, was enabled when exporting, e.g., *export delimited using "F:\data\WVS-TS4_A170bin.csv", no label replace*.

```
. label list A170
A170:
          -5 Missing; Unknown
          -4 Not asked
          -2 No answer
          -1 Don't know
           1 Dissatisfied
           2 2
           3 3
           4 4
           5 5
           6 6
           7 7
           8 8
           9 9
          10 Satisfied
```

. tabulate A170

Satisfaction with your life	Freq.	Percent	Cum.
Missing; Unknown	123	0.03	0.03
Not asked	1,982	0.44	0.47
No answer	1,245	0.28	0.74
Don't know	2,602	0.58	1.32
Dissatisfied	17,616	3.91	5.23
2	11,700	2.59	7.82
3	20,622	4.57	12.40
4	24,586	5.45	17.85
5	60,679	13.46	31.31
6	49,149	10.90	42.21
7	67,701	15.02	57.22
8	83,358	18.49	75.71
9	46,460	10.30	86.02
Satisfied	63,046	13.98	100.00
Total	450,869	100.00	

Running ***1remove_DKNA.do*** for Stata **(**at https://tinyurl.com/ddz8239j)

. tabulate A170

Satisfaction with your life	Freq.	Percent	Cum.
Dissatisfied	17,616	3.96	3.96
2	11,700	2.63	6.59
3	20,622	4.64	11.22
4	24,586	5.53	16.75
5	60,679	13.64	30.39
6	49,149	11.05	41.44
7	67,701	15.22	56.65
8	83,358	18.74	75.39
9	46,460	10.44	85.83
Satisfied	63,046	14.17	100.00
Total	444,917	100.00	

Figure 1. The target variable and the frequency of its values before and after removing (1remove_DKNA.do for Stata) the artificial increase of scales due to the original encoding of DK/NA values as negative numbers by WVS

The next step was to load these .csv exports into the Rattle interface, version 5.5.1 – started using two commands in R, namely *library* and *rattle*, then set A170bin as the target, ignore its source A170 from the list of inputs and apply the Adaptive Boosting technique for the decision tree classifiers (Karabulut and Ibrikci, 2014). This step ran (Chen *et al.,* 2008; Williams, 2011) for four versions of this most comprehensive dataset of WVS - v4.0, 3.0, 2.0, and 1.6 - using default settings. The purpose was to discover the most resilient related variables at the intersection of those four versions, cross-validation considerations. The latter was the 1st selection round, 9 resulting variables.

Other alternative selections applied only to the most recent and comprehensive version (4.0) and starting after the same DK/NA treatments considered:

- The use of the Naïve Bayes classification algorithm inside the Microsoft DM add-in for spreadsheets that works together with SQL (Structured Query Language) Server Analysis Services 2016 on a machine running Windows 10 Professional X64;
- The use of filter options applied to the results of a correlation command (PCDM or Pairwise Correlation-based Data Mining) for selections in Stata 17 (invoked for the scale form of the target variable, namely A170, Figure 2) on the same machine.

First, they meant a minimum threshold of 0.1 (Schober *et al.,* 2018) for the absolute values of pairwise correlation coefficients (Homocianu and Airinei, 2022) between each recoded variable from the previous step and the one to analyze. In addition, a maximum accepted p-value, max p=0.001, and a minimum support afferent to a minimum number of valid observations for the target variable, at least half the total corresponding number – 444917/2, Figure 2, for each pair.

Only seven (7) variables proved to be the most resilient at the intersection of Adaptive Boosting (Rattle in R), Naïve Bayes (Analysis Services), and PCDM (Stata). From these seven, only the first six were confirmed, successive invocations until no loss in selection, when using CVLASSO or Cross-Validation LASSO for performing random cross-validations and RLASSO or Rigorous LASSO for removing overfitting available after installing the LASSO - Least Absolute Shrinkage and Selection Operator pack (Ahrens *et al.,* 2020), and the BMA (Bayesian Model Averaging) command in Stata 17 for both forms of the target variable (S002VS set as auxiliary influence in BMA).

Additionally, some socio-demographic variables served non-random cross-validations. For the first (non-random cross-validations), these variables helped mixed-effects models (Roberts *et al.,* 2017; DeBruine and Barr, 2021) in Stata 17 MP or Multi-Processing, 64-bit version. Such models included both fixed-effects, the remaining six variables after the previous selection phases and random ones, clusters on gender, age, marital status, number of children, education level, income level, professional situation, settlement size, country, and survey year, all as socio-demographic variables, bottom of Table A1, the Appendix A section.

The immediate selection phase measured the existing collinearity between the remaining influences, the six above. First, a matrix with correlation coefficients augmented with intensity bars has been generated only for these six remaining influences (Schober *et al.,* 2018). In addition, Ordinary Least Squares (OLS) regressions served the same purpose by measuring the computed VIF (Variance Inflation Factor) against (eq.1) the maximum accepted VIF threshold of the model (Vatcheva *et al.,* 2016) for all combinations of two influences of those 6 (combinations of n=6 taken by k=2, meaning 15 possibilities – eq.2). E235 and E236 emerged as being collinear at this point.

Model's maximum accepted VIF = 1/ (1– model's R-squared) (1)

C (n, k) = n! / (k! * (n-k)!) (2)

Where: C (n, k) is the number of combinations of n taken by k,
n! is the factorial of n,
k! is the factorial of k,
and (n-k)! is the factorial of (n-k).

In addition, to choose between these two, logistic regressions have been used. The variable that is responsible for generating models with more explanatory power/larger R-squared (Irandoukht, 2021) and more information gain/smaller values for both AIC (Akaike Information Criterion) and BIC or Bayesian Information Criterion (Lai, 2020) was preserved, e.g. E236.

Additionally, two prediction nomograms (Zlotnik and Abraira, 2015) resulted one simple and another one augmented with additional details to become self-explanatory when using the nomolog command (after its previous install using a specific installation syntax, namely net install st0391, and considering most stalwart remaining influences.

Moreover, reverse causality checks were performed using ordinal logit/logistic or ologit (Ordered LOGIT) regressions and the scale form of the target variable A170osc. In each of these regressions that considered only one of the remaining input variables, the latter served both as input and outcome interchanging these roles with A170osc, regression pairs. A larger R-squared, representing smaller differences between the observed data and the fitted values/theoretical model, and/or a lower AIC and BIC, better fit and smaller information loss for the resulting models are an indication that each of the remaining variables to further select are more likely to be determinants of A170osc rather than vice versa, determined by it.

Finally, for each variable in the core, three determinants, a two-way graphical representation scatter chart was automatically generated by considering each corresponding relationship with the outcome variable, life satisfaction in its scale format tabulated on average by peculiar criteria using the tabstat command in Stata. The reporting of results mainly benefited from the estout prerequisite package, with support for both the *eststo* and *esttab* commands (Jann, 2005; Jann, 2007), allowing the direct generation of tables, in the console and as external files, respectively) with default performance metrics and some additional ones (Homocianu and Tîrnăucă, 2022) for well-known statistical models.

3. RESULTS AND DISCUSSION

After performing the first selection step using Adaptive Boosting on four versions of the WVS dataset, a set of nine intersecting variables resulted A008, A009, A173, C006, D002, E235, E236, S002/S002VS, and X047/X047_WVS, as seen in all four sources. This acted as the first selection round based on cross-validations considering different versions, with different numbers of observations of the source dataset.

The results of applying the first alternative selection based on Naïve Bayes classification in Microsoft DM and Analysis Services on version 4.0, the most comprehensive one of the WVS dataset, after removing DK/NA values, was a dependency network. Only eight of those nine influences resulting at the previous step, all at the intersection of those four sources above, except for D002, are present in this network.

Next, some filters served the selections when performing correlations using the PCDM custom command in Stata (Homocianu and Airinei, 2022) on the same WVS dataset (the most recent and comprehensive version, 4.0). For instance, min.abs.correl.coeff.=0.10, min.N=222459(=round of 444917/2, where 444917 is the number of valid observations for the target variable, as seen on the top-right and bottom of Figure 2, and maximum p-value of 0.001. The results in Figure 2 indicate only seven (A008, A009, A173, C006, E235, E236, and X047_WVS) of those nine remaining variables above, all except for D002 – low support, meaning just 26459 observations as seen in the description of variables and general statistics, Tables A1 and A2, Appendix A-https://tinyurl.com/nb9xkcrw, and S002VS – low correlation coefficients below the threshold value of 0.10.

The next concern was to start from the same nine robust common influences above and perform random cross-validations (cvlasso), selections based on removing overfitting (rlasso), and BMA selections, which reports Posterior Inclusion Probabilities or PIP, preferably as close to 1 as possible, all three until convergence (no loss) and considering both forms (binary and scale) of the target variable (A170 and A170bin). Cvlasso used both the *lse* option, largest lambda for which MSPE or the Mean Squared Prediction Error is within one standard error of the minimal MSPE, and the *lopt* one, the lambda that minimizes MSPE. After this stage, those seven variables remaining after PCDM persisted, all nine except for D002 and S002VS.

Next, three rounds of non-random cross-validations run using mixed-effects modeling. For the first such round, just one variable, namely X047_WVS, scale of incomes of the remaining seven, the ones bolded in Figure 2, acting as fixed-effects lost significance, nine from eleven models/scenarios with A170 set as target. And this was observed because of considering many clustering criteria/random effects the socio-demographic variables mentioned in the previous section, and two mixed-effects regression types, both *melogit* (Mixed Effects LOGIT) for the binary form of the response variable and *meologit* (Mixed Effects Ordered LOGIT) for the one having values on a scale. If considering only the remaining six as fixed-effects, all bolded in Figure 2 except for X047_WVS - 2nd round of non-random cross-validations, there was no loss in significance no matter the clustering criteria.

Outcome(y)	Input(x)	Correl.Coef.(CC)	Abs.Val.CC(ACC)	No.Obs.(Nobs)	Signif.(p)
A170	A170	1	1	444917	0
A170	COW_NUM	-0.105104312	0.105104312	444917	0
A170	A003	-0.122021491	0.122021491	419991	0
A170	**A008**	-0.466223018	0.466223018	436729	0
A170	**A009**	-0.302639294	0.302639294	433318	0
A170	A030	-0.114502351	0.114502351	440004	0
A170	A124_07	-0.10253212	0.10253212	405676	0
A170	A124_09	-0.126325523	0.126325523	397690	0
A170	**A173**	0.40825782	0.40825782	427474	0
A170	**C006**	0.566084129	0.566084129	431278	0
A170	D059	0.114466403	0.114466403	375051	0
A170	E037	-0.143889074	0.143889074	414830	0
A170	E069_06	-0.104019174	0.104019174	417463	0
A170	E124	-0.170445435	0.170445435	317632	0
A170	E218	0.100172563	0.100172563	240338	0
A170	E234	0.110268045	0.110268045	243962	0
A170	**E235**	0.124808794	0.124808794	253597	0
A170	**E236**	0.200398372	0.200398372	242184	0
A170	F118	0.111835555	0.111835555	400329	0
A170	G006	-0.131029692	0.131029692	428960	0
A170	X044	-0.168521588	0.168521588	370305	0
A170	X045	-0.215984693	0.215984693	374130	0
A170	**X047_WVS**	0.222011672	0.222011672	406573	0
A170	X047R_WVS	0.203961615	0.203961615	407680	0
A170	Y020	0.112914118	0.112914118	432158	0
A170	Y022	0.118219199	0.118219199	386674	0
A170	Y011B	-0.131037177	0.131037177	428960	0
A170	Y014B	-0.103993829	0.103993829	417463	0
A170	Y022B	0.114356785	0.114356785	375051	0
A170	Y023A	0.111835552	0.111835552	400329	0

Figure 2. Results of a selection command (PCDM) based on pairwise correlation and additional filters on magnitude, support, and significance for the WVS dataset (v 4.0)

To additionally validate the simultaneous removal of both X047_WVS and S002VS at the previous steps, D002 no longer considered due to its low number of valid observations, an additional set of non-random cross-validations (3rd round of non-random cross-validations) based on both *melogit* and *meologit* has been performed (eight fixed-effects and other ten clustering variables in Table A3, Appendix A-https://tinyurl.com/nb9xkcrw).

Those six remaining influences above proved to be robust, in terms of no loss of significance, in this additional round, namely A008, A009, A173, C006, E235, and E236. The other two failed at least in one scenario: X047_WVS when cross-validating using most socio-demographic variables as cluster criteria except for the age (X003) and the number of children (X011) and considering the scale form of the target variable (A170), while S002VS (chronology of EVS-WVS waves) when cross-validating using the highest educational level attained (X025), the country code (S003), and the survey year (S020) as cluster criteria and considering both forms of the target variable (A170bin and A170).

Next, when verifying the existing collinearity using the first method, a matrix with correlation coefficients and a minimum visual augmentation using intensity bars for the remaining six influences emerged, in Figure 3 all Pearson correlation coefficients are significant at 0.1‰. The latter, as absolute values, shows no evidence of collinearity if considering 0.1 and 0.39 as the lower and upper limits for weak correlation, while 0 and 0.1 as the ones for negligible correlation (Schober *et al.,* 2018).

	A008	A009	A173	C006	E235	E236
A008	1					
A009	0.3699	1				
A173	-0.249	-0.1972	1			
C006	-0.3459	-0.2549	0.3193	1		
E235	-0.0479	-0.0518	0.1556	0.0933	1	
E236	-0.1415	-0.0977	0.1414	0.2105	0.2107	1

Figure 3. Collinearity view using a matrix with correlation coefficients coupled with intensity bars only for the remaining six influences and the *pwcorr* command in Stata

In addition, OLS max.Comput.VIF against OLS max.Accept.VIF (eq.1) for models with all six previously tested influences (Figure 3) at once (model 1 in Table A4) and additionally taken each two (models 2-16 in Table A4) in all 15 combinations (eq2.) served discovering further evidence of collinearity. The removal decision considered one of the two variables. It is about E235 and E236, namely the importance of democracy as own value and democracy as perceived in own country, respectively model 16 in Table A4, namely the only one for which OLSmaxComputVIF > OLSmaxAcceptVIF (Homocianu and Tîrnăucă, 2022).

After performing additional logit regressions, E236 brought higher accuracies (AUC-ROC or Area under the ROC Curve (Receiver Operating Characteristic Curve) of 0.8350 and 0.8351) and R-squared values (0.2645 and

0.2649) together with better fit due to lower AIC and BIC values than E235 (AUC-ROC of 0.8340 and 0.8345, R-squared of 0.2624 and 0.2638). And this was recorded when considering the binary form of the target variable, model 3 vs. model 4 for comparable support due to the same number of observations using a filtering condition on the variable dropped, and model 5 vs. 6 for all but different numbers of available responses and no filtering condition - Table A5. In the case of additional ologit regressions, the models keeping E235 and dropping E236 were better as R-squared than those keeping E236 and dropping E235. Not the same applied in terms of information gain. Consequently, the balance is inclined towards keeping E236. This was at the expense of removing E235, the other democracy-related variable.

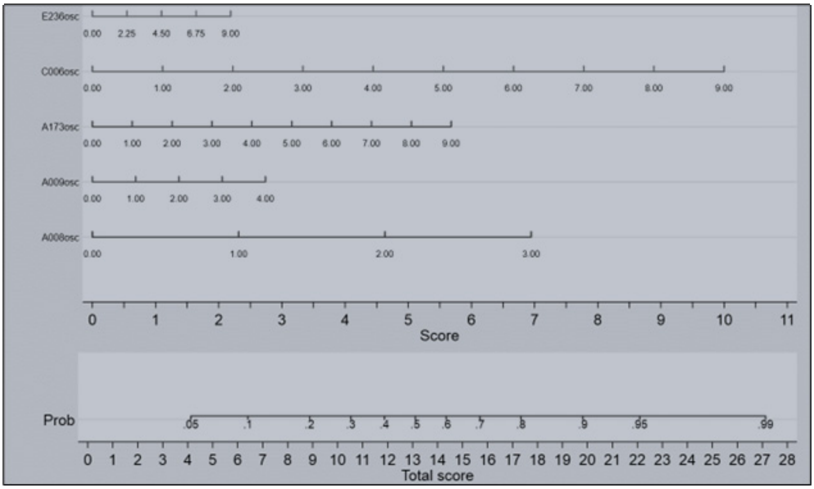

Figure 4. Prediction nomogram for a model with five influences

As support, 234,223 valid intersecting observations, meaning 51.95% of the total number of records for the entire dataset) corresponding to the last three waves were behind the first core model (model 5, Table A5). This is because all five most resilient influences and the response variable were considered simultaneously only in these three waves (2005-2009, 2010-2014, and 2017-2022), E236 having no observations for the first four. The same happened if removing E236 and preserving E235 with a slight increase in the number of responses (more than 245,000 valid intersecting observations - model 6, Table A5, and a slight decrease in terms of accuracy of classification (AUC-ROC=0.8345). If removing both E235 and E236 (model 2, Table A5), the support increases to 410,513 non-null intersecting observations, meaning 91.05% from the total number, namely 450,869, and 92.26% from those 444,917 valid for the target variable, while covering all seven waves and increasing the accuracy of classification (AUC-ROC=0.8458). Furthermore, the four remaining influences are now fully included in the list of the mightiest links in a Naïve Bayes dependency network.

Next, a simple Stata script design supports the alignment of the scales to 0 for the target variable and those corresponding to some solid influences of it. Another purpose of the latter was to optimize the following two prediction nomograms (Figure 4, *nomolog* command in Stata) for better readability. Both fundaments on binary logistic regressions. The corresponding two models are identical to those numbered 2 and 5 (Table A5) in terms of performance metrics and values of coefficients and errors for the top five influences except the sign of the first two, namely A008 and A009, due to reversed scales. These serve the visual interpretation of all remaining most potent influences. The first nomogram is simple, meaning the exact way it results after generating it using the *nomolog* command. It corresponds to a model with five resilient influences, with lower support, 51.95% of the total number of observations because of E236osc, but still generating a considerable R^2 (0.2649) and a good accuracy of classification (AUC-ROC of 0.8351). The second one corresponds to a model with only those four most resilient influences and high support (91.05% of the total observations of the WVS dataset, version 4.0), generating an R^2 of 0.2884 and a good accuracy of classification (AUC-ROC of 0.8458). This second nomogram is augmented with metadata about the individual score at the intersection with the X-axis, perpendicular lines drawn next to each possible value of the associated influences, respectively with suggestions for interpreting the input values, their corresponding scores, and the resulting total score and afferent likelihood, so that the nomogram is self-explanatory. The maximum theoretical probability for the most advantageous combination of variable values, extreme right in both nomograms is high. It indicates a value of more than 0.95, middle and bottom of Figure 4. These nomograms also reflect the magnitude of marginal effects, better comparability than with raw coefficients, for the corresponding variables. In addition, they serve to understand the cumulated effect size by considering the amplitude of any scale easily noticeable in these visual representations.

Final cross-validations considered models with seven, the quad-core plus the marital status/X007osc, social class/X045osc, and settlement size/X049osc or eight influences, the penta-core plus the same three above) and a reasonable number of criteria for cross-validations. They infirmed the last three influences added to those two cores when considering cross-validation criteria such as gender, employment status, the chronology of waves, country, and survey year in the case of the last two from those three, social class and settlement size, or the number of children in the case of marital status. This happens even if the two overall models with 7 and 8 influences did not show multi-collinearity and recorded significance for all corresponding variables and accuracy and R^2 scores better than the two core models with four and five components that passed already all the cross-validation tests.

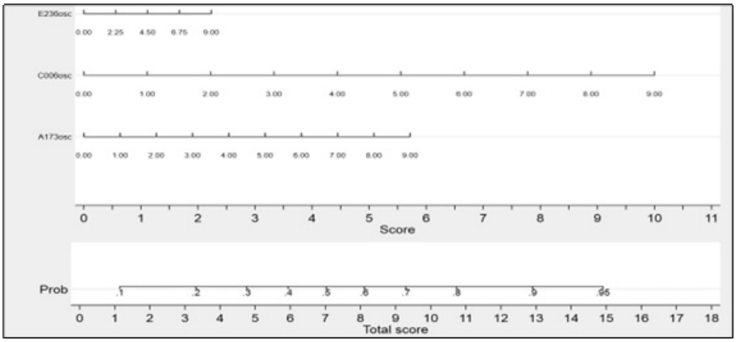

Figure 5. Prediction nomogram corresponding to a robust model with only three determinants (after performing reverse causality checks)

After performing some reverse causality checks (Table A6), only three variables from those five most robust influences (model 5, Table A5) confirmed also as determinants (A173osc, C006osc, and E236osc). Both a separate binary logistic model and a corresponding prediction nomogram (Figure 5) only for this triad of determinants were generated. The performance metrics of this model with only three components indicated an AUC-ROC of 0.814, lower than the one of model 5, Table A5, but still meaning a good accuracy of classification, and an R-squared of 0.2205. Moreover, the maximum theoretical probability for the most advantageous combination of those three determinant values (right edge of each line in Figure 5) still indicates a value of more than 95% (18 or the sum of 5.75, 10 and 2.25 on the score axis corresponds to a lot more than 0.95).

Some tabulations by mean support the to-way graphical representations between the target variable and each of the core model with three determinants (Figure 5) as depicted in the representation below (Figure 6). As magnitude (descending order of scale amplitudes), the first and most important of these three determinants (triad) corresponds to satisfaction with the household financial situation. It indicates that people more satisfied in such terms are more likely to show more contentment with their lives (positive influence or the maximum recoded value of 9 for C006osc, the right side of Figure 5). The latter means that this type of financial satisfaction (household-related) is among the best associated with life satisfaction according to WVS data (complete validation of H2). This finding is in line with the already documented relationship between both the financial costs and benefits and their well-being implications (Xiao *et al.,* 2009; Ng *et al.,* 2019; Brzozowski and Spotton Visano, 2020; Barrington-Leigh, 2021).

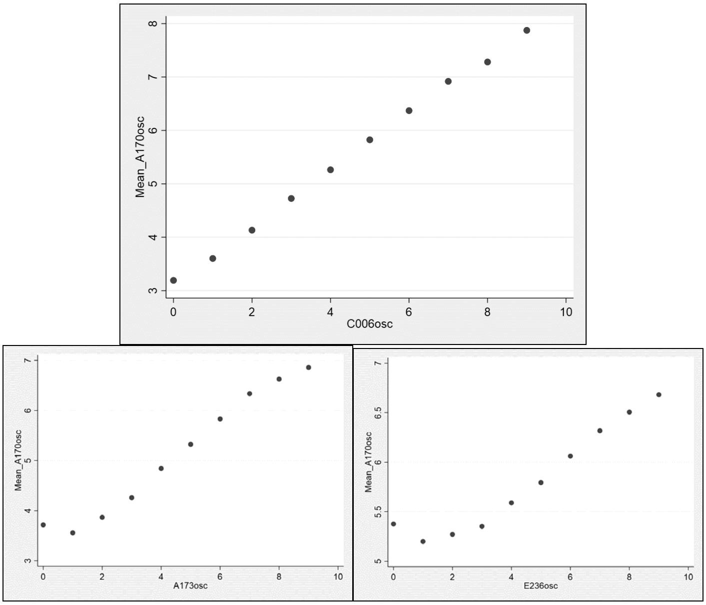

Figure 6. Two-way graphical representations of the relations between each variable from the core model

The second most potent determinant found is afferent to the level of freedom of choice and control. It means that the people with a higher level for this type of freedom are also more likely, positive influence or the maximum recoded value of 9 for A173osc - the right side of Figure 5, to be satisfied with their lives. The latter is in line with the findings of other scholars (Abbott *et al.,* 2016; Ngamaba, 2016; Li *et al.,* 2017) and contributes to the validation of H1.

The third overpowering determinant relates to considerations about democracy. E236osc corresponds to the perceived level of democracy in own country, and it is a positive influence, the maximum recoded value of 9, the right side of Figure 5. E235 (model 4 in Table A5) also indicates the importance of democracy, as reflected in the WVS survey responses. The latter is also positively correlated with the response variable and shows that people more inclined to declare the overall importance of democracy are also more likely to be satisfied with their lives. These two findings are compatible with other similar discoveries from the scientific literature (Orviska *et al.,* 2014; Loubser and Steenekamp, 2017). The latter means a complete validation of H3.

The specific features of some countries will be the object of future research on the same topic. For instance, a dummy variable referring to ex-communist countries or not (Homocianu *et al.,* 2022), some country-dependent measures of economic activity such as GDP or the ratio between Stock Market Capitalization and GDP (Gross Domestic Product) defined in The World Bank Data Catalog, or even the Worldwide Governance Indicators defined by Kaufmann *et al.* in 2010 and used in many other studies including recent ones (Abegaz *et al.,* 2023; Antón *et al.,* 2023).

The reverse causality checks indicated only three determinants (a triad) from the penta-core model, namely the satisfaction with the household financial situation (C006osc), the level of freedom of choice and control (A173osc), and the perceived level of democracy in own country (E236osc), in this specific order given by the descending order of magnitude of effects corresponding to these three (Figure 5).

A limitation of this study is the impossibility of applying the obtained models to a specific list of countries. For instance, the core model with three determinants does not apply to respondents from Israel (no responses for variables A173, and C006). The same happens in the case of 16 other countries from a total of 108, namely Albania, Bosnia-Herzegovina, Croatia, Dominican Republic, El Salvador, Israel, Kuwait, Latvia, Lithuania, Montenegro, Qatar, Saudi Arabia, Uganda, North Macedonia, Tanzania, and Uzbekistan (this time, no responses for E236).

4. CONCLUSIONS

Following an exploratory approach to World Values Survey data, including the latest and the most comprehensive dataset (version 4.0, time-series 1981-2022), an accurate model with three strong non-redundant determinants fully supported by most data emerged. It indicates that a specific type of respondent is more likely to be satisfied with life. The latter means being more content with the household financial situation, showing a superior level of freedom of choice and control, and exhibiting more exposure to democracy and democratic values. Prediction nomograms presented in this article acted as powerful ranking instruments and probability identification tools. All conclusions related to the identified determinants stand on a triad logit model with a good classification accuracy (AUC-ROC greater than 0.8) and a high maximum probability (more than 95% for the most advantageous combination of values for those three most resilient identified determinants). They resulted after performing initial value treatment and, in addition, many selection rounds, and robustness tests, including three types of cross-validation and overfitting, collinearity, and reverse causality checks.

References

1) Abbott, P., Wallace, C., Lin, K. et al. (2016). The Quality of Society and Life Satisfaction in China. *Social Indicators Research,* 127, pp. 653–670.
2) Abegaz, M.B., Debela, K.L. and Hundie, R.M. (2023). The effect of governance on entrepreneurship: from all income economies perspective. *J Innov Entrep,* 12, pp. 1.

3) Ahrens, A., Hansen, C.B. and Schaffer, M.E. (2020). Model selection and prediction with regularized regression in Stata. *The Stata Journal: Promoting communications on statistics and Stata*, 20(1), pp. 176–235.
4) Antón, JI., Grande, R. and Muñoz de Bustillo, R. (2023). Gender Gaps in Working Conditions. *Social Indicators Research*, 166(1), pp. 53–83.
5) Argan, M., Argan, M.T. and Dursun, M.T. (2018). Examining relationships among well-being, leisure satisfaction, life satisfaction, and happiness. *Int J Med Res Health Sci,* 7(4), pp. 49-59.
6) Barrington-Leigh, C.P. (2021). Life satisfaction and sustainability: a policy framework. *SN Social Sciences* 1, pp. 176.
7) 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*, 74, pp. 333–351.
8) Bartram, D. (2020). Age and Life Satisfaction: Getting Control Variables under Control. *Sociology*, 1(17), pp.421-437.
9) Becchetti, L. and Conzo, G. (2022). The Gender Life Satisfaction/Depression Paradox. *Social Indicators Research*, 160(1), pp. 35–113.
10) Bittmann, F. (2021). Beyond the U-shape: Mapping the functional form between age and life satisfaction for 81 countries utilizing a cluster procedure. *Journal of Happiness Studies*, 22(5), pp. 2343-2359.
11) Boyce, C.J., Brown, G.D. and Moore, S.C. (2010). Money and happiness: Rank of income, not income, affects life satisfaction. *Psychological science*, 21(4), pp. 471-475.
12) Brzozowski, M. and Spotton Visano, B. (2020). "Havin'Money's not everything, not havin'it is": The importance of financial satisfaction for life satisfaction in financially stressed households. *Journal of Happiness Studies,* 21, pp. 573-591.
13) Chen, Y.-K., Wenlong, L. and Xiaofeng, T. (2008). *Parallelization of AdaBoost algorithm on multi-core processors.* [online] Available at: https://ieeexplore.ieee.org/document/4671775 [Accessed 01.04.2024].
14) DeBruine, L.M. and Barr, D.J. (2021). Understanding Mixed-Effects Models Through Data Simulation. *Advances in Methods and Practices in Psychological Science,* 4(1), pp. 1–15.
15) Denton, P.H. (2015). The End of Democracy. *Essays in Philosophy*, 16(1), pp. 70–88.
16) Diener, E. and Diener, M. (2009). Cross-Cultural Correlates of Life Satisfaction and Self-Esteem. *Culture and Well-Being*, 68(4), pp. 71–91.
17) Dirzyte, A., Patapas, A. and Perminas, A. (2022). Associations between Leisure Preferences, Mindfulness, Psychological Capital, and Life Satisfaction. *International Journal of Environmental Research and Public Health,* 19(7), pp. 21-41.
18) Dorn, D., Fischer, J.A.V., Kirchgässner, G. and Sousa-Poza, A. (2007). Direct democracy and life satisfaction revisited: new evidence for Switzerland. *Journal of Happiness Studies*, 9(2), pp. 227–255.
19) Feldman, F. (2008). Whole Life Satisfaction Concepts of Happiness. *Theoria*, 74(3), pp. 219–238.

20) Gundelach, P. and Kreiner, S. (2004). Happiness and Life Satisfaction in Advanced European Countries. *Cross-Cultural Research*, 38(4), pp. 359–386.
21) Halvorsen, I. and Heyerdahl, S. (2006). Girls with anorexia nervosa as young adults: Personality, self-esteem, and life satisfaction. *International Journal of Eating Disorders,* 39(4), pp. 285–293.
22) Hellevik, O. (2017). The U-shaped age–happiness relationship: Real or methodological artifact? *Quality & Quantity*, 51, pp. 177-197.
23) Homocianu, D. and Tîrnăucă, C. (2022). MEM and MEM4PP: New Tools Supporting the Parallel Generation of Critical Metrics in the Evaluation of Statistical Models. *Axioms,* 11(10), pp. 549.
24) Homocianu, D. and Airinei, D. (2022). PCDM and PCDM4MP: New Pairwise Correlation-Based Data Mining Tools for Parallel Processing of Large Tabular Datasets. Mathematics, 10(15), pp. 26-71.
25) Homocianu, D., Dospinescu, O. And Sireteanu, NA. (2022). Exploring the Influences of Job Satisfaction for Europeans Aged 50 + from Ex-communist vs. Non-communist Countries. *Social Indicators Research,* 159, pp. 235–279.
26) Irandoukht, A. (2021). Optimum ridge regression parameter using R-squared of prediction as a criterion for regression analysis. *Journal of Statistical Theory and Applications,* 20(2), pp. 242.
27) Jann, B. (2005). Making regression tables from stored estimates. *Stata J.,* 5, pp. 288–308.
28) Jann, B. (2007). Making regression tables simplified. *Stata J.,* 7, pp. 227–244.
29) Kaiser, M., Otterbach, S. and Sousa-Poza, A. (2022) Using machine learning to uncover the relation between age and life satisfaction. *Sci Rep,* 12, p. 5263.
30) Karabulut, E.M. and Ibrikci, T. (2014). Analysis of Cardiotocogram Data for Fetal Distress Determination by Decision Tree-Based Adaptive Boosting Approach, *Journal of Computer and Communications,* 2, pp. 32-37.
31) Kassenboehmer, S. C. and Haisken-DeNew, J. P. (2012). Heresy or enlightenment? The well-being age U-shape effect is flat. *Economics Letters,* 117(1), pp. 235-238.
32) Kaufmann, D., Kraay, A. and Mastruzzi, M. (2010). *The worldwide governance indicators: Methodology and analytical issues.* [online] Available at: https://www.worldbank.org/en/publication/worldwide-governance-indicators [Accessed 01.04.2024].
33) Keane, L., Pacek, A. and Radcliff, B. (2012). Organized labor, democracy, and life satisfaction: A cross-national analysis. *Labor Studies Journal*, 37(3), pp. 253-270.
34) Keng, K.A., Jung, K., Jiuan, T.S. and Wirtz, J. (2000). The influence of materialistic inclination on values, life satisfaction and aspirations: An empirical analysis. *Social Indicators Research,* 49, pp. 317-333.
35) Koohbanani, S.E., Dastjerdi, R., Vahidi, T. and Far, M.H.G. (2013). The Relationship Between Spiritual Intelligence and Emotional Intelligence with Life Satisfaction Among Birjand Gifted Female High School Students. *Procedia-Social and Behavioral Sciences,* 84, pp. 314-320.
36) Lai, K. (2020). Using Information Criteria Under Missing Data: Full Information Maximum Likelihood Versus Two-Stage Es-timation. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(2), pp. 278–291.

37) Li, C., Wang, S., Zhao, Y., Kong, F. and Li, J. (2017). The Freedom to Pursue Happiness: Belief in Free Will Predicts Life Satisfaction and Positive Affect among Chinese Adolescents. *Frontiers in Psychology*, 7, p. 2027.
38) Loubser, R. and Steenekamp, C. (2017). Democracy, well-being, and happiness: A 10-nation study. *Journal of Public Affairs*, 17, pp. 1-2.
39) Lu, L. (1999). Personal or environmental causes of happiness: A longitudinal analysis. *The Journal of Social Psychology*, 139(1), pp. 79-90.
40) Łużyński, W. (2019). The Moral Principles of Democracy: Reflections based on Catholic Social Teaching, *Cuadernos Salmantinos de Filosofía*, 46, pp. 309-322.
41) Mangels, D. (2009). The science of happiness. *Berkeley Scientific Journal*, 12(2).
42) Meléndez, J.C., Tomás, J.M., Oliver, A. and Navarro, E. (2009). Psychological and physical dimensions explaining life satisfaction among the elderly: A structural model examination. *Archives of Gerontology and Geriatrics*, 48(3), pp. 291–295.
43) Ng, W., Russell Kua, W.S. and Kang, S.H. (2019). The relative importance of personality, financial satisfaction, and autonomy for different subjective well-being facets. *The Journal of Psychology*, 153(7), pp. 680-700.
44) Ng, Y.K. (2022). Happiness or Life Satisfaction? In: Ng, Y.K. eds., *Happiness—Concept, Measurement and Promotion.* Singapore: Springer.
45) Ngamaba, K.H. (2016). Happiness and life satisfaction in Rwanda. *Journal of Psychology in Africa,* 26(5), pp. 407-414.
46) Oishi, S., Diener, E., Lucas, R.E. and Suh, E.M. (2009). Cross-Cultural Variations in Predictors of Life Satisfaction: Perspectives from Needs and Values. In: Diener, E. eds., *Culture and Well-Being*, Dordrecht: Springer.
47) Orviska, M., Caplanova, A. and Hudson, J. (2014). The impact of democracy on well-being. *Social indicators research,* 115, pp. 493-508.
48) Owen, A.L., Videras, J. and Willemsen, C. (2008). Democracy, Participation, and Life Satisfaction. *Social Science Quarterly*, 89(4), pp. 987–1005.
49) Pacek, A. and Radcliff, B. (2008). Assessing the Welfare State: The Politics of Happiness. *Perspectives on Politics*, 6(2), pp. 267-277.
50) Rajani, N.B., Skianis, V. and Filippidis, F.T. (2019). Association of environmental and sociodemographic factors with life satisfaction in 27 European countries. *BMC Public Health*, 19(1), pp. 1-8.
51) Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schrö-der, B., Thuiller, W., Warton, D.I., Wintle, B.A., Hartig, F. and Dormann, C.F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography,* 40, pp. 913–929.
52) Roth, B., Hahn, E. and Spinath, F.M. (2017). Income inequality, life satisfaction, and economic worries. *Social Psychological and Personality Science*, 8(2), pp. 133-141.
53) Ruggeri, K., Garcia-Garzon, E. and Maguire, Á. (2020). Well-being is more than happiness and life satisfaction: a multidimensional analysis of 21 countries. *Health Qual Life Outcomes*, 18, pp. 192.
54) San Martín, J., Perles, F. and Canto, J.M. (2010). Life satisfaction and perception of happiness among university students. *The Spanish Journal of Psychology*, 13(2), pp. 617-628.

55) Schober, P., Boer, C. and Schwarte, L.A. (2018). Correlation Coefficients. *Anesthesia & Analgesia*, 126(5), pp. 1763–1768.
56) Schröder, M. (2016). How income inequality influences life satisfaction: Hybrid effects evidence from the German SOEP. *European Sociological Review*, 32(2), pp. 307-320.
57) Silbermann, I. (1985). On "happiness". *The Psychoanalytic Study of the Child*, 40(1), pp. 457-472.
58) Sirgy, M.J., Gurel-Atay, E., Webb, D., Cicic, M., Husic-Mehmedovic, M., Ekici, A. and Johar, J.S. (2013). Is materialism all that bad? Effects on satisfaction with material life, life satisfaction, and economic motivation. *Social Indicators Research*, 110, pp. 349-366.
59) Sohier, L. (2018). Do involuntary longer working careers reduce well-being? *Applied Research in Quality of Life,* 14, pp. 171–196.
60) Suldo, S.M., Riley, K.N. and Shaffer, E.J. (2006). Academic correlates of children and adolescents' life satisfaction. *School Psychology International*, 27(5), pp. 567-582.
61) Sundriyal, R. and Kumar, R. (2014). Happiness and wellbeing. *The International Journal of Indian Psychology*, 1(4).
62) Supervía, P.U., Bordás, C.S. and Lorente, V.M. (2020). Exploring the Psychological Effects of Optimism on Life Satisfaction in Students: The Mediating Role of Goal Orientations. *International Journal of Environmental Research and Public Health*. 17(21), pp. 78-87.
63) Tsikriktsis, N. (2005). A review of techniques for treating missing data in OM survey research. *Journal of Operations Management*, 24, pp. 53–62.
64) Vatcheva, K.P., Lee, M., McCormick, J.B. and Rahbar, M.H. (2016). Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. *Epidemiology (Sunnyvale),* 6(2), pp. 227.
65) Veenhoven, R. (1996). The study of life-satisfaction. [online] Available at: http://hdl.handle.net/1765/16311 [Accessed 01.04.2024].
66) Veenhoven, R. (2012). Happiness: Also Known as "Life Satisfaction" and "Subjective Well-Being". In: Land, K., Michalos, A., Sirgy, M.J. (eds) *Handbook of Social Indicators and Quality of Life Research*. Dordrecht: Springer.
67) Verme, P. (2009). Happiness, freedom and control. *Journal of Economic Behavior & Organization,* 71, pp. 146–161.
68) Williams, G. (2011). *Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery*, New York: Springer, pp. 269-291.
69) Xiao, J.J., Tang, C. and Shim, S. (2009). Acting for happiness: Financial behavior and life satisfaction of college students. *Social Indicators Research*, 92, pp. 53-68.
70) Yin, K. (2022). On the Human Rights Implications of the Whole-Process People's Democracy. *The Journal of Human Rights*, 21(1), pp. 91.
71) Zhang, F., Zhang, C. and Hudson, J. (2018). Housing conditions and life satisfaction in urban China. *Cities*, 81, pp. 35-44.
72) Zlotnik, A. and Abraira, V. (2015). A general-purpose nomogram generator for predictive logistic regression models. *The Stata Journal: Promoting communications on statistics and Stata,* 15, pp. 537–546.