Evaluation of the Employees' Current Salary: Multivariate analysis

Svjetlana Vranješ

Faculty of Economics, University of Banja Luka, svjetlana.vranjes@ef.unibl.org

Abstract. Using a sample size of 200 R&D employees, this paper examines the relationship between the current salary and starting salary, previous experience, education, employees' age and gender. The results provided by this study show that current salary is positively associated with employees' salary at the beginning of the career and years of education. The author finds strong evidence that current salary is negatively associated with employees' age, previous experience and gender. Furthermore, conducting cluster analysis, results provide two different groups. The first group consists of employees who are more likely to be included in the clerical type of jobs and the second group is specific to the other types of job.

Key words: salary, standard of living, job satisfaction, expertise, multivariate analysis.

INTRODUCTION

The first factor connected to standard of living and quality of life is the amount of money or salary. Salary is a component that influences employee's satisfaction regarding their job and it is a predictor of life as well. The aim of this study is to consider salary as one of the highly important factors of standard of living.

Job satisfaction was the most common concept used by authors to test the different relationships which could appear in the analysis of the employee data (Lee & Wilbur, 1958; Roy, 2017; Huang et al. 2016; Khamisa, 2015; Fu & Deshpande, 2014). Lee and Wilbur (1958) investigate a relation between employee's age and job satisfaction. The main aim of their study lies in the observation how the different variables such as salary, education, tenure and job attributes influence the relationship between employees' age and job satisfaction. This article presents the results where is a positive and statistically significant relationship between job satisfaction and worker age. Beattie and Spencer (1971) have conducted a study to test the influence of the different variables on the salary attainment. These variables include the age, seniority, education and informal factors of men who work in the bureaucracy of the Canadian federal administration. This study provides the results of a positive connection between salary and expertise (education), age and seniority as the independent variables. In their study, Graddy and Pistaferri (2002) examined the outgoing salaries to test the differences in the starting salary of women and men graduates who started working in the United Kingdom. Their results present the unexplained differential which is equal to 8.6 percent. Therefore, the main aim of their study was to compare these results gained from women and men graduates of London Business School with other results obtained in the UK.

The main purpose of this research is to examine how the current salary is related to starting salary (salary at the beginning of the career), employees' age, gender, previous experience and education.

RESEARCH DESIGN

In this study, the intention is focused on researching how every value of the independent variable is associated with a value of current salary. As there are more than one independent variable (explanatory variable), multiple linear regression is the most suitable model for this study. The main advantage of multiple linear regression is the usage of several explanatory variables to predict current salary which presents the depend-

ent variable. Since the ANOVA could provide information regarding variability within a regression model, this study includes the ANOVA calculation for multiple regression. What we could also use in this analysis and interpretation of the results is MANOVA. Multivariate analysis of variance is specific in the situation when there are two or more dependent variables related to one or more independent or predictor variables. Therefore, we could use current salary and category of the job to observe how these dependent variables are influenced by the other independent variables.

Furthermore, this study shows the interesting results in terms of grouping all participants into two groups. The first and the second group consist of participants who have lower and higher average salary, respectively. Since hierarchical clustering needs computation of all distances between the pairs of cases, we decided to use k-means clustering.

DATASET

This study includes several independent variables such as employees' starting salary, age, gender, previous experience, education and job category. There is one dependent variable defined as current salary (or salary now). The quantitative data is collected from the database of the companies who operate in the information technology industry in Germany. Thus, this data consists of information about annual salary, age, gender, work experience and education of R&D employees. According to the sample size, this study analyzes the complete information for 200 R&D employees.

RESULTS AND DISCUSSION

Table 1 presents descriptive statistics and correlations for study variables. There was a strong, positive correlation between dependent variable and worker's salary at the beginning of the career, r=0.898, p=0.000, N=200. There is presented a positive correlation between current salary and years of education as well, r=0.644. p=0.000, N=200. Therefore, both correlations are a high, positive and statistically significant. Overall, increases in current salary were correlated with increases in worker salary at the beginning. Also, increases in current salary are related to increases in the years of education.

Table 1. Intercorrelations among study variables

	M	SD	1	2	3	4	5	6
Salary now	33529.95	16195.59	1.000	0.898**	-0.159*	-0.138*	0.644**	-0.454**
Salary beginning	17274.85	7538.34		1.000	-0.030	-0.120*	0.626**	-0.474**
Previous experience	99.23	104.28			1.000	-0.052	-0.387**	-0.128*
Age	41.97	10.58				1.000	-0.055	0.033
Years of education	13.34	2.986					1.000	-0.381**
Gender	1.46	0.500						1.000

Note. N = 200.
* p < .05   ** p < .01

Correlation coefficients between dependent variable and age (r= -0.138, p=0.025, N=200), previous experiences (r=-0.159, p=0.012, N=200) and gender (r=-0.454, p=0.000, N=200) are lower, negative, but also statistically significant.

Table 2. Analysis of Variance for the Regression Equation

Source	df	SS	MS	F	p
Regression	4	43237058176.698	10809264544.175	235.245	0.000
Residual	195	8960042472.802	45948935.758		
Total	199	52197100649.500			

Conducting the analysis of variance, our data meets the criteria for the use of multiple linear regression due to the F-ratio which is statistically significant (F (4,195) = 235,245, p < 0.01).

According to the statistical theory, R-square tells us the goodness of fit of the model. We can consider it as a percentage. Therefore, knowing the participants' years of education, age, previous experience, salary at the beginning and gender, we can explain 82.8 % difference in current salary. The total variance is statistically significant (F (4,195) = 235.245, p < 0.01).

Table 3. Multiple *regression analysis of* age, previous experience, salary at the beginning, years of education and gender as predictors of current salary

Variable	Beta (β)	t value	P	Model Summary
Salary beginning	0.832	20.157	0.000	R^2=0.83
Previous experience	-0.123	-3.535	0.001	Adjusted R^2= 0.826
Age	-0.040	-1.344	0.180	ΔF= 189.961**
Years of education	0.053	1.188	0.236	
Gender	-0.054	-1.531	0.127	

Table 3 shows the multiple regression analysis of age, previous experience, salary at the beginning, years of education and gender as predictors of dependent variable. The results clearly showed that significant predictors of dependent variable are only salary at the beginning of the career (t (β) = 0.848; p < 0.01) and years of previous experience (t (β) = -0.109; p < 0.01). Thus, current salary is positively related to employees' salary at the beginning of the career and employees' previous experience. Furthermore, age and years of education were not statistically significant predictors of the dependent variable.

In this paper, k-means cluster analysis is used to present a non-hierarchical method of defining clusters. Since it was known how many clusters the author expected to have and there is a moderately sized data set, the most appropriate method for this study is k-means clustering.

Table 4. The division of cases into clusters

	Number of Cases in each Cluster	
Cluster	1	169.000
	2	31.000
Valid		200.000
Missing		0.000

By using cluster analysis, the author divides participants into two independent groups by knowing their salary at the beginning, current salary, previous experience, job category and gender. One of the reasons for the usage of k-means method is the number of groups/clusters that was defined a priori (2 groups). As it is depicted in Table 4, after conducting cluster analysis, there were 169 subjects in the first group and 31 were in the second group.

Table 5. Clusters of variables (current salary, salary at the beginning, previous experience, job category and gender)

	Cluster	
	1	**2**
Salary now	27542.69	66170.16
Salary beginning	14591.36	31904.19
Previous experience	104	72
Job category	1	3
Gender	2	1

The average salary for participants in the first cluster is lower than average salary for participants in the second cluster. Also, salary at the beginning of the career is lower for participants in the first cluster. On the other hand, participants in the first cluster reported a higher level of previous experience which was (due to lower salary rates) a bit unusual/unexpected.

In terms of the job category, the results show that participants from the first cluster are more likely to be employed in "clerical" type of job (work in an office, especially routine documentation and administrative tasks), whilst participants from the second cluster are more likely to be employed in a job category "other". The results in terms of gender depict that participants from the first cluster are more likely to be females, whilst participants from second cluster are more likely to be males.

STATISTICAL POWER AND CONFIDENCE INTERVALS

The subjects of our statistical analysis are also statistical power and confidence intervals. Statistical power is associated with three specific factors such as the magnitude of effects, the sample size and the statistical significance. Examining the dataset, the results present the the power of 1.000. Due to this maximum level, there is a strong possibility to find out significant correlation between explanatory variables and current salary.

Table 6. Confidence intervals

Salary1	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	**Upper Bound**
Gender	1.46	0.035	1.390	1.530
Age	41.97	0.748	40.490	43.440
Years of education	13.34	0.211	12.924	13.756
Salary beginning	17274.85	533.041	16223.717	18325.983
Previous experience	99.23	7.374	84.684	113.766

Table 6 shows the confidence intervals of five variables where 95% CI [40.49, 43.44] presents that the lower level of the confidence is equal to 41.49 and the upper limit for age of participants is 43.44. Therefore, the conclusion is that there is 95% confident that the mean age of participants is between 41 and 43. The interpretation of the other variables could be done in the same way.

CONCLUSION

In this study, multiple linear regression and cluster analysis are applied to test the influences on salary trends. The findings show a great importance of the salary at the beginning of the career as the main pre-

dictor for current salary. Applying cluster analysis, the author concludes that women are more likely to be employed in the clerical type of job. Also, this study finds out that accumulated previous experience has an important role in the evaluation of our salary.

Earnings of an employer are one of the main predictors for job and life satisfaction as well. According to the study, gained results interpret that beginning salary can be the good predictor of the current salary. This implies that employees should carefully choose their first job. The main focus should be on the satisfaction the job position provides as well how life satisfaction would be on the highest level possible.

REFERENCES

Beattie, C. & Spencer, G. (1971). Career Attainment in Canadian Bureaucracies: Unscrambling the Effects of Age, Seniority, Education, and Ethnolinguistic Factors on Salary. American Journal of Sociology, Vol. 77, No. 3, pp. 472-490.

Fu, W. & Deshpande, S.P. (2014). The impact of caring climate, job satisfaction, and organizational commitment on job performance of employees in a China's insurance company. Journal of Business Ethics, 124(2), pp. 339-349.

Graddy, K. & Pistaferri, L. (2000). Wage differences by gender: Evidence from recently graduated MBAS. Oxford Bulletin of Economics and Statistic, 62, 837–53.

Huang, Y.H., Lee, J., McFadden, A.C., Murphy, L.A., Robertson, M.M., Cheung, J.H. & Zohar, D. (2016). Beyond safety outcomes: An investigation of the impact of safety climate on job satisfaction, employee engagement and turnover using social exchange theory as the theoretical framework. Applied ergonomics, 55, pp. 248-257.

Khamisa, N., Oldenburg, B., Peltzer, K. & Ilic, D. (2015). Work related stress, burnout, job satisfaction and general health of nurses. International journal of environmental research and public health, 12(1), pp. 652-666.

Lee, R & Wilbur, E. (1985). Age, education, job tenure, salary, job characteristics and job satisfaction: A multivariate analysis. Human Relations, 38, pp. 781 – 791.

Roy, D. F. (2017). ""Banana time" job satisfaction and informal interaction." The Anthropology of Organisations. Routledge, pp. 31-41.