Various Approaches to Measuring Effectiveness of Tertiary Education

Józef Dziechciarz, Marta Dziechciarz-Duda, Anna Król and Marta Targaszewska

Abstract This paper aims at assessing selected approaches to measuring the effectiveness of investment in tertiary education and their applicability. It summarizes various results obtained in the research project *Methods of Measuring the Return on Investment in Higher Education*. The applied methods, include classical methods (ANOVA, Mincerian earnings function, correspondence analysis, hierarchical agglomerative clustering) as well as new ideas (application of the Wilcoxon Matched-Pairs Signed-Rank Test to determine the significance of differences in incomes before and after reaching the tertiary education). The research is based on data coming both from Polish (Social Diagnosis, Study of Human Capital) as well as German databases (Social Economic Panel, SOEP). The obtained results support the hypothesis that tertiary education influences the level of incomes. Moreover, the estimated pseudo rates of return to education

Józef Dziechciarz
Wrocław University of Economics, Poland
✉ jozef.dziechciarz@ue.wroc.pl

Marta Dziechciarz-Duda
Wrocław University of Economics, Poland
✉ marta.dziechciarz@ue.wroc.pl

Anna Król
Wrocław University of Economics, Poland
✉ anna.krol@ue.wroc.pl

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

ARCHIVES OF DATA SCIENCE, SERIES A
KIT SCIENTIFIC PUBLISHING
Vol. 1, No. 1, S. 103–127, 2016

DOI 10.5445/KSP/1000058747/07
ISSN 2363-9881

provide the basis for the evaluation of the effectiveness of private investment in education.

1 Introduction

Since the year 1998, when the Sorbonne Declaration was signed and the European Higher Education Area was established, the development and modernization of higher education have been priorities of the policies of the European Union (EU) policies. Both, the *Bologna Declaration* and the *Lisbon Strategy*, have emphasized the following aims: improving the quality of education, building the knowledge-based society and economy, adapting the education system to the needs of the labour market, lifelong learning, and supporting the acquisition of skills to compete in a global environment. According to the strategy presented by the European Commission in 2006, the project of the Modernization Agenda for Universities (entitled *Delivering on the Modernisation Agenda for Universities: Education, Research and Innovation*) should be based on three reforms: curricula, governance and funding (cf. European Commission, 2006).

The latest EU strategy – "Europe 2020" – is yet another step of reforming the higher education system in Europe. Its main priority is to support the creation of a knowledge-based and balanced economy which favors social inclusion and cohesion. Tertiary education is one of the essential factors in achieving the main goals of this strategy. In order to define and realize all the educational aims of the strategy, the European Commission issued the Higher Education Modernisation Agenda which recommends – among other measures – increasing the number of universities' graduates, encouraging people from various social groups to undertake studies, increasing the quality of tertiary education, adjusting the curricula to labour market needs, directing higher education on financial crisis issues, as well as introducing outcome oriented funding of the universities (output-budgeting).

Moreover, the contemporary research and education market with its increasing number of students, globalization, rapid technological development, growing research costs, emergence of specialized university-independent B+R centres, and the increasing significance of commercialization and entrepreneurship poses many challenges for traditional universities and enforces their transformation (cf. Etzkowitz and Peters, 1991; Wissema, 2009; Jongbloed, 2010). In order to ensure the better quality, effectiveness and accessibility of higher

education among all EU countries the shift from the traditional Humboldtian University towards the modern entrepreneurial university is essential. Conventionally functioning areas of the university – education and research – should be supplemented by other fields such as research commercialization, application for external grants, and projects as well as co-operation with industry.

A significant part of the postulated reforms in the functioning of the universities require changes in the area of funding, in particular encouraging a shift from a centralized input oriented funding mechanism towards a decentralized outcome oriented financing. All this causes the necessity of measuring the effectiveness of various aspects of higher universities' activities, including education (cf. Dziechciarz, 2011).

This paper aims at assessing selected approaches to measuring effectiveness of investment in tertiary education and their applicability, making use of the data from Polish (Social Diagnosis, Study of Human Capital), as well as German databases (SOEP). It summarizes various results obtained in the framework of the research project *Methods of Measuring the Return on Investment in Higher Education*.

2 Rate of return to education concepts

In this paper *effectiveness* refers to a relationship between higher education, resources used in education and outcomes – labour productivity and graduates' employability (cf. Aubyn et al, 2008, p. 55). One of the concepts used in measuring effectiveness in the education system is the rate of return on investment to education. The most widely and commonly used approach is the concept of private returns, measured from the point of view of individuals (students), where benefits are increased earnings and costs are foregone earnings, education fees, cost of attendance or other incidental expenses during the period of studies (cf. Psacharopoulos, 1995). The returns to education may also be measured from the social perspective. The costs are in this case the state's and the society's large spending on education and the benefits are based on productivity (cf. Psacharopoulos, 1995). Table 1 presents various types of benefits of education from both private and social perspectives.

The focus of this study is placed on measuring the effectiveness of tertiary education, defined as post-secondary education obtained at both universities and colleges.

Table 1 Classification of the benefits of education

	Private Benefits	Social Benefits
Market	employability	higher productivity
	higher earnings and savings	higher net tax revenue
	less unemployment	less reliance on government
	labor market flexibility	financial support
	greater mobility	technological development
Non-market	increased happiness	reduced crime
	better personal and family health	less spread of infectious diseases
	better child cognitive development	lower fertility
	greater longevity	better social cohesion
	greater satisfaction from consumption decisions	voter participation

See e.g. Psacharopoulos (2009), McMahon (1997).

3 Datasets

The described research is based on three datasets: German database Socio-Economic Panel Study (SOEP) (Wagner et al, 2007), and two Polish bases: Social Diagnosis (Rada Monitoringu Społecznego, 2003-2011) and the Study of Human Capital (BKL) (Bilans Kapitału Ludzkiego, 2012).

The SOEP is an annual wide-ranging representative longitudinal study on private households which started in 1984. The data provides information on households and its members and some of the many aspects include household composition, occupational biographies, employment, earnings, health and satisfaction indicators.

The Polish Social Diagnosis is a panel study investigating households and their members aged 16 and above. The project takes into account all the significant aspects of life, both the economic ones (i.e. income, material wealth, savings and financing), and the not strictly economic ones (i.e. education, medical care, problem-solving, stress, psychological well-being, lifestyle, pathologies, engagement in the arts and cultural events). The first sample was taken in the year 2000. The following study took place three years later, and since then has been repeated every two years. The database is open and may be accessed through the internet site of the panel[1].

The BKL is a labour market monitoring project carried out by the Polish Agency for Enterprise Development (Polska Agencja Rozwoju Przedsiębiorczości, PARP) in collaboration with the Jagiellonian University Krakow. In the

[1] http://www.diagnoza.com/

years 2010-2014 the project traced how the structure of competences changed in the labour market and sought answers to the key questions related to human capital at both the national and regional level. The project provides access to its results and gathered data without any limitations and fees. [2]

4 Various reseach approaches to measuring effectiveness of tertiary education

4.1 Application of the Mincer model in the analysis of the influence of tertiary education on the level of incomes

The first approach to measure effectiveness of tertiary education is a two step procedure. In the first step we examine the significance of the influence of education on monthly net incomes. Additionally we investigate whether factors such as sex, the class of residence, region, study major, occupation, age, tenure of employment, tenure of employment with current employer etc. significantly differentiate the income level among persons with higher education of those with lower education.

The number of observations for the analysis of the influence of education level on monthly net incomes is 9756, and 2022 for the investigation of the influence of additional factors on incomes. Only those respondents which are of working age, and who currently work (according to the variable tenure of employment with current employer), and who declared salaries at least as high as the minimum wage in the year 2009 were chosen. The research also excluded those respondents which declared extreme incomes. The list of the variables with mean standard derivation and number of observations used in the research is given in Tables 2 and 3.

In the described research one-way analysis of variance (ANOVA) is applied, except in cases of heterogeneous variances in groups of independent variables, in which the Welch test was used (cf. Proust, 2009, p. 141). ANOVA is used to examine the equality of group means for a quantitative outcome. The goal of one-way ANOVA is to verify the hypothesis that the analysed variable is influenced by independent (grouping) variables by rejecting the null hypothesis that all of the group means are equal (cf. Walesiak and Gatnar, 2012, p. 104). The application of one-way ANOVA is limited by the following assumptions:

[2] http://en.bkl.parp.gov.pl/

Table 2 Independent variables characteristics of the data set Social Diagnosis (2003–2011)

Variable name	Groups	Number of observations	Mean[1] [PLN]	Standard deviation
education level	Higher education	2221	2331	1013
	Post-secondary education	409	1695	548
	Secondary vocational education	2504	1761	636
	Secondary general education	789	1617	570
	Basic vocational education	3101	1637	602
	Lower secondary, primary or unfinished primary education and without education	932	1423	437
age	working mobile age (18–44 years)	1408	2286	987
	working immobile age (females 45–59, males 45–64 years)	614	2589	1031
sex	male	785	2720	1089
	female	1237	2161	891
the class of residence	big cities (100000 and more inhabitants)	924	2567	1076
	small and medium cities (less than 100000)	650	2251	2251
	villages	448	2171	2171
region [2]	central (without Warsaw)	226	2298	1003
	south (without Silesia)	150	2296	855
	east	392	2128	909
	north-west	319	2374	993
	south-west	198	2436	1075
	north	329	2476	1007
	Warsaw sub region	210	2778	1121
	Silesia	198	2386	942
tenure of employment [3]	less than 5 years	366	1964	943
	at least 5 but less than 20 years	959	2389	987
	at least 20 years	690	2579	1026
tenure of employment with current employer	5 years or less	977	2240	1021
	more than 5 years	1045	2507	982

[1] Monthly net income
[2] Warsaw sub region and Silesia were analysed separately due to higher income levels than observed in other regions in Poland (`www.wynagrodzenia.pl/dane_gus.php`, [14.11.2012])
[3] Ranges indicated by: Ustawa o promocji zatrudnienia i instytucjach rynku pracy z dnia 20 kwietnia 2004 r. [Dz. U. 2004 nr 99, poz. 1001].

the dependent variable should be normally distributed and the variance should be homogeneous in all group of independent variables (cf. Ntoumanis, 2001, pp. 73, 74).

Table 3 Independent variables characteristics of the data set Social Diagnosis (2003–2011) (cont.)

Variable name	Groups	Number of observations	Mean [1] [PLN]	Standard deviation
study major [2]	education	375	2137	826
	arts, humanities	203	2272	915
	social sciences, journalism, information sciences, economy and administration, law	717	2367	1024
	biological sciences, physics, mathematics, statistics, computer sciences	193	2565	1105
	technical sciences, production and processing, architecture and engineering	271	2631	1035
	agriculture, forestry, fishing, veterinary medicine, public health, health care, social welfare, services for population, transportation services, protection of environment and sanitary, municipal services, protection and safety	257	2434	1090
occupation [3]	parliamentarians, high officials and managers	223	2984	1078
	specialists	1080	2399	947
	technicians and other mid-level staff	336	2308	1028
	office workers	156	1958	8778
	personal services staff and salesmen, farmers, gardeners, foresters, fishermen, industry workers, craftsmen, operators and mechanics for machines, simple work staff, armed forces	166	2054	986

[1] Monthly net income

[2] Classification indicated by: Rozporządzenie Rady Ministrów w sprawie Polskiej Klasyfikacji Edukacji z dnia 6 maja 2003 r. [Dz. U. 2003 nr 98, poz. 895].

[3] 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. [Dz. U. 2010 nr 82, poz. 537].

The research was conducted for the data from the Social Diagnosis data set in the year 2009, and all hypotheses (influence of an independent variable on income) were verified at the 95% confidence level. The test statistics and significance levels for one-way ANOVA or Welch test are presented in Table 4.

The results of the analysis show that the level of education significantly influences the monthly net income. The highest incomes were characteristic for persons with tertiary education degree, and lowest for persons with at most a lower-secondary education level. Moreover, all independent variables

Table 4 Test of influence of independent variable on income for 2009 of the data set Social Diagnosis

Variable name	Statistic	Significance – ANOVA/ strong tests for means equality
education level	245.500	0.000
age	39.193	0.000
sex	145.514	0.000
the class of residence	31.126	0.000
region	8.774	0.000
tenure of employment	46.358	0.000
tenure of employment with current employer	35.695	0.000
major	11.027	0.000
occupation	31.317	0.000

significantly (at the level 0.05 for the post-hoc Games-Howell test (cf. Morgan et al (2004, p. 152), Field (2005, p. 341)) differentiate personal monthly net income of persons with higher education degree, where the average monthly income is 2331.16 PLN. The results of research also shows that females achieve lower (on average of 550 PLN) wages than males. Lower incomes are specific for persons of working mobile age. This could be explained by the lower experience of those persons.

There is also a significant difference between the wages of persons living in big cities (population ≥ 100000) and the persons living in small and medium cities (population < 100000) or villages, where the wages of persons from the first two categories are the highest. In case of the variable "region" high income is characteristic for the respondents from the Warsaw sub-region, and the lowest income is observed for the Eastern provinces (Lublin, Podkarpackie, Podlaskie, Świętokrzyskie).

Both, "tenure of employment" and "tenure of employment with current employer", significantly influence the achieved income - persons with higher experience (more than 5 years) earn more money. The last two variables – "study major" and "occupation" - are closely related. Our research shows that the graduates of educational studies, humanities or art studies obtain lower incomes than graduates of technical and theoretical science studies. Moreover, the average incomes in the group of social sciences, journalism and information sciences, economy, administration and law studies are similar to the group which contains studies such as 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. The level of monthly net income is also influenced by occupation,

where the highest incomes (800 PLN higher than for other professions) are specific to parliamentarians, high officials and managers. Similar earnings are characteristic for technicians, mid-level staff and specialists. For more details on the result of this analysis see Targaszewska (2013).

These results were used to support the process of cmodel selection in the second step of the research - the estimation of the private pseudo rate of return to education with Mincer's earning function. Additionally, the analysis of the dynamics of the influence of the education level on wages was performed. The empirical research for Germany was based on the SOEP database (years 1995, 2000, 2005, 2010) and for Poland on the Social Diagnosis data set (years 2003, 2005, 2007, 2009, 2011). The description of variables for the SOEP data set and its summary statistics are presented in Table 5, whereas the presentation of the variables from the Social Diagnosis data set is given in Table 6. For the sake of conciseness, the summary statistics for the SOEP data set are presented only for the year 2010 (the most recent of the analyzed years) and for the Social Diagnosis data set only for the year 2009 (as in Table 4). Similarly, the various specifications of the estimated models are presented in detail in Tables 7 and 8 only for the selected years (the SOEP data set in year 2010 and the Social Diagnosis data set in year 2009 accordingly), and Tables 9 and 10 present only the most important results for each year of the study.

To analyze the dynamics of influence of education level on the wages the commonly applied Mincerian earnings function (cf. Mincer, 1958, 1974) was used in each year separately:

$$lnEAR_i = X_i^T \beta + \varepsilon_i, \tag{1}$$

where *EAR* - earnings, X - vector of variables influencing wages, β - vector of unknown parameters, ε - error term. The elements of X describe education (represented by number of years of education or dummies for level of education, the latter providing the estimation of the pseudo rate of return to education) and professional experience, as well as auxiliary characteristics such as gender, region, place of work, position etc. As the dependent variable real hourly gross earnings for SOEP data set and real monthly net earnings for Social Diagnosis data set were used. Adjustment for inflation led to creation of new variables: *RHGEAR* and *RMNEAR* accordingly. Afterwards the inflation-adjusted earnings were transformed to natural logarithms. The log-linear functional form proved to be correct in many previous studies (cf. Heckman et al, 2003). In addition, a Box-Cox transformation which allows to choose between linear ($\alpha = 1$) and log-linear ($\alpha = 0$) specification was used (cf. Box and Cox, 1964):

Table 5 Description of variables for the SOEP data set and summary statistics in year 2010

Variable name	Description	Mean	Standard deviation	Distribution of categories (%)
HGEAR	Hourly gross earnings in [EUR]	16.80	15.01	
YOET	Years of education and training	12.82	2.75	
AGE	Age	43.70	12.52	
PWE	Potential work experience (*PWE=AGE-YOET*-6)	25.60	12.18	
HEDU	Higher education			
	1 if obtained higher education diploma			26.60
	0 otherwise			73.40
MEDU	Secondary education			
	1 if obtained secondary education diploma			98.08
	0 otherwise			1.92
SEN	Tenure of employment with current employer	11.53	10.57	
FEM	Gender			
	1 if female			48.90
	0 otherwise			51.10
TYPE	Work position type			
	APP (trainee)			12.74
	SPEC (specialist)			12.17
	PROF (freelancer/professional)			44.69
	MAN (manager)			6.74
	OTHER (other)			23.66
SIZE	Size of current employer			
	SMALL (less than 20 employees)			31.66
	MEDIUM (20 - 2000 employees)			47.29
	LARGE (more than 2000 employees)			21.05

$$B(EAR_i, \alpha) = \begin{cases} \frac{EAR_i^{\alpha}-1}{\alpha} \text{ for } \alpha \neq 0 \\ \ln EAR_i \text{ for } \alpha = 0 \end{cases}. \tag{2}$$

For all tested specifications the parameter α was close to 0 indicating the better fit of the models resulting from the log-linear transformation. Fig. 1 presents the results of searching the parameter α which maximizes the logarithm of the likelihood function for the model specification SOEP4 in Table 7. The horizontal line in Fig. 1 marks the 95% confidence interval for parameter α. Its lower boundary (0.1051), center (0.1213) and upper boundary (0.1353) are shown by vertical lines.

Table 6 Description of variables for the Social Diagnosis data set and summary statistics in year 2009

Variable name	Description	Mean	Standard deviation	Distribution of categories (%)
MNEAR	Monthly net earnings in [PLN]	1350.00	707.72	
HE	Higher education			
	1 for tertiary education			10.83
	0 otherwise			89.17
ME	Secondary education			
	1 for secondary education			36.63
	0 otherwise			63.37
YOE	Years of education	11.54	3.31	
AGE	Age	48.77	18.01	
EXP	Professional experience years	22.36	13.91	
WSEC	Work sector			
	public (*PUB*)			14.21
	private (*PRIV*)			24.08
	own business (*ENT*)			3.20
	other (*OTH*)			58.51
CTYPE	The type of residence			
	BCITY (more than 100 thousand occupants)			23.92
	MCITY (less than 100 thousand occupants)			32.90
	VIL (villages)			43.18
F	Gender			
	1 if female			54.79
	0 otherwise			45.21
EAST	Geographical localization[1]			
	1 if eastern Poland			26.52
	0 otherwise			73.48

[1]Eastern Poland includes the Lubelskie, Podkarpackie, Podlaskie, Świętokrzyskie and Warmińsko-Mazurskie regions in accordance with the division incorporated in the European Operational Programme Development of Eastern Poland. These provinces are considered to have lower living standards, a lower dynamic of economic development, poorly developed and inadequate transport infrastructure and insufficient growth factors, which might be reflected in the earnings of their residents.

Since heteroscedasticity of the error term was detected in the large majority of cases (using White's test (cf. White, 1980)), for model estimation the weighted least squares method was applied. For a more detailed analysis and additional information see Król (2014).

Table 7 presents estimation results of five different specifications of Mincer's model in the year 2010 based on the SOEP data set (dependent variable

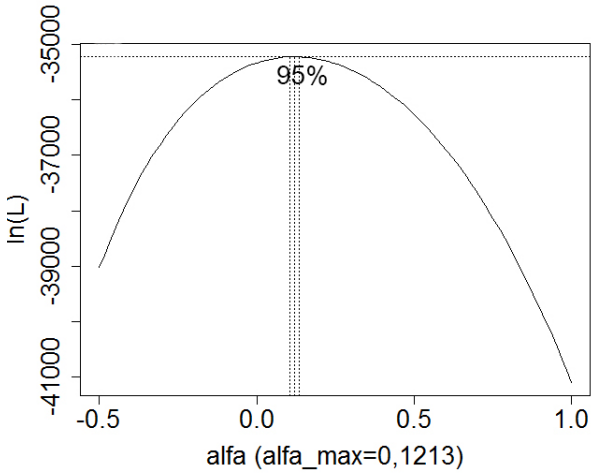

Fig. 1 Values of the logarithm of likelihood function for various values of α parameter for specification (SOEP4).

ln(RHGEAR)), starting from the simplest classic Mincer model SOEP1 describing earnings by the number of years of education and experience to the most complex model SOEP5. Model SOEP5 was chosen for further interpretation, since it has the highest goodness-of-fit measure and the lowest AIC information criterion, moreover all its variables are statistically significant. The interpretation of results of estimation of the model SOEP5 shows that in the year 2010 the Germans with higher education could earn about 25% more in comparison to similar (in terms of gender, work experience, work type, size of the company etc.) persons. Women in Germany earned in 2010 on average about 15% less than men on similar work positions and with similar professional experience. The professionals and freelancers in the year 2010 could earn about 31% more, the managers about 33% more and the trainees about 31% less than regular employees (specialists and other employees). This result shows that labour market requirements extend beyond simple higher education diploma, and that additional qualifications and skills are also important. The differences in earnings in big and small firms may also be observed: employees of small companies in the year 2010 earned about 24% less and in medium companies about 10% less than employees in big corporations, *ceteris paribus*.

Table 8 presents estimation results of five specifications of Mincer's model in the year 2009 based on the model Social Diagnosis data set (dependent

Table 7 Estimation results of five different specifications of Mincer's model (year 2010) based on the SOEP data set (dependent variable *ln(RHGEAR)*)

	SOEP1	SOEP2	SOEP3	SOEP4	SOEP5
constant	0.8636***	1.839***	1.518***	2.207***	2.090***
	(0.03461)	(0.02436)	(0.07071)	(0.02290)	(0.06951)
YOET	0.08217***				
	(0.001909)				
PWE	0.05246***	0.05220***	0.05225***	0.02572***	0.02577***
	(0.001943)	(0.002006)	(0.002001)	(0.001813)	(0.001803)
PWE^2	−0.0008456***	−0.0008823***	−0.0008815***	−0.0004860***	−0.0004867***
	(3.867e-05)	(3.916e-05)	(3.912e-05)	(3.522e-05)	(3.508e-05)
HEDU		0.4192***	0.4148***	0.2231***	0.2229***
		(0.01225)	(0.01226)	(0.01046)	(0.01041)
MEDU			0.3236***		0.1181*
			(0.06703)		(0.06579)
SEN				0.02148***	0.02151***
				(0.001403)	(0.001396)
SEN^2				−0.0003143***	−0.0003161***
				(3.659e-05)	(3.644e-05)
FEM				−0.1679***	−0.1684***
				(0.008819)	(0.008774)
APP				−0.3813***	−0.3749***
				(0.02136)	(0.02131)
PROF				0.2694***	0.2682***
				(0.01036)	(0.01031)
MAN				0.2893***	0.2881***
				(0.01771)	(0.01765)
SMALL				−0.2836***	−0.2847***
				(0.01369)	(0.01361)
MEDIUM				−0.1084***	−0.1081***
				(0.01028)	(0.01021)
n	9534	9534	9534	8787	8787
$\bar{R}^2$	0.2219	0.1711	0.1731	0.4153	0.4172
AIC	41081.68	40710.35	40745.02	38393.00	38289.02

Significance levels ('***' = 0.01; '**' = 0.05; '*' = 0.1), n = number of observations , $\bar{R}^2$ adjusted R^2 and AIC = Akaike Information Criterion.

variable *ln(RMNEAR)*). Again the most complex specification (SD5) is taken for interpretation and further research. The premium for higher education in Poland in the year 2009 was about 29%. Note that in the Social Diagnosis data set the values of variables *HE* and *ME* for the persons with higher education are 1, whereas in the SOEP data set for the persons with higher education *HEDU*=1 and *MEDU*=0. The difference in earnings between males and females in Poland is bigger than in Germany. In the analysed period women earned on average about 19% less than men doing similar work in much the same work

Table 8 Estimation results of five different specifications of Mincer's model (year 2009) based on the Social Diagnosis data set (dependent variable *ln(RMNEAR)*)

	(SD1)	(SD2)	(SD3)	(SD4)	(SD5)
constant	6.133***	6.873***	6.738***	6.916***	6.795***
	(0.01718)	(0.01060)	(0.01213)	(0.01385)	(0.01437)
YOE	0.06484***				
	(0.001002)				
EXP	0.01195***	0.01404***	0.01354***	0.01105***	0.01288***
	(0.0009639)	(0.0009927)	(0.0009839)	(0.0008900)	(0.0008705)
EXP^2	−0.0001213***	−0.0002492***	−0.0002102***	−8.475e-05***	−0.0001101***
	(1.964e-05)	(2.104e-05)	(2.004e-05)	(1.867e-05)	(1.793e-05)
HE		0.4807***	0.6024***	0.3470***	0.4699***
		(0.009995)	(0.01068)	(0.009400)	(0.01008)
ME			0.2887***		0.2116***
			(0.008038)		(0.007129)
MCITY				−0.08492***	−0.06582***
				(0.008162)	(0.008016)
VIL				−0.2194***	−0.1637***
				(0.008138)	(0.008188)
PUB				0.4295***	0.3852***
				(0.009388)	(0.009046)
PRIV				0.3909***	0.3662***
				(0.008321)	(0.008243)
ENT				0.5190***	0.4745***
				(0.01904)	(0.01881)
EAST				−0.08472***	−0.08729***
				(0.007100)	(0.006818)
F				−0.1933***	−0.2138***
				(0.006624)	(0.006485)
n	18417	18426	18426	18370	18370
$\bar{R}^2$	0.1904	0.1182	0.1642	0.3330	0.3660
AIC	77441.29	76479.20	75850.83	77973.57	77847.24

Significance levels ('***' = 0.01; '**' = 0.05; '*' = 0.1), n = number of observations , $\bar{R}^2$ adjusted R^2 and AIC = Akaike Information Criterion.

place. The influence of residence on the level of wages was significant as well. Eastern Polish regions, which are considered to have lower living standards, lower dynamic of economic development and insufficient growth factors, show significantly lower earnings. In comparison to Central and West Poland the people from eastern provinces earned about 8% less, *ceteris paribus*. Moreover, the residents of small and medium cities earn about 6% and residents of villages about 15% less than the inhabitants of big cities.

Tables 9 and 10 present the final estimated models for the years 1995–2010 for the SOEP data set and for the years 2003 – 2011 for the Social Diagnosis data

Table 9 Estimation results of Mincer's model in years 1995, 2000, 2005, 2010 based on the SOEP data set (dependent variable *ln(RHGEAR)*)

	(SOEP1995)	(SOEP2000)	(SOEP2005)	(SOEP2010)
constant	2.421**	2.502**	2.243**	2.090**
PWE	0.02106**	0.001609**	0.03010**	0.02577**
PWE^2	−0.0004066**	−7.616e-07**	−0.0005706**	−0.0004867**
HEDU	0.09978**	0.1385**	0.1826**	0.2229**
MEDU	−0.05301	0.04387	0.02129	0.1181*
SEN	0.01831**	0.02322**	0.02000**	0.02151**
SEN^2	−0.0004006**	−0.0004539**	−0.0003263**	−0.0003161**
FEM	−0.1897**	−0.1899**	−0.1789**	−0.1684**
APP	−0.3114**	−0.4339**	−0.4124**	−0.3749**
PROF	0.2190**	0.2301**	0.2636**	0.2682**
MAN	0.1461**	0.1861**	0.2544**	0.2881**
SMALL	−0.2908**	−0.2744**	−0.2894**	−0.2847**
MEDIUM	−0.1080**	−0.1101**	−0.1202**	−0.1081**
n	7018	12423	9697	8787
$\bar{R}^2$	0.3365	0.3474	0.4088	0.4172
AIC	29487.92	53348.74	42004.25	38289.02

Significance levels ('***' = 0.01; '**' = 0.05; '*' = 0.1), n = number of observations , $\bar{R}^2$ adjusted R^2 and AIC = Akaike Information Criterion.

Table 10 Estimation results of Mincer's model in years 2003, 2005, 2007, 2009, 2011 based on the Social Diagnosis data set (dependent variable *ln(RMNEAR)*)

	(SD2003)	(SD2005)	(SD2007)	(SD2009)	(SD2011)
constant	6.010***	6.067***	6.131***	6.795***	6.677***
HE	0.5349***	0.4349***	0.4805***	0.4699***	0.4879***
ME	0.2171***	0.2226***	0.2286***	0.2116***	0.2353***
MCITY	−0.09271***	−0.07576***	−0.07770***	−0.06582***	−0.04613***
VIL	−0.2105***	−0.1382***	−0.1452***	−0.1637***	−0.1384***
PUB	0.4884***	0.4087***	0.4276***	0.3852***	0.4390***
PRIV	0.4486***	0.3173***	0.3862***	0.3662***	0.3959***
ENT	0.6216***	0.4372***	0.5128***	0.4745***	0.4716***
EAST	−0.05617***	−0.05287***	−0.03427***	−0.08729***	−0.07752***
F	−0.2324***	−0.1882***	−0.2030***	−0.2138***	−0.2219***
n	6707	5802	9205	18370	18661
$\bar{R}^2$	0.3497	0.3028	0.3081	0.3660	0.3545
AIC	29570.07	24745.31	39952.75	77973.57	80272.48

Significance levels ('***' = 0.01; '**' = 0.05; '*' = 0.1), n = number of observations , $\bar{R}^2$ adjusted R^2 and AIC = Akaike Information Criterion.

set. The obtained results allow for the evaluation of the dynamics of influence of higher education and other factors on the earnings in Germany and Poland accordingly.

In Germany in the last 15 years we observe a quite stable increase in the value of the premium for higher education (from 11% to 25%). The auxiliary factors whose influence changed the most are the ones connected with the type of work. For example, the premium for managers increased from about 15% in the year 1995 to about 33% in the year 2010. Another interesting trend observed in the analyzed period is the slight decrease of gender-related work discrimination (from about 17% to about 15%).

The analysis of Polish data shows the stabilization of the influence of tertiary education on the level of earnings. In the years 2005 – 2011 the premium for higher education oscillates around the level of 30%. Similarly, the earnings of women in the analyzed period remain lower than those of men of about 20%. There is a slight improvement in the reduction of regional differences. The difference in earnings of the residents of villages in comparison to the inhabitants of big cities decreased from about 19% in 2005 to about 15% in 2011.

4.2 Determination of the significance of differences in incomes before and after reaching the higher education

Another part of our research was to check if there is a significant difference in incomes before and after reaching a higher education degree and, in addition, to measure the rate of return to education in both groups (cf. Targaszewska, 2014). To achieve those goals the Wilcoxon matched-pairs signed-rank test and the classical Mincerian function were applied. The Wilcoxon matched-pairs signed-rank test is non-parametric test used to compare two paired (dependent) samples – each observation of the first sample has a unique connection with an observation in the second sample. The null hypothesis states the equality of median difference in paired observations. This means: samples have identical distributions (cf. Jackson, 2011, pp. 266, 267). The research was based on the group of respondents which fulfilled the following conditions:

- They participated in the SD-project in the years since 2003,
- they declared a lower than tertiary level of education,

- and they declared in 2011 to have a higher level of education than in the previous studies.

For both groups the rate of return to education is measured and compared (Targaszewska, 2014). Because of the nature of the data some assumptions were made. Firstly, in a situation where over the years, some respondents changed their level of education more than once, the research included only the most recent change. Secondly, the variable denoting years of experience in 2011 was estimated. Experience in 2011 is equal to experience in 2009 plus two years. Lastly, incomes were corrected by the inflation indicator (with the base year 2003). Moreover, the cases with incomes under the minimum wages in each year were removed from the research. Finally 152 cases were taken into account. The p-value for the executed test was almost 0.0. This allows to reject the null hypothesis (of equality of median difference in paired observations) which means that there is a significant difference in incomes between groups. Wages of persons with a higher education degree are on average higher by 773 PLN in comparison to persons without this kind of education. 2453 [PLN] and 3226 [PLN] are the means "before" and "after" reaching the level of higher education, respectively.

Subsequently, the rate of return to education and rate of return to experience for both groups "before" and "after" were estimated by the classical Mincer model (cf. Mincer, 1974):

$$\log(Y) = \alpha + \rho s + \beta_0 x + \beta_1 x^2 + \xi, \tag{3}$$

where Y is earnings, s is schooling level or years of study, x is work experience. The parameter ρ can be interpreted as the average private rate of return to schooling, β is related to the financial return to experience, and α is related to initial earnings capacity (cf. Polachek, 2008). The estimated models are presented in Table 11.

In the first model for the group "before" the parameter p is not significant at the 5% significance level. It seems that for persons without higher education degree the most important variable is experience. After rejecting the variable "years of study" the model was estimated once again. From the new model for group "before" one can conclude that the rate of return to experience, after 10 years of working is nearly 1.4%. For the group "after" each of the parameters is significant. The rate of return to education is about 6.6% and rate of return to experience after 10 years of working is about 1.1%.

Table 11 Parameter estimates of Mincer's model based on the data set Social Diagnosis (2003–2011)

	Group					
	before		before without years of study		after	
Variable	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Constant	7.474	0.000	7.458	0.000	6.584	0.000
Years of study (s)	-0.001	0.943	–	–	0.066	0.002
Experience (x)	0.034	0.000	0.034	0.000	0.031	0.001
Quadratic experience (x^2)	-0.001	0.006	-0.001	0.006	-0.001	0.017
n	152		152		152	
$\bar{R}^2$	0.1465		0.1522		0.1792	
AIC	158.39		156.36		143.84	

n = number of observations , $\bar{R}^2$ adjusted R^2 and AIC = Akaike Information Criterion.

4.3 Examination of non-monetary benefits of tertiary education

Our next research step for measuring the effectiveness of education was to capture the intangible benefits of higher education, particularly non-monetary private and social rates of return on investment in education. Empirical studies were carried out on data from the Social Diagnosis 2011 data set. As shown in Table 1, non-monetary returns are an important part of the benefits of education. It is commonly believed that better educated people have a better life. This general opinion can be empirically confirmed in two ways. Firstly, by people's personal experience and, secondly by the statements of the respondents concerning their life quality perception and expectations along with their level of education. Fig. 2 visualizes the output of the correspondence analysis performed on this data. A comprehensive description of the algorithm of correspondence analysis, computational details, and its applications can be found in the classic text by Greenacre (1984). Fig. 2 shows the coincidence of the respondents' education level with a subjective evaluation of the happiness with his/her life in the last years.

The position of higher education in Fig. 2 is close to the most positive assessment of one's life in the last year. The percent of total inertia described in the two first dimensions is almost 100%. In Figs. 2 and 3, the first dimension describes about 98% of inertia. In Fig. 2, 'very happy' is next to 'high education', and further to the right the 'education level' is lower, and the 'assessment of life in the last year' is also getting worse. Basic education is near to negative assessment. The conclusion is that better education is associated with a more

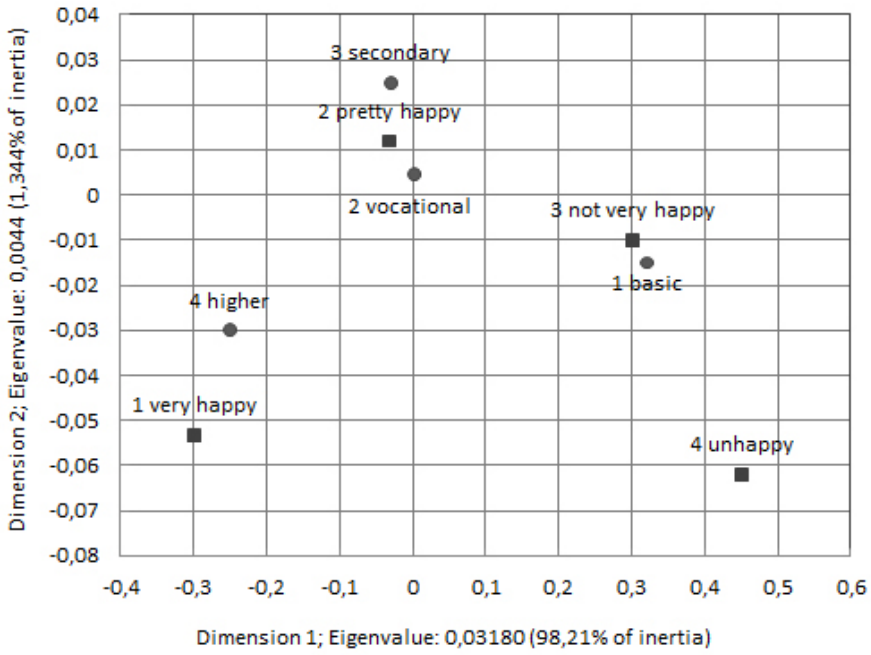

Fig. 2 Correspondence analysis of education variable and assessment of life in the last year, the Social Diagnosis 2011, sample size: 26332. Education Levels: 1 – basic education, 2 – vocational, 3 – secondary and 4 – higher. Frequency of health problems: 1 – often, 2 – sometimes and 3 – never.

positive perception of the past. But when respondents were asked to name the three most important conditions for a successful and wonderful life, the first five positions were health, children, happy marriage, work and money. The education level was mentioned somewhere between the 13^{th} and 10^{th} place out of 14 possible places (higher position for better educated respondents). This surprising phenomenon can be explained by the association of education with higher earnings and better work.

Private non-monetary returns of tertiary education include the impact of education on personal health, the ability to enjoy leisure and the capacity to make personal choices. Obviously, education tends to improve income which affects health positively. People with a higher education level are more aware of healthy behaviour and demonstrate more tendencies to seek treatment when needed. More results of the analysis of non-monetary benefits of tertiary education can be found in Dziechciarz-Duda and Król (2013).

According to the WHO Regional Office for Europe (2012), the male population (age of 30) with higher education will live on average another 48.5 years. While the male population (age of 30) with primary education will live on average another 36.5 years and for secondary education another 43 years. For women, life expectancy is in general higher: for better educated women on average 83.2 years and for the least educated women 5 years shorter. Moreover, differences in the risk of death related to the educational level are greater in the case of men than women for all causes of death (except cardiovascular diseases). Death rates from all main causes tend to be lower among people with higher education levels. All diseases contribute to shortening the lives of less-educated people more than the lives of better educated individuals. The cause that is most responsible for shortening the lives of less-educated people when compared with better educated individuals are cardiovascular diseases, external causes and cancer. Numerous research results confirm that higher education contributes to increased longevity and better health in terms of severe and fatal diseases, partly through the increased earnings that enable the purchase of better health care and a better diet (cf. WHO Regional Office for Europe, 2012).

Fig. 3 visualizes the output of the correspondence analysis for the level of education of Polish respondents in 2011 with the subjective evaluation of self-well-being expressed as the frequency of health problems that hinder a positive perception of the quality of life. The results support the hypothesis of the positive impact of education on personal health. In Fig. 3, "often" is next to "basic education" and further, to the right the level of education is growing and the assessment of health is better.

4.4 Analysis of employment status and professional profiles of universities graduates

The goal of our last research step was to analyse the professional situation of young people with tertiary education. For this purpose a hierarchical classification method (Ward) was applied to the data from the Study of Human Capital in Poland 2012, and 8 homogenous classes of university graduates were distinguished based on the dendrogram. The analysis of the characteristics of each class is a valuable source of information about the factors that have an impact on the level of unemployment in this group. The following variables describing the situation of graduates on the labour market were used: profes-

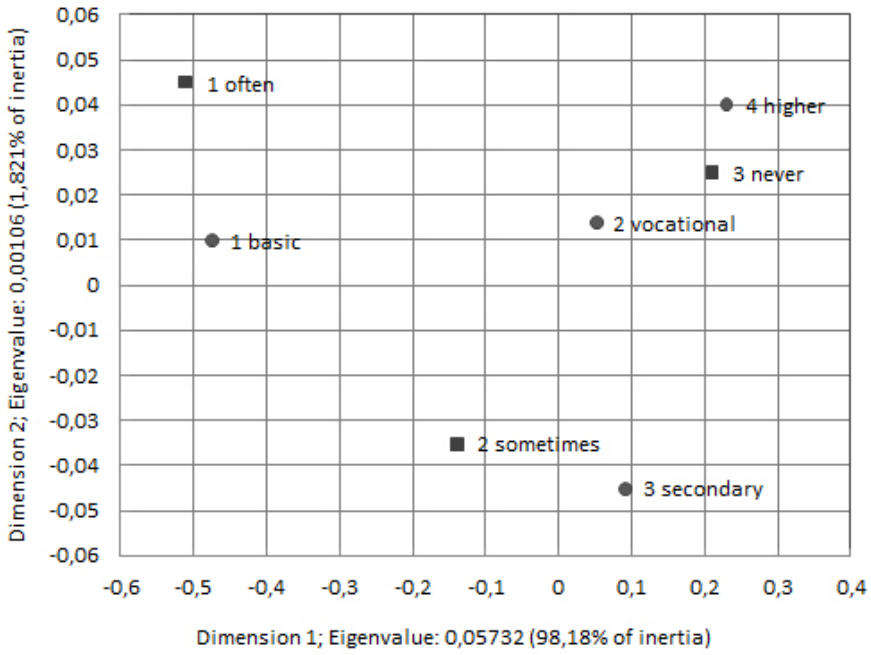

Fig. 3 Correspondence analysis of education and the frequency of health problems, the Social Diagnosis 2011, sample size: 26332. Education Levels: 1 – basic education, 2 – vocational, 3 – secondary and 4 – higher. Frequency of health problems: 1 – often, 2 – sometimes and 3 – never.

sional status (full-time job, part-time job, unemployed, housework, etc.), the type of university (private, public), the type (full-time studies, evening studies, extramural studies), the level of studies (bachelor, engineer, master, etc.) as well as the average level of net income.

The analysis of the professional situation and characteristics of the graduates in separate classes allowed for the assessment of how well the representatives of each group cope with the labour market challenges (see Fig. 4). The worst groups, in terms of the percentage of employed and earnings level, were the young, out-of-work people with a bachelor's degree, graduated from private universities (class 5), young people without work and experience (class 8), as well as young people from small towns and villages, graduated from agriculture and service studies (class 2). The level of employment and earnings in class 3 (teachers and humanistic studies graduates) is similar to the average in the whole population. Whereas the situation of the classes 1 (well-paid engineers),

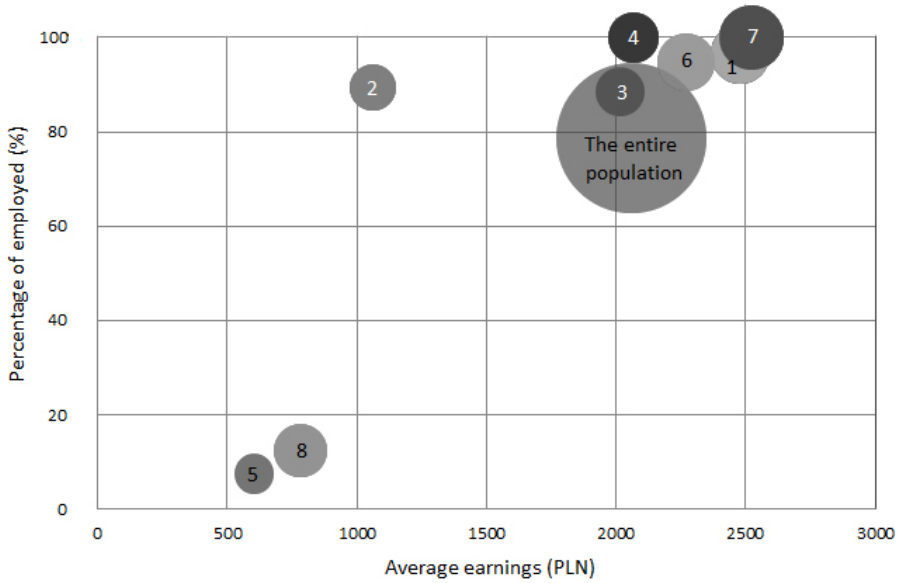

Fig. 4 Assessment of groups according to the percentage of employed and earnings, Study of Human Capital in Poland 2012.

4 (employed economists with significant experience), 6 (working with a bachelor's degree, graduated from private universities) and 7 (entrepreneurial with a master's degree) is significantly better. Comprehensive results of the research can be found in Dziechciarz-Duda and Przybysz (2014).

These results show that graduates with a bachelor degree are in less favorable condition compared to graduates with a master degree. The percentage of employment in the group of bachelor degree holders is only 63,8%, whereas for graduates with master degrees it increases to 80,7%. A similar situation may be observed for graduates of technical universities – the employment rate of undergraduate engineers is 76,5%, while in the group of engineers with master degrees it is 85,3%. The graduates of engineering and technical studies, as well as mathematics, statistics, physics and medicine grads occupy the strongest position on the labour market. The students of the most popular majors (economics, pedagogics and social studies) face an employment rate of about 80% and an unemployment rate of almost 15%.

5 Final remarks

This paper summarizes various results obtained in the framework of the research project *Methods of Measuring the Return on Investment in Higher Education*. The goal of the project was to analyse the problem of measuring the effectiveness of investment into higher education in its various forms. The research approaches, included classical methods (ANOVA, Mincerian earnings function, correspondence analysis, hierarchical agglomerative clustering), as well as new ideas (application of Wilcoxon Matched-Pairs Signed-Rank Test to determine the significance of differences in incomes before and after reaching higher education). All obtained results support the hypothesis that higher education influences the level of income. Moreover, the estimated pseudo rates of return to education provide the basis for the evaluation of the effectiveness of private investment in education.

Acknowledgements The study was conducted in the framework of the research project entitled *Rate of Return Measurement Methods in Higher Education* (*Metody pomiaru stopy zwrotu z inwestycji na edukację w szkołach wyższych*). The project has been financed by the National Science Centre on the basis of decision no. DEC-2011/01/B/HS4/02328.

References

Aubyn M, Pina A, Garcia F, Pai J (2008) Study on the efficiency and effectiveness of public spending on tertiary education. `http://ec.europa.eu/economy_finance/publications/publication16267_en.pdf`, accessed: 2014-12-30

Bilans Kapitału Ludzkiego (2012) `http://bkl.parp.gov.pl`, accessed: 2013-06-03

Box GEP, Cox DR (1964) An analysis of transformations. Journal of the Royal Statistical Society Series B (Methodological) 26(2):211–252

Dziechciarz J (2011) On rate of return measurement in education. Econometrics 31:49–66

Dziechciarz-Duda M, Król A (2013) On non-monetary benefits of tertiary education. Econometrics 41(3):78–94

Dziechciarz-Duda M, Przybysz K (2014) Wykształcenie a potrzeby rynku pracy. Klasyfikacja absolwentów wyższych uczelni. Taksonomia Klasyfikacja i analiza danych – teoria i zastosowania 22(327):303–312

Etzkowitz H, Peters LS (1991) Profiting from knowledge: Organisational innovations and the evolution of academic norms. Minerva 29(2):133–166, DOI 10.1007/BF01096406

European Commission (2006) Communication from the Commission to the Council and the European Parliament. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2006:0208:FIN:en:PDF, accessed: 2013-12-03

Field A (2005) Discovering Statistics Using SPSS. SAGE, London

Greenacre M (1984) Theory and applications of correspondence analysis. Academic Press, New York

Heckman J, Lochner L, Todd P (2003) Fifty Years of Mincer Earnings Regressions. NBER Working Papers (9732), National Bureau of Economic Research

Jackson S (2011) Research Methods and Statistics: A Critical Thinking Approach. Wadsworth Publishing Company, Belmont

Jongbloed B (2010) Funding Higher Education: A View across Europe. University of Twente, Enschede

Król A (2014) An analysis of the dynamics of higher education's influence on the level of wages. Econometrics 43(1):60–73

McMahon WW (1997) Recent advances in measuring the social and individual benefits of education. International Journal of Education Research 27(6):447–531, DOI 10.1016/S0883-0355(97)00047-5

Mincer J (1958) Investment in human capital and personal income distribution. Journal of Political Economy 66(4):281–302

Mincer J (1974) Schooling, Experience and Earnings. Columbia University Press, New York

Morgan GA, Leech NL, Gloeckner GW, Barrett KC (2004) SPSS for Introductory Statistics: Use and Interpretation. Lawrence Erlbaum Associates, Mahwah

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

Polachek SW (2008) Earnings over the life cycle: The mincer earnings function and its applications. Foundations and Trends in Microeconomics 4(3):165–272, DOI 10.1561/0700000018

Proust M (2009) Statistics and Graphics Guide. SAS, Cary

Psacharopoulos G (1995) The Profitability of Investment in Education: Concepts and Methods. HCO Working Papers 63

Psacharopoulos G (2009) Returns to Investment in Higher Education. A European Survey. Center for Higher Education Policy Studies, Enschede

Rada Monitoringu Społecznego (2003-2011) Diagnoza Społeczna: zintegrowana baza danych, 2003-2011. www.diagnoza.com, accessed: 2012-08-04

Targaszewska M (2013) An attempt at identification sources of variation in monthly net incomes among persons with tertiary education. Econometrics 39(1):210–220

Targaszewska M (2014) An attempt at measuring the effectiveness of higher education in Poland. Econometrics 43(1):50–59

Wagner GG, Frick JR, Schupp J (2007) The German Socio-Economic Panel Study (SOEP) - scope, evolution and enhancements. Schmollers Jahrbuch 127(1):139–169

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

White H (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48(4):817–838

WHO Regional Office for Europe (2012) Social inequalities in health in Poland. http://www.euro.who.int/__data/assets/pdf_file/0008/177875/E96720.pdf, accessed: 2013-09-17

Wissema JG (2009) Towards the Third Generation University. Managing the University in Transition. EE Publishing, Cheltenham