



Faculty of Economics, University of Niš
18 October 2018

49th International Scientific Conference
**QUANTITATIVE AND QUALITATIVE
ANALYSIS IN ECONOMICS**

USING FACTOR ANALYSIS AND MULTIPLE LINEAR REGRESSIONS: AN EXAMPLE IN SATISFACTION WITH LIFE IN SLOVENIA

Anton Vorina, PhD*

Abstract: *This paper presents the research about satisfaction with life in Slovenia. There were analysed the impact of gender, age, level of education, employment status and the amount of salary on the satisfaction with life. The main goal of the research is to determine whether (and how) the impact of five independent variables (gender, age, level of education, employment status and the amount of salary) influences the satisfaction with life (dependent variable). A written "online" survey (with Google-drive) was conducted from 13/4/2017 to 7/5/2017. Based on the multilinear regression model (R- square is only 0,037) we found out that there is no statistically correlation between gender, age, level of education, employment status, the amount of salary and the satisfaction with life.*

Keywords: *life, satisfaction, multiple linear regression, factor analysis, salary, gender.*

1. Introduction

There is a research on satisfaction with life (hereinafter referred to as SWL) or well-being life in Slovenia. We have research questions: *Is there any correlation between gender, age, level of education, employment status, amount of salary and satisfaction with life in Slovenia?*

We're assuming:

Main hypothesis: *The satisfaction with life depends on five variables (gender, amount of salary, age, level of education, employment status).*

In the theoretical part, we used secondary sources from data collection, such as google.com, Google scholar and the "online" library. For the research part, we used the questionnaire, which was analysed using the SPSS program. We're limited only to Slovenia citizens.

* Vocational College of Economics, Celje, Slovenia; ✉ anton.vorina@guest.arnes.si
UDC 519.2:316.6(497.4)

2. Literature review

Satisfaction is a pleasant emotion associated with quality of life. Otherwise, you can survive without the satisfaction and happiness, but that life is miserable and without any real quality. People are looking for satisfaction and they strive for satisfaction and happiness. It is very important that we feel satisfied, fulfilled and happy (Vorina, A. and N. Vorina, 2013)

“Happiness” began gaining importance as a field in economics in the 1970s, particularly with the work of Easterlin (1974, Kacapyr, 2008).

Oswald (1997) reports the general conclusions of country-specific studies: income contributes, but only marginally, to happiness; unemployment lowers happiness dramatically; happiness is positively related to marriage, education, being female, being white and being self-employed; happiness declines until an adult reaches age 30 and then begins to increase with age.

Borooah’s (2006) study indicates that self-reported happiness depends on religion, income, employment, family and social life, and good health.

Easterlin (1995) finds that income is not a significant determinant of happiness in the USA with a correlation coefficient of 0.2.

Kacapyr (2008) finds that social and economic factors explain about 66 % of the variance in self-reported happiness across nations. Respondents in former socialist-bloc countries report surprisingly low levels of SWL given their circumstances while Latin Americans report higher-than-expected levels of happiness.

The Satisfaction With Life Scale (SWLS) is a global measure of happiness, defined as the cognitive, judgmental component of satisfaction with life. Research indicates that job satisfaction is significantly related to life satisfaction. Since the 1930s there has been a great deal of interest in the relationship between employee’s well-being and productivity.

The results of the study (Rathi, 2011) demonstrated a positive relationship between psychological well-being and organizational commitment and its components, namely affection, continuance, and normative commitment.

Vorina (2013) study (n=1006) shows that the engagement of employees would increase if the satisfaction with life increases.

Hersey (1932, Zelenski et. al, 2008) reported a positive relationship between daily emotions and performance.

3. Methodology

Sample. We chose to investigate the population – people (from 16 to 73 years old) from Slovenia (n=336). The study included 117 (34.8 %) men and 219 (65.2%) women. There were 8 respondents (2.4 %) with completed primary school level of education or less, 162 respondents (48.2 %) with completed secondary school level of education and 166 people (49.4 %) with completed vocational school level of education or more. The average age of respondents is 44.5 years, standard deviation is 11.968 years. The average monthly income is 825,12 euros.

**Using Factor Analysis and Multiple Linear Regressions:
An Example in Satisfaction with Life in Slovenia**

Questionnaire. This questionnaire was made by researches. The investigation started on the 13th of April, and was finished on the 7th of May 2017. The questionnaire was anonymous and took time less than five minutes. The questionnaire consists of 10 questions, 5 are about the satisfaction with life and the other questions are about the demographical data (gender, amount of the salary, age, level of education, employment status). The first 5 questions were based on statements that were measured in a scale 1-7 (1 - strong disagreement, to 7 - strong agreement), and the other five questions were individually answered by the respondents. The satisfaction with life (SWL) was measured by SWLS (Diener et al, 1985) consists of 5-items: In most ways my life is close to my ideal; The conditions of my life are excellent; I am satisfied with my life; So far I have got the important things I want in life; If I could live my life over, I would change almost nothing.

Used methods, data processing. The collected data were analysed using IBM SPSS, version 20. There were also used the Microsoft tools Word and Excel. Regarding the purpose and objectives of the research we used descriptive statistics, *Multiple Linear Regression and Factor Analysis*.

Verification of the adequacy of the measurement instrument. The reliability of a questionnaire, which measure SWL is Cronbach's alpha 0.870 (Tables 1 and 2), which means a lot or exemplary reliability of measurement (Table 3).

Table 1: Case Processing Summary, SWL

	N	%
Valid	336	100,0
Cases excluded	0	0,0
Total	336	100,0

Source: SPSS 20, Author's creation

Table 2: Reliability Statistics, SWL

Cronbach's Alpha	N of items
0.870	5

Source: SPSS 20, Author's creation

Table 3: Internal Consistency Using Cronbach's Alpha

Cronbach's alpha	Internal consistency
$0.9 \leq \alpha$	Excellent
$0.8 \leq \alpha < 0.9$	Good
$0.7 \leq \alpha < 0.8$	Acceptable
$0.6 \leq \alpha < 0.7$	Questionable
$0.5 \leq \alpha < 0.6$	Poor
$\alpha < 0.5$	Unacceptable

Source: Cortina, J.M. (1993)

4. Factor Analysis for questionnaire “SWL”

Factor analysis is a technique that is used to reduce a large number of variables into fewer numbers of factors. This technique extracts maximum common variance from all variables and puts them into a common score (<http://www.statisticssolutions.com/factor-analysis-sem-factor-analysis/>).

The Kaiser-Meyer-Olkin Measure of Sampling Adequacy is a statistic that indicates the proportion of variance in your variables that might be caused by underlying factors. High values (close to 1.0) generally indicate that a factor analysis may be useful with your data. If the value is less than 0.50, the results of the factor analysis probably won't be very useful.

Bartlett's test of sphericity tests the hypothesis that your correlation matrix is an identity matrix, which would indicate that your variables are unrelated and therefore unsuitable for structure detection. Small values (less than 0.05) of the significance level indicate that a factor analysis may be useful with your data. (https://www.ibm.com/support/knowledgecenter/en/SSLVMB_sub/spss/tutorials/fac_telco_kmo_01.html). The value of the indicator is greater than 0.5 and KMO by the risk of less than 0.05 to reject the null hypothesis that the correlation matrix (Table 5) is the same as array unit (Bartlett's test of sphericity, Table 4).

Table 4: KMO and Bartlett's test for the questionnaire “SWL”

Kaiser-Meyer-Olkin test		0,842
Bartlett's Test	885,311	3959,496
	10	66,000
	0,000	0,000

Source: SPSS 20, Author's creation

Table 5: Correlation matrix between variables measuring the SWL

	If I could live my life over, I would change almost nothing.	So far I have gotten the important things I want in life	I am satisfied with my life.	The conditions of my life are excellent	In most ways my life is close to my ideal.
If I could live my life over, I would change almost nothing.	1	0,561	0,518	0,465	0,533
So far I have gotten the important things I want in life	0,561	1	0,561	0,485	0,571
I am satisfied with my life.	0,518	0,561	1	0,675	0,705
The conditions of my life are excellent	0,465	0,485	0,675	1	0,773
In most ways my life is close to my ideal.	0,533	0,571	0,705	0,773	1

Source: SPSS 20, Author's creation

**Using Factor Analysis and Multiple Linear Regressions:
An Example in Satisfaction with Life in Slovenia**

Table 5 gives a correlation matrix between variables. From the size of the correlation coefficients there can be seen a medium strong correlation between all the variables.

In Table 6 presents Eigenvalues only to a single factor. *The first factor has the inherent value greater than one and is equal to 3.352. It explains 67,035 % of the total variance.* Therefore, we have got only one factor, which means that the questionnaire is exclusively used for measuring satisfaction with life, and that it cannot be carried out hereinafter referred to as the Varimax method.

Table 6: Eigenvalues and explain the variance using the “SWL”

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3,352	67,035	67,035	3,352	67,035	67,035
2	,663	13,252	80,287			
3	,439	8,778	89,065			
4	,330	6,592	95,657			
5	,217	4,343	100,000			

Extraction Method: Principal Component Analysis.

Source: SPSS 20, Author's creation

5. Findings of the research-Multiple linear regression models

The main hypothesis states *The satisfaction with life depends on five variables (gender, amount of salary, age, level of education, employment status).* SWL was measured as a summary of five items as noted above (maximum score is 35, minimum score is 5) and was called Summary Satisfaction With Life (SSWL). We calculate a new variable in SPSS program using the technique *Compute a new variable* SSWL. The dependent variable (Y-SSWL) was measured as the sum of 5 factors. The main hypothesis was tested by *multiple linear regression.*

Variables are defined as follows:

Y- A depend variable is SSWL (Summary satisfaction with life).

Independent variables are:

X₁ – Amount of salary (in euros)

X₂ – Dummy gender; X₂=1 for a man, X₂=0 for women,

X₃ – Age (in years)

X₄ – Dummy education; X₄=1 for finished primary school or less, X₄=0 for others (finished secondary school or more),

X₅ – Dummy employment; X₅=1 for respondents who have a job on contract, X₅=0 for others.

Let's investigate Pearson coefficients of correlation between pair of variables, see the Table 7.

The strongest negative correlation coefficient appears for the following pairs:

X₁ and X₂ with $r_{X_1, X_2} = -0.108$.

The strongest linear positive correlation appears for:

X₁ and X₅ with $r_{X_1, X_5} = 0.506$,

X₄ and X₅ with $r_{X_4, X_5} = 0.444$.

Table 7: Correlation matrix for K=5, n=336

	X ₁	X ₂	X ₃	X ₄	X ₅	Y
X ₁	1					
X ₂	-,108*	1				
X ₃	-,069	,050	1			
X ₄	,358**	,003	-,056	1		
X ₅	,506**	,086	-,070	,444**	1	
Y	-,053	-,013	,048	-,068	,095	1

Source: SPSS 20, Author's creation

In general, multiple regression allows the researcher to ask (and hopefully answer) the general question "what is the best predictor of ...". Of course, the limitation of all regression techniques is that one can only ascertain relationships, but never be sure about the underlying causal mechanism.

For testing the first research hypotheses, we apply a **multiple linear regression analysis**, using software SPSS 20 program. The general multiple linear regression population Model applied (1) is:

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{j,i} + e_i, \quad i = 1, 2, \dots, n \quad (1)$$

Regression parameters are estimated using minimum least squares (OLS) method. The Model with the OLS estimated parameters (2) is:

$$\hat{y}_i = \hat{\beta}_0 + \sum_{j=1}^k \hat{\beta}_j x_{j,i}, \quad i = 1, 2, \dots, n \quad (2)$$

The random error term (3) is estimated using:

$$\hat{e}_i = y_i - \hat{y}_i \quad (3)$$

The most important limit of the regression method is that only probable relationships are assumed, and one may never be sure about the true causal mechanism that lies behind. Regression diagnostics was conducted using coefficient of determination R^2 , the t -test for testing the significance of each of the independent variables, and F -test in testing for overall regression (Vorina, 2012, 406). The model assumption was tested as well. Results for multiple regression models are given in Tables 9. The multiple OLS regression model was analysed, for studying the following:

**Using Factor Analysis and Multiple Linear Regressions:
An Example in Satisfaction with Life in Slovenia**

Model I- Impact of K=5 regressors on dependent variable Y, (n=336). The linear multiple regression Model I with estimated parameters is:

$$(\hat{Y} = 22,171 + 0,003 * X_1 - 0,533 * X_2 - 0,061 * X_3 + 2,127 * X_4 - 1,595 * X_5)$$

(1,068) (0,001) (0,738) (0,035) (2,265) (0,785)

n = 336, R-squared = 0.037, Adjusted R-squares =0.022, standard Error=6,301

Table 8: Regression Model I, Coefficients, K=5, n=336

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	22,171	1,068		20,765	0		
	Please, write the amount of your salary:	0,003	0,001	0,21	3,112	0,002	0,648	1,544
	dummy sex	-0,533	0,738	-0,04	-0,722	0,471	0,958	1,044
	Your age:	-0,061	0,035	-0,115	-1,778	0,076	0,699	1,431
	dummy education	2,127	2,265	0,051	0,939	0,348	0,991	1,009
	dummy employment	-1,595	0,785	-0,125	-2,032	0,043	0,78	1,282

a. Dependent Variable: Summary SSWL

Source: SPSS 20, Author's creation

Look at P-value=0,002 and find out that only the amount of salary is statistically significant at 5% level, the other variables are not statistically significant because their value is higher than 0,05.

In Model I (F (5,333) = 2.528, p-value=0.029) only 3.7 % of total sum of squares is explained by the estimated model (Table 9 and Table 10).

Table 9: Regression Model: K=5, n=336

R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
0,193 ^a	0,037	0,022	6,301	1,991

- a. Predictors: (Constant), dummy employment, dummy sex, dummy education, Your age:, Please, write the amount of your salary:
b. Dependent Variable: Summary SSWL

Table 10: Regression Model, F-test

Model	Sum of Squares	df	F	Sig.
Regres.	501,826	5	2,528	0,029
Resid.	13021,626	328		
Total	13523,452	333		

The main hypothesis is rejected, because the model explains only 3.7% of the total sum of square.

The other four test: The Jarque-Berra test for normality of residuals, The White heteroscedasticity LM test and Breusch-Pagan-Godfrey, LM test for homoscedasticity, the Variance Inflation Factor (VIF) for judging about multicollinearity. The Durbin-Watson test of the first order autocorrelation in residuals was not conducted because there is only one variable (amount of salary) which is statistically significant and also the model is very poor (R-square is only 3.7%).

7. Conclusion

Life satisfaction has been conceptualized as a cognitive constituent of subjective well-being. High satisfaction suggests that the quality of life, among the population concerned, is good.

We investigated the satisfaction with life in Slovenia. We made online questionnaires (see annex) which consisted of 10 questions. The questionnaire started on the 13th of April 2017, and finished on the 7th of May 2017. The total number of the respondents was 336.

We rejected hypothesis in the research, because we found out that there isn't a statistically significant correlation between the satisfaction with life and five independent variables as mentioned.

The questionnaire was fulfilled mainly by students, who live with their parents or are financed by them, hereby their level of life satisfaction cannot be objective. For future research can be interesting to explore this area by making a bigger questionnaire for different people of different ages and from countries. In the future research, it would be interesting to study whether there are statistically significant differences between people from other countries (Serbia, Croatia).

References

- Borooah, V.K. (2006). How much happiness is there in the world? A cross-country study, *Applied Economics Letters*, 13 (8), 483–488.
- Cortina, J.M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*. 78: 98–104.
- Diener, E., Emmons, R. A., Larsen, R. J., and Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, 49, 71–75.
- Easterlin, R. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization*, 27 (1): 35–47.

Using Factor Analysis and Multiple Linear Regressions: An Example in Satisfaction with Life in Slovenia

- IBM knowledge center, KMO and Bartlett's Test, Retrieved from: https://www.ibm.com/support/knowledgecenter/en/SSLVMB_sub/spss/tutorials/fac_telco_kmo_01.html, Accessed on 4 January 2018.
- Kacapyr, E. (2008). Cross-country determinants of satisfaction with life. *International Journal of Social Economics*, 35 (6): 400–416.
- Oswald, A. (1997). Happiness and economic performance. *The Economic Journal*, 107 (445): 1815–1831.
- Psychological Well-Being and Organizational Commitment: Exploration of the Relationship*, Retrieved from: <http://www.amrita.edu/sdg/pdf/ASB%20Working%20paper%20series/MBA%20book%20106-2011%20Psychological%20Well-Being%20and%20Organizational%20Commitment%20Exploration%20of%20the%20Relationship.pdf>, Accessed on 18 December 2017.
- Rathi, N. (2011). Psychological Well-Being and Organizational Commitment: Exploration of the Relationship. *Working Paper No.106/2011*, Retrieved from: <http://www.amrita.edu/sdg/pdf/ASB%20Working%20paper%20series/MBA%20book%20106-2011%20Psychological%20Well-Being%20and%20Organizational%20Commitment%20Exploration%20of%20the%20Relationship.pdf>, Accessed on 26 August 2017.
- Statistics Solutions: Factor Analysis*, Retrieved from: <http://www.statisticssolutions.com/factor-analysis-sem-factor-analysis/>, Accessed on 4 January 2018.
- Vorina, A. (2012). Hedonic regression model for estimate price for a vehicle brand. *Proceedings of the 9th International Conference on Logistics & Sustainable Transport*. 403–409.
- Vorina, A. (2013). The relationship between satisfaction with life and employee engagement, *Journal of Process Management – New Technologies, International*, 1 (2): 77–81.
- Vorina, A. Vlasova, M. (2017). Satisfaction with life: comparison between Slovenia and Russia. *Contemporary Approaches in the Analysis of Economic Performances: [proceedings of the international scientific conference]*, *Prolom Spa*. October 11–12.
- Vorina, A. and N., Vorina. (2013). The impact of gender, level degree of education and the amount of salary on the satisfaction with life and political culture - an example of Slovenia, *Journal of Process Management*, 4 (1), 122–127.
- Zelenski, J. Murhpy. S. A. & Jenkins D.A. (2008). The Happy-Productive Worker Thesis Revisited. *Springer Science+Business Media B.V.* 521-537. Retrieved from: http://www.sel.eesc.usp.br/informatica/graduacao/material/etica/private/the_happy-productive_worker_thesis_revisited.pdf, Accessed on 25 December 2018.

KORIŠČENJE FAKTORSKE ANALIZE I MULTILINEARNE REGRESIJE NA PRIMERU ZADOVOLJSTVA ŽIVOTOM U SLOVENIJI

Rezime: Rad opisuje istraživanja u vezi sa zadovoljstvom životom u Sloveniji. Analizirani su uticaj roda, starosti, nivoa obrazovanja, statusa zaposlenja i visine zarade na zadovoljstvo životom. Glavni cilj istraživanja je da utvrdi da li (i kako) uticaj pet nezavisnih varijabli (rod, starost, nivo obrazovanja, status zaposlenja i visina plate) utiče na zadovoljstvo životom (zavisna varijabla). Pisana "on-lajn" anketa (sa Google-diskom) je sprovedena od 13.4.2017. do 7. 5. 2017. godine. Na osnovu multilinearne regresionog modela (R -kvadrat je samo 0,037), otkrili smo da nema statističke korelacije između roda, starosti, nivoa obrazovanja, statusa zaposlenja, visine plate i zadovoljstva životom.

Ključne reči: život, zadovoljstvo, multilinearne regresija, faktorska analiza, plata, rod.