EKONOMETRIA ECONOMETRICS 1(39) • 2013

ISSN 1507-3866

Marta Targaszewska
Wrocław University of Economics, Wrocław, Poland
marta.targaszewska@ue.wroc.pl

AN ATTEMPT AT IDENTIFICATION SOURCES OF VARIATION IN MONTHLY NET INCOMES AMONG PERSONS WITH TERTIARY EDUCATION[1]

Abstract: The paper presents the results of the analysis of factors differentiating the level of monthly net incomes in Poland. In the first step of the research, one-way ANOVA was applied in order to verify whether the level of education influences incomes. Subsequently, it was investigated whether such variables as sex, age, the class of residence, region, experience, study major and occupation additionally differentiate the incomes of persons with a higher education. The analysis makes use of the individual data from Social Diagnosis gathered in 2009. All the calculations were performed with the application of IBM SPSS STATISTICS 20 package.

Keywords: analysis of variance, ANOVA, rate of return to education, income, higher education.

1. Introduction

In the year 2000, the European Council accepted a new development plan for the European Union called the Lisbon Strategy. The goal of this initiative was to create the most dynamic, competitive and knowledge-based economy in the world within a period of 10 years [Gorynia, Łaźniewska (eds.) 2009, p. 274]. The key to achieving the appointed goal was to be education, and therefore investment in human capital [*The Returns to Various*… 2005]. In the following years, the trends of higher education development in the European Union were determined by the strategy called "Europe 2020" which, similarly to the Lisbon Strategy, emphasized the significance of educational policy both in economic development and in the fight against unemployment [www. ec.europa.eu/europe2020/index_pl.htm]. By human capital one understands the productive capabilities of a human being, in other words a human's ability to perform work [Dobija (Ed.) 2009, p. 9]. One can therefore claim that the benefits from investment in education are gained not only by the state or the

[1] The research is part of the grant No. 2011/01/B/HS4/02328 *Metody pomiaru stopy zwrotu z inwestycji na edukację w szkołach wyższych*.

economy (as was stated in the Lisbon Strategy) but by the employer and the individual and his family as well. The accurate measure of benefits for individuals are incomes [*The Returns to Various*… 2005].

The most common method applied in studies on financial benefits of education is Mincer's function, which takes into account such factors as [Dziechciarz 2011, p. 60]: number of years of education or level of education and professional experience. The estimated parameters of the function provide an evaluation of the private rate of return to education. Since the level of incomes is additionally influenced by work supply, determined in turn by such factors as for example the major of the studies, the population structure and its propensity to retrain, as well as place of residence, attempts to modify Mincer's model with those factors have been frequently undertaken [*Edukacja dla pracy*… 2007]. The aim of this research is to investigate whether the level of education significantly influences personal monthly net incomes and examining what factors additionally differentiate incomes' level among persons with a higher education. The obtained results could provide the basis for the estimation of the rate of return to education.

2. Dataset

In order to distinguish the factors influencing incomes, the data from Social Diagnosis were used [Rada Monitoringu Społecznego 2009]. Social Diagnosis is survey research on households and its members. The research has the form of a panel and reflects the conditions and quality of life of selected Polish households. The questions included in the questionnaire concern in particular economic conditions such as income, investments, contracted credits, but also non-economic factors mainly in the areas of education, lifestyle and health. The data have been gathered since the year 2000. The measurements take place biennially in March (since 2007) [www.diagnoza.com].

Due to the data availability, the research focuses on the year 2009 [Rada Monitoringu Społecznego 2009]. The analysis encompasses the following variables: level of education, sex, the class of residence, region, study major, occupation, age, tenure of employment (in years), tenure of employment with current employer and personal monthly net income. In Social Diagnosis databases the first six characteristics are qualitative (measured on nominal or ordinal scales) and the following four are quantitative and measured on a ratio scale. Since the method used to distinguish the factors influencing incomes was ANOVA, the variables: age, tenure of employment and tenure of employment with current employer, being grouping variables, were transformed into qualitative. Moreover, a few quantitative variables were regrouped into new, more numerous classes. In the analysis of influence of individual variables on incomes, only those respondents were chosen which are of working age, who currently work (according to the variable tenure of employment with current employer) and who declare salaries at least as high as the minimum wage, which in

2009 was equal to net 954 [PLN] (gross 1276 PLN) [www.zus.pl/default.asp?p=1&id=1019]. The analysis also excluded those respondents which in 2009 declared incomes significantly higher than the rest of the population (outliers and extreme observations[2]). With such assumptions, the monthly net income depending on level of education was analysed for 9756 respondents, and for 2022 respondents with higher education the influence of additional factors on income was investigated.

3. One-way analysis of variance

One-way analysis of variance (ANOVA) is a statistical method developed by R.A. Fisher in 1925 [Stanisz 2007a, p. 272] which allows to compare a few populations in terms of characteristics measured on a metric scale. These comparisons indicate whether the mean values of the analysed dependent variable are the same for the respective populations [Walesiak, Gatnar (eds.) 2012, p. 104]. The goal of one-way ANOVA is therefore to verify the hypothesis that the analysed variable is influenced by chosen factors called grouping variables (independent/explanatory). A tested null hypothesis (stating the equality of means in groups) is given by the following formula [Stanisz 2007a, p. 274]:

$$H_0: \mu_1 = \mu_2 = \ldots = \mu_k, \tag{1}$$

where μ_i ($i = 1, 2, \ldots, k$) is the expected value, and k is the number of levels for the given grouping variable. The significance of differences between means in groups is evaluated based on the ratio of between-group variation and within-group variation (test statistics F) [Walesiak, Gatnar (eds.) 2012, p. 106]. Accordingly, the test statistics are expressed by the formula [Walesiak, Gatnar (eds.) 2012, p. 107]:

$$F = \frac{SSM}{SSW} \cdot \frac{n-k}{k-1}, \tag{2}$$

where: SSM – between-group variation,
SSW – within-group variation,
n – number of observations,
k – number of levels for given grouping variable.

One-way analysis of variance model represents each observation x_{ij} (the result of the effect of i-th level of factor on j-th observation) as a linear combination of a certain constant, unknown and common to all population value (μ), factor characteristic for i-th population (α) and error term, as given by the following formula [Stanisz 2007a, p. 283]:

$$x_{ij} = \mu + \alpha_i + \varepsilon_{ij}. \tag{3}$$

[2] All the observations which exceed 1,5 times interquartile range (in each of analyzed populations) were marked as extreme or outliers. And they were in all about 132 cases (respondents which declared incomes greater than or equal to 5400 [PLN]).

The application of analysis of variance is limited by the validity of the following assumptions [Ntoumanis 2001, pp. 73, 74]: the samples are random and independent of each other in each population, the dependent variable follows normal distribution in each group of independent variable, the variance is homogeneous in all groups of independent variable.

In cases of a big sample size for each population (over 100 observation) there is no need to examine the normality of distribution – according to the Central Limit Theorem the probability distribution of any sample mean has an approximate normal distribution [Wywiał 2004, p. 44]. What is more, if the sample size is large, ANOVA is quite robust to the violation of normality assumption [Warner 2008, p. 510]. If the equal variance assumption has been violated, an adjusted F statistic can be used. Such a type for adjustment is provided by the strong test for means equality, namely the Welch test [Field 2005, p. 346]. The Welch statistic is expressed by the formula [Proust 2009, p. 141]:

$$W = \frac{\frac{\sum_{i=1}^{k} w_i (\bar{x}_i - \hat{x})^2}{k-1}}{1+\left[\frac{2(k-2)}{k^2-1}\right]\sum_{i=1}^{k}\frac{\left(1-\frac{w_i}{W}\right)^2}{n_i-1}}, \quad (4)$$

where: $w_i = \frac{n_i}{s_i^2}$, $\hat{x} = \sum_{i=1}^{k} \frac{w_i x_i}{W}$, $W = \sum_{i=1}^{k} w_i$.

If the measurement variable does not meet the assumptions of an ANOVA (normality and homogeneity of variance) the Kruskal-Wallis test is commonly used [Stanisz 2007 b, p. 386].

4. The ANOVA results

The first step of the research verified whether the level of education influences the level of monthly net incomes. Because of the size of the sample the examination of distribution normality was not necessary. The number of observations, percentage shares and means for individual groups of independent variables are given in Table 1.

Table 1. Number of observations and basic statistics for variable education level

The level of grouping variable	Number of observations	(%)	Mean (PLN)
Higher education	2 221	23	2 331
Post-secondary education	409	4	1 695
Secondary vocational education	2 504	26	1 761
Secondary general education	789	8	1 617
Basic vocational education	3 101	32	1 637
Other[a)]	732	7	1 423
Total	9 756	100	1 811

[a)] The level *others* comprises lower-secondary education, primary education, unfinished primary education and without education.

Source: own elaboration based on dataset [Rada Monitoringu Społecznego 2009].

The Levene test was used in order to verify whether the variances for each level of education are homogeneous (Levene's test tests the null hypothesis that the variances in different groups are equal) [Stanisz 2007a, p. 342]. The significance level (p-value) for the executed test was equal to 0.000, which allows us to reject the null hypothesis in favour of the alternative hypothesis, stating the lack of monthly net income variance homogeneity in each group. In the given situation, to evaluate the influence of education level on monthly net income, the Welch test was used. For the executed test the border significance level was 0.000, which indicates the rejection of the null hypothesis of means equality in groups. Means in groups according to education level are presented in Figure 1.

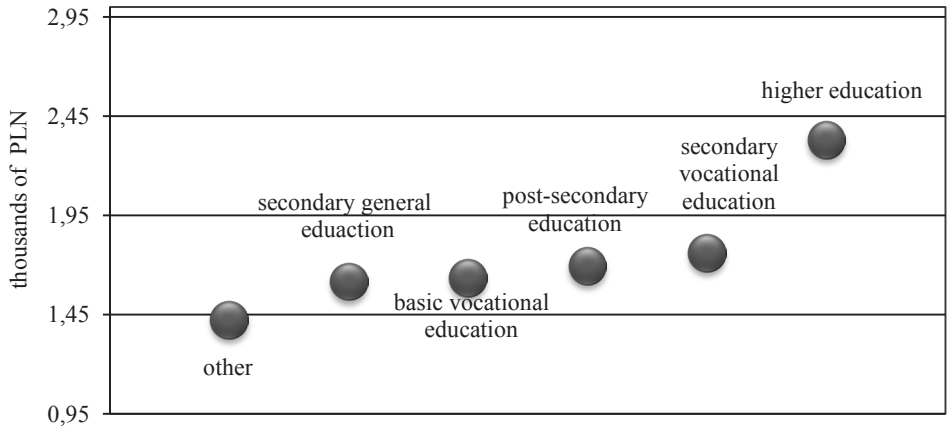

Figure 1. Means in groups according to education level

Source: own elaboration.

A *post-hoc* Games-Howell test (applied in cases of variance heterogeneity [Morgan et al. 2004, p. 152] and non-equinumerous groups [Field 2005, p. 341]) allowed us to single out three levels of monthly net income according to the level of education. The first group is composed of respondents with higher education and highest incomes. The average monthly income in this group amounts to 2331.16 PLN. A separate group is created by the respondents with post-secondary education, secondary vocational education, basic vocational education and secondary general education. In this group the highest average income is achieved by the respondents with secondary vocational education. The group with the lowest average income of 1423.31 PLN is formed by the persons with at the most lower-secondary education level. The research conclusions show that the level of education significantly differentiates the achieved incomes.

The next step of the research investigated which factors additionally differentiate the incomes of the respondents with higher education.

Table 2. Grouping variables' characteristics (age, sex, class of residence, region)

Name of the variable	Classes	Number of observations	(%)	Mean [PLN]
Age[a)]	• working mobile age	1 408	70	2 286
	• working immobile age	614	30	2 589
Sex	• male	785	39	2 720
	• female	1 237	61	2 161
The class of residence	• big cities – number of inhabitants 100 thousands and more	924	46	2 567
	• small and medium cities - number of inhabitants less than 100 thousands	650	32	2 251
	• villages	448	22	2 171
Region[b)]	• central (without Warsaw)	226	11	2 298
	• south (without Silesia)	150	8	2 296
	• east	392	19	2 128
	• north-west	319	16	2 374
	• south-west	198	10	2 436
	• north	329	16	2 476
	• Warsaw sub region	210	10	2 778
	• Silesia	198	10	2 386

[a)] According to the GUS classification: working mobile age (18–44 years); working immobile age (for females 45–59 and for males 45–64 years); post-working age (females minimum 60 and males minimum 65 years); [b)] Warsaw sub region and Silesia were analyzed separately due to higher income levels than observed in other regions in Poland (www.wynagrodzenia.pl/dane_gus.php, [14.11.2012]).

Source: own elaboration based on dataset [Rada Monitoringu Społecznego 2009].

Prior to the application of ANOVA, each analysed factor was tested for validity of the assumptions. Since the size of the sample for each independent variable and on every level was at least 100, the examination of distribution normality was not necessary. The Levene test was applied to verify the homogeneity of variance. The significance levels for variance homogeneity test and one-way ANOVA or strong test for means equity (in cases of the lack of homogeneity of variance) are presented in Table 4.

All the analysed independent variables significantly (at the level 0.05) differentiate personal monthly net income. The results indicate that males achieve a higher income than females. Their wages are on average 550 PLN higher. Higher wages are achieved as well by persons of working immobile age, which may be connected with their more extensive experience. The average monthly net income for the respondents of working mobile and immobile age in 2009 amounted to 2286 PLN and 2589 PLN respectively. In the case of the variables, region and class of residence, a high monthly net income is specific for big cities and the Warsaw sub region. Moreover, the *post-hoc* Games-Howell test showed that there is no significant difference

Table 3. Grouping variables' characteristics (tenure of employment, tenure of employment with current employer, study major, occupation)

Name of the variable	Classes	Number of observations	(%)	Mean [PLN]
Tenure of employment[a)]	• less than 5 years	366	18	1 964
	• at least 5 but less than 20 years	959	48	2 389
	• at least 20 years	690	34	2 579
Tenure of employment with current employer	• 5 years or less	977	48	2 240
	• more than 5 years	1 045	52	2 507
Study major[b)]	A. education	375	19	2 137
	B. arts, humanities	203	10	2 272
	C. social sciences, journalism and information sciences, economy and administration, law	717	36	2 367
	D. biological sciences, physics, mathematics and statistics, computer sciences	193	10	2 565
	E. technical sciences, production and processing, architecture and engineering	271	13	2 631
	F. agriculture, forestry, fishing, veterinary medicine, public health, health care, social welfare, services for population and transportation services, protection of environment and sanitary municipal services, protection and safety	257	13	2 434
Occupation[c)]	I. parliamentarians, high officials and managers	223	11	2 984
	II. specialists	1 080	55	2 399
	III. technicians and other mid-level staff	336	17	2 308
	IV. office workers	156	8	1 958
	V. personal services staff and salesmen, farmers, gardeners, foresters and fishermen, industry workers and craftsmen, operators and mechanics for machines and devices, simple work staff, armed forces	166	8	2 054

[a)] Ranges indicated by: *Ustawa o promocji zatrudnienia i instytucjach rynku pracy z dnia 20 kwietnia 2004 r*. [DzU 2004 nr 99, poz. 1001]; [b)] Classification indicated by: *Rozporządzenie Rady Ministrów w sprawie Polskiej Klasyfikacji Edukacji z dnia 6 maja 2003 r*. [DzU 2003 nr 98, poz. 895]; [c)] Classification indicated by: *Rozporządzenie Ministra Pracy i Polityki Społecznej w sprawie klasyfikacji zawodów i specjalności na potrzeby rynku pracy oraz zakresu jej stosowania z dnia 27 kwietnia 2010 r.* [DzU 2010 nr 82, poz. 537].

Source: own elaboration based on dataset [Rada Monitoringu Społecznego 2009].

Table 4. Significance levels for Levene test, one-way ANOVA and strong tests for means equality

Name of the variable	Significance – Levene test	Significance – ANOVA/ strong tests for means equality
Age	0.369	0.000
Sex	0.000	0.000
The class of residence	0.000	0.000
Region	0.004	0.000
Tenure of employment	0.002	0.000
Tenure of employment with current employer	0.124	0.000
Major	0.000	0.000
Occupation	0.004	0.000

Source: own elaboration.

between incomes in small and medium cities (number of inhabitants less than 100 thousand) and villages. Group means according to the regions are shown in Figure 2.

The lowest income is specific for respondents living in the eastern region, including the following voivodeships: Lublin, Podkarpackie, Podlaskie and Świętokrzyskie. The average income in this region amounts to 2128 PLN, which is below the country's average income level. Incomes in the eastern region are similar to those in central (voivodeships Masovian without Warsaw and Łódź) and southern regions (Lesser Poland voivodeship). Border significance level in the Games-Howell test for the mentioned regions and eastern region were 0.422 and 0.587 respectively. In contrast, a small differentiation of monthly net income is present among

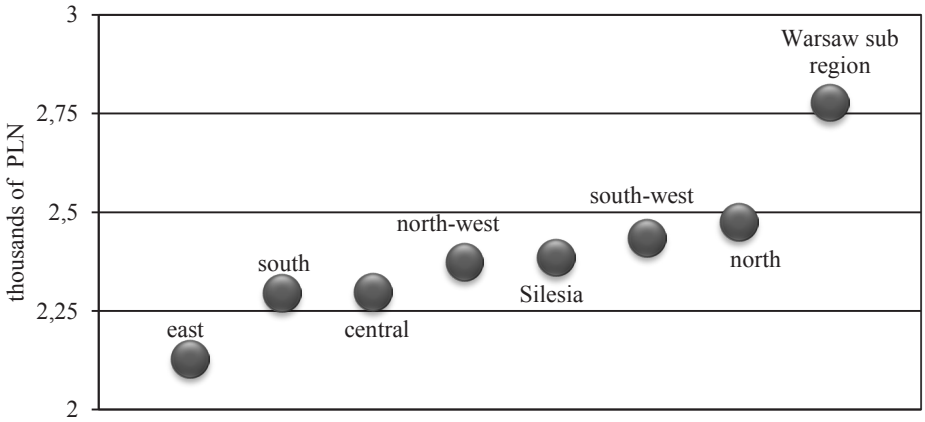

Figure 2. Group means according to the regions

Source: own elaboration.

respondents from the north-west, south-west, northern regions and Silesia. In the case of variables denoting experience, namely tenure of employment and tenure of employment with current employer, the conducted analysis showed that both factors with a 95% probability significantly influence achieved income. In other words, the more experienced the respondent with a higher education, the higher the income. A *post-hoc* test for tenure of employment showed that all three levels of variable are significantly differentiated (*p*-value was lower than 0.05 in case of each compared pair). The level of monthly net income is also influenced by the study major. Average incomes according to the study major are presented in Figure 3.

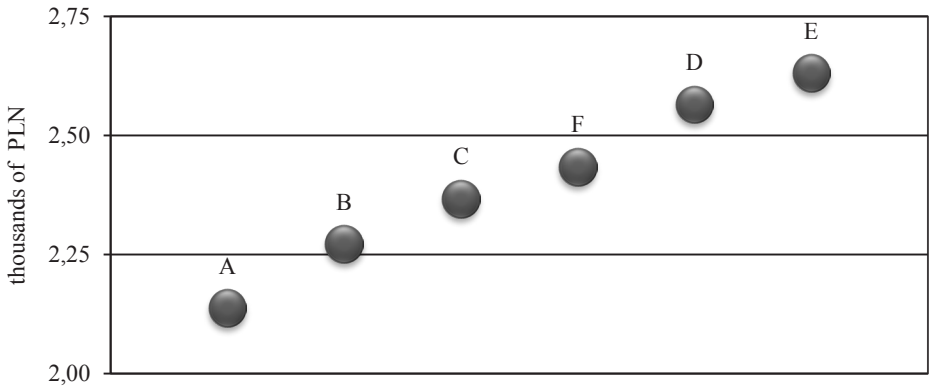

Figure 3. Group means according to major[3]

Source: own elaboration.

The highest incomes (2631 PLN) are achieved by respondents who majored in technical sciences including production, processing, architecture or engineering. This group is significantly different from such majors as education, arts, humanities and social studies. The lowest income is specific for the respondents who graduated in arts, humanities and education. Their monthly net income is situated below the country's average income level. The border significance level ($p = 0.957$) for the *post-hoc* Games-Howell test indicates that the hypothesis of means equality for majors from group C (social sciences, journalism and information sciences, economy, administration and law) as well as majors from group F (agriculture, forestry, fishing, veterinary medicine, public health, health care, social welfare, services for population, transportation services, protection of environment, sanitary municipal services, protection and safety), where the difference in income is only 67 PLN, should not be rejected. Since the major of the study is closely related to the occupation, it was investigated whether profession influences the difference in achieved incomes.

[3] For a description of the variable names refer to Table 3.

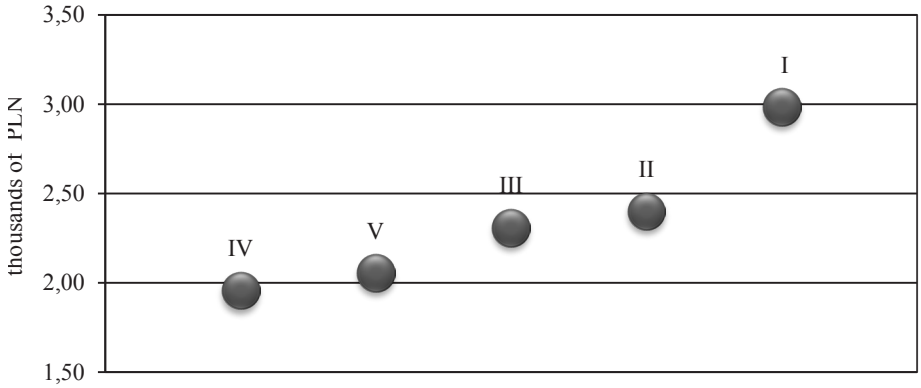

Figure 4. Group means according to occupation[4]

Source: own elaboration.

Based on Figure 4 one can conclude that the highest incomes are specific to the first group, namely parliamentarians, high officials and management. This group earns considerably more than the country's average income level. On average their wages are 800 PLN higher than for other professions. For technicians, mid-level staff and specialists, the null hypothesis stating the equity of means was not rejected (with a Games-Howell test significance level of 0.607). A similar situation is present for the groups represented by office workers and personal services staff, salesmen etc.

5. Conclusions

Based on the conducted research, one can conclude that the level of education significantly influenced monthly net incomes in 2009. The highest incomes were observed for persons with a higher education, and the lowest for persons with at most a lower-secondary education level. Moreover, the results allowed us to conclude that among the persons with higher education:

- all analysed independent variables significantly differentiate monthly net incomes,
- females earn less than males,
- higher incomes are specific for the persons of immobile working age,
- there is a significant difference between the wages of persons living in big cities and the persons living in small and medium cities and villages,
- monthly net incomes of Warsaw citizens are considerably higher than incomes of the other Polish citizens, and are higher by 300 PLN in comparison to other regions,
- the graduates of technical and theoretical science studies achieve higher incomes than graduates of educational studies or humanities,

[4] For a description of the variable names refer to Table 3.

– income increases as experience increases (measured by tenure of employment and tenure of employment with current employer),
– the highest income is achieved by the respondents working as high officials and managers.

Literature

Dobija M. (ed.), *Kapitał ludzki w perspektywie ekonomicznej*, Uniwersytet Ekonomiczny w Krakowie, Kraków 2009.

Dziechciarz J., On rate of return measurement in education, *Econometrics* 2011, no. 194, pp. 49–66.

Edukacja dla pracy. Raport o rozwoju społecznym Polska 2007, United Nations Development Programme, www.unic.un.org.pl/nhdr/nhdr2007/Edukacja 20dla 20pracy.pdf (accessed 27.11.2012).

Field A., *Discovering Statistics Using SPSS*, SAGE, London 2005.

Gorynia M, Łaźniewska E. (eds.), *Kompendium wiedzy o konkurencyjności*, PWN, Warszawa 2009.

Morgan G.A., Leech N.L., Gloeckner G.W., Barrett K.C., *SPSS for Introductory Statistics: Use and Interpretation*; Lawrence Erlbaum Associates, Mahwah 2004.

Ntoumanis N., *A Step-by-Step Guide to SPSS for Sport and Exercise Studies*, Routledge, London 2001.

Proust M., *Statistics and Graphics Guide*, SAS, Cary 2009.

Rada Monitoringu Społecznego, *Diagnoza społeczna: zintegrowana baza danych*, 2009, www.diagnoza.com (accessed 04.08.2012).

Rozporządzenie Ministra Pracy i Polityki Społecznej w sprawie klasyfikacji zawodów i specjalności na potrzeby rynku pracy oraz zakresu jej stosowania z dnia 27 kwietnia 2010 r. [DzU 2010 nr 82, poz. 537].

Rozporządzenie Rady Ministrów w sprawie Polskiej Klasyfikacji Edukacji z dnia 6 maja 2003 r. [DzU 2003 nr 98, poz. 895].

Stanisz A., *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 2. Modele liniowe i nieliniowe*, StatSoft, Kraków 2007a.

Stanisz A., *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 1. Statystyki podstawowe*, StatSoft, Kraków 2007b.

The Returns to Various Types of Investment in Education and Training (2005), European Commission. www.ec.europa.eu/education/policies/2010/studies/invest05_en.pdf (accessed 30.11.2012).

Ustawa o promocji zatrudnienia i instytucjach rynku pracy z dnia 20 kwietnia 2004 r. [DzU 2004 nr 99, poz. 1001].

Walesiak M., Gatnar E. (Eds.), *Statystyczna analiza danych z wykorzystaniem programu R*, PWN, Warszawa 2012.

Warner R.M., *Applied Statistics*, Sage Publications, Thousand Oaks 2008.

Wywiał J., *Wprowadzenie do wnioskowania statystycznego*, Akademia Ekonomiczna w Katowicach, Katowice 2004.

Internet sources

www.ec.europa.eu/europe2020/index_pl.htm (accessed 30.11.2012).

www.diagnoza.com (accessed 14.09.2012).

www.zus.pl/default.asp?p=1&id=1019 (accessed 14.09.2012).

www.wynagrodzenia.pl/dane_gus.php (accessed 14.11.2012)

PRÓBA IDENTYFIKACJI ŹRÓDEŁ ZRÓŻNICOWANIA MIESIĘCZNEGO DOCHODU NETTO WŚRÓD OSÓB Z WYKSZTAŁCENIEM WYŻSZYM

Streszczenie: W niniejszym artykule zaprezentowano wyniki badań nad czynnikami różnicującymi wysokość miesięcznego dochodu netto w Polsce. Za pomocą jednoczynnikowej analizy wariancji ANOVA w pierwszym kroku sprawdzono, czy poziom wykształcenia ma wpływ na osiągany dochód. Następnie zbadano, czy takie zmienne, jak: płeć, wiek, klasa miejscowości, region Polski, doświadczenie, ukończony kierunek studiów oraz wykonywany zawód, dodatkowo różnicują dochód osób z wykształceniem wyższym. Do analizy wykorzystano dane indywidualne Diagnozy Społecznej zabrane w 2009 r. Wszystkie obliczenia wykonano za pomocą pakietu IBM SPSS STATISTICS 20.

Słowa kluczowe: analiza wariancji, ANOVA, stopa zwrotu z edukacji, dochód, wykształcenie wyższe.