EKONOMETRIA ECONOMETRICS 1(35) • 2012

ISSN 1507-3866

Iwona Skrodzka
University of Bialystok
i.skrodzka@uwb.edu.pl

THE USAGE OF SOFT MODELLING TO RESEARCH THE RELATIONSHIP BETWEEN HUMAN CAPITAL AND THE ECONOMIC DEVELOPMENT LEVEL IN POLISH REGIONS[1]

Abstract: The aim of this article is to present research results of the relationship between human capital and economic development level in Polish regions. The method which was used in this study was soft modelling. Soft modelling enables to investigate the relationship among unobserved variables (latent variables). Human capital and economic development can be classified into these types of categories. The presented soft model contained four latent variables: human capital, investments in human capital, physical capital and the level of economic development. The results of the research allowed us to verify which of two forms of capital: human or physical, had the stronger influence on the level of economic development in Polish regions. Moreover they enabled to put regions into order in terms of the human capital stock, human capital investments and the economic development level.

Keywords: human capital, economic development, soft modelling.

1. Introduction

Human capital is defined as a stock variable that represents the capacity of an individual (household, nation) to generate a sustained flow of earned income [Dagum 2004, p. 1]. Recent studies conclude that human capital is the key growth factor [Próchniak 2006,; Cichy, Malaga 2007; Florczak 2007]. However, existing research, focusing on cross-country growth performance, has produced contradictory results. Several economists argue that it derived from poor quality data used in some of those studies [de la Fuente, Domenech 2006, p. 1]. The choice of the human capital proxy is barely reflected upon and depends on data availability very much in most empirical growth research [Florczak 2007, p. 112].

Human capital should be considered as a complex, multifaceted concept with various intangible dimensions that are not directly observable and that cannot be

[1] Research cost was covered by "Wiedza i praktyka – klucz do sukcesu w biznesie" project, co-financed by the European Union from the European Social Fund and the state budget.

measured with precision by a single attribute, a set of attributes, or their combined sum [Le, Gibson, Oxley 2005, p. 4; Łukasiewicz 2009, p. 96].

The aim of this article is to present research results of the relationship between human capital and economic development in Polish regions in 2008.[2] Human capital is defined as an unobserved variable reflected by education, knowledge, skills, work experience and health embodied in a region's society [Domański 1993, p. 19; Marciniak 2000, pp. 157, 158]. The method which was used in this research is soft modelling.

Soft modelling enables to investigate relationships among unobserved variables (latent variables) [Wold 1980; Rogowski 1990; Mierzyńska 1999; Perło 2004]. Values of these variables cannot be measured because there is a lack of general definition or measurement method. The soft model consists of two submodels:

– internal submodel which presents the system of relationships between latent variables,
– external submodel which contains latent variables definitions [Rogowski 1990, s. 33].

Latent variables can be defined on the basis of deductive or inductive approach. Deductive approach assumes that indicators reflect latent variable. Inductive approach assumes that indicators form latent variable. The choice of approach depends on the theory or intuition of researcher [Rogowski 1990, pp. 25, 26].

The estimation of soft model parameters can be done based on the Partial Least Square [Wold 1980, pp. 337, 338; Rogowski 1990, pp. 37–44]. Test Stone-Geisser and "2s" rule are used to verify the model [Rogowski 1990, pp. 47–54].

2. Specification of the internal model

Estimated model consisted of two following equations

$$\mathrm{HC}_t = \alpha_0 + \alpha_1 \mathrm{INHC}_{t-1} + \xi_1, \tag{1}$$

$$\mathrm{LED}_t = \beta_0 + \beta_1 \mathrm{HC}_t + \beta_2 \mathrm{PC}_t + \xi_2, \tag{2}$$

where: HC – human capital, INHC – investments in human capital, PC – physical capital, LED – the level of economic development, α_0, α_1, β_0, β_1 – structural parameters, ξ_1, ξ_2 – error terms, t – 2008.

The aim of equation (1) was to investigate the relationship between human capital investments and its stock. Equation (2) was used as the tool to identify the relationship between human capital and the level of economic development in Polish regions. It also enabled to verify which of two capital forms, human or physical, has the stronger influence on the level of economic development. The scheme of internal model relations was presented in Figure 1.

[2] Data availability influence on the year choice.

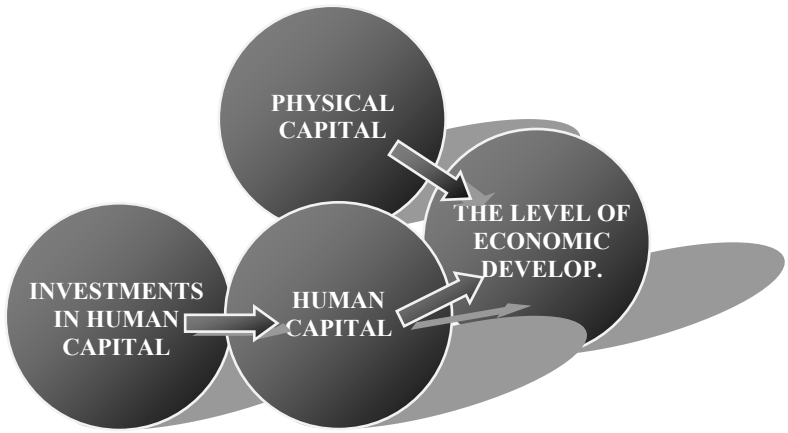

Figure 1. Internal model scheme

Source: own elaboration.

3. Specification of the external model

The model shown in Figure 1 contained four latent variables. Three of them: "human capital", "investments in human capital" and "the level of economic development" were not directly observed. The fourth – "physical capital" was quasi-unobserved. Unobserved variables were defined by the group of indicators. A deductive approach was used to define the above variables. In this approach, it is assumed that the latent variable is a theoretical category and it is necessary to investigate empirical data (observed variables) reflecting this variable [Rogowski 1990, p. 33; Mierzyńska 1999, p. 34].

Data used to specify the model were obtained from Local Data Bank[3] and they refer to 2008. Many indicators were covered in the GUS statistics database. Analysis of all indicators would be unclear and difficult to interpret, hence a selection was necessary. The criteria were the following:

– universality (commonly respected indicators),
– comparability (indicators as coefficients of intensity),
– variety (coefficient of variation higher than 10%).

The internal model of HC latent variable contains observed variables which reflect education, knowledge, skills and health in the region. Fifty four indicators were considered but finally thirteen of them were qualified to the model. They are presented in Table 1.

Most of the HC indicators are stimulants (their high value indicates the high stock of human capital). Only three indicators: HC11, HC12, HC13 are destimulants

[3] http://www.stat.gov.pl/bdlen/app/strona.html?p_name=indeks (12.12.2011).

(their high value reflects the low stock of human capital). They represent the health component.

Table 1. Human capital indicators

Indicator code	Indicator name
HC01	Percentage of population with tertiary education level.
HC02	Percentage of employed with tertiary education level.
HC03	Life-long learning of persons aged 25–64 (%).
HC04	Students of post-graduate studies per 1 thousand population.
HC05	Students of doctoral studies per 10 thousand population.
HC06	Students per 1 thousand population.
HC07	Graduates per 1 thousand population.
HC08	Percentage employed in R&D sector.
HC09	Employed in R&D in enterprises sector per 10 thousand employee.
HC10	Percentage of employed using computer with Internet access at least once a week.
HC11	Deaths because of diseases of nervous system and sense organs per 100 thousand population.
HC12	Infant deaths per1 thousand live births.
HC13	Persons receiving retirement pay and pension from non-agricultural social security system per 1 thousand population, pension due to an inability to work.

Source: own elaboration.

The internal model of the INHC latent variable contains six observed variables which refer to investments in education, health and knowledge (see Table 2). All of them are stimulants.

Table 2. Investments in human capital indicators

Indicator code	Indicator name
INHC01	Public and private expenditures on education per capita.
INHC02	Households expenditures on education per capita.
INHC03	Public and private expenditures on health care per capita.
INHC04	Households expenditures on health care per capita.
INHC05	Expenditures on R&D per capita.
INHC06	Enterprises expenditures on R&D per capita.

Source: own elaboration.

The internal model of the LED latent variable contains ten indicators (see Table 3). They reflect the economic potential of the region, structure of employment, investment potential, entrepreneurship and society well-being. Only one indicator (LED03) is a destimulant.

Table 3. Level of economic development indicators

Indicator code	Indicator name
LED01	Gross domestic product per capita.
LED02	Gross value addend per employee.
LED03	Percentage of employed in agriculture.
LED04	Percentage of employed in services.
LED05	Investment outlays per capita.
LED06	Investment outlays in enterprises per capita.
LED07	Entities entered in the REGON register per 10 thousand population.
LED08	New entities of the national economy recorded in the REGON register per 10 thousand population.
LED09	Retail sales per capita.
LED10	Average monthly income per capita (disposable).

Source: own elaboration.

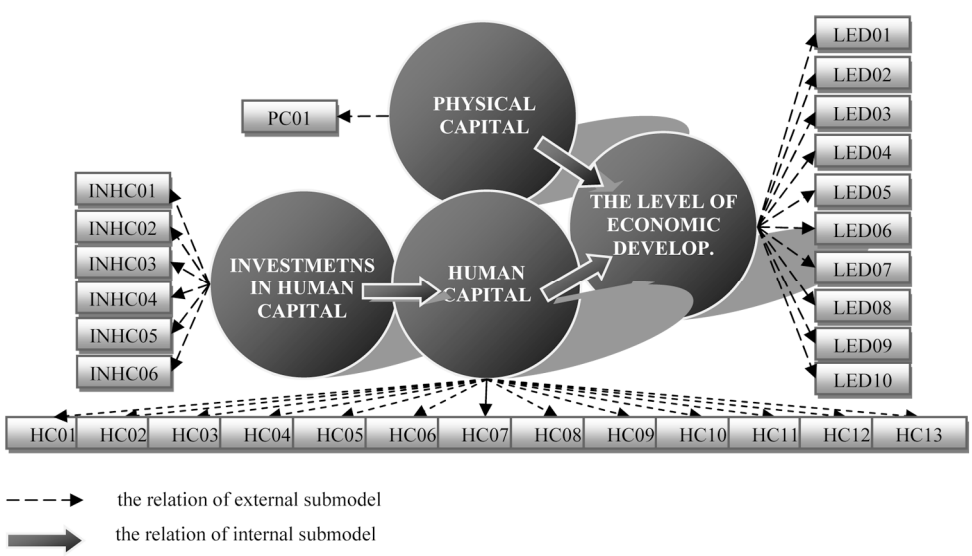

Figure 2. Internal and external relations scheme of soft model

Source: own elaboration.

Figure 2 presents internal and external relations scheme of the soft model. The model contains four latent variables and thirty observed variables. Four indicators: HC11, HC12, HC13, LED03 are destimulants, the rest of them are stimulants.

4. Estimation results

The soft model was estimated with the aid of PLS software created by J. Rogowski.[4]

Table 4 contains weight and loadings estimates with regard to the HC latent variable external model. The errors calculated based on the Tukey cut method are given in brackets. The results are consistent with expectations. Stimulants have positive weights and loadings and destimulants have negative ones. Moreover all parameters are statistically significant.[5]

Table 4. Estimates of HC variable external model

Indicator code	Weight (*error*)	Loading (*error*)
HC01	0,0950 (*0,0023*)	0,7752 (*0,0025*)
HC02	0,1187 (*0,0016*)	0,8620 (*0,0011*)
HC03	0,1025 (*0,0020*)	0,7285 (*0,0017*)
HC04	0,1308 (*0,0007*)	0,9408 (*0,0002*)
HC05	0,1132 (*0,0007*)	0,8830 (*0,0014*)
HC06	0,1083 (*0,0008*)	0,8671 (*0,0008*)
HC07	0,0964 (*0,0011)*	0,7858 (*0,0006)*
HC08	0,1174 (*0,0012)*	0,8867 (*0,0013)*
HC09	0,1218 (*0,0008)*	0,8545 (*0,0005)*
HC10	0,1356 (*0,0003)*	0,9136 (*0,0003)*
HC11	–0,0484 (*0,0019)*	–0,2773 (*0,0027)*
HC12	–0,0042 (*0,0019)*	–0,0843 (*0,0031)*
HC13	–0,0278 (*0,0020)*	–0,2704 (*0,0028)*

Source: own elaboration based on soft model results.

[4] PLS software is available at Faculty of Economics and Management University of Bialystok.

[5] Parameter is statistically significant when value of double error is higher than value of estimator ("2s" rule).

Indicators: "students of post-graduate studies per 1 thousand population" (HC04) and "percentage of employed using computer with Internet access at least once a week" (HC10) are the most strongly correlated with the HC variable. The indicators: "percentage employed in R&D sector" (HC08), "students of doctoral studies per 10 thousand population" (HC05), "students per 1 thousand population" (HC06), "percentage of employed with tertiary education level" (HC02) and "employed in R&D in enterprises sector per 10 thousand employees" (HC09) have a strong influence on the HC variable. Indicators which represent the health component (HC11, HC12, HC13) reflect the HC variable poorly. To sum up, education and knowledge are the most significant components of human capital in Polish regions.

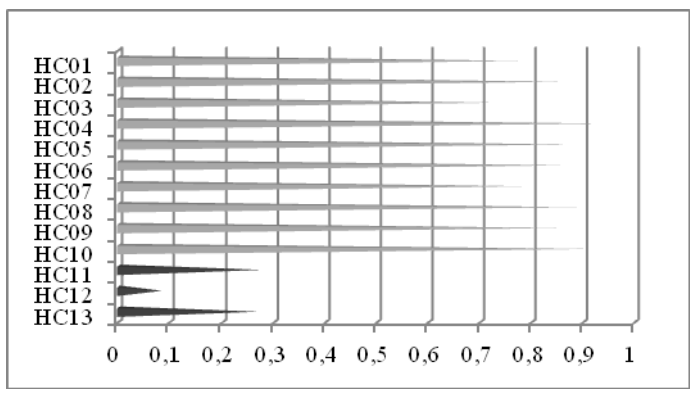

Figure 3. Loadings of HC variable (the absolute value)*

* Darker shade relates to destimulants.

Source: own elaboration based on the results of soft model.

Table 5. Estimates of INHC variable external model

Indicator code	Weight (*error*)	Loading (*error*)
INHC01	0,1733 (*0,0006*)	0,6394 (*0,0004*)
INHC02	0,2136 (*0,0002*)	0,8700 (*0,0001*)
INHC03	0,1259 (*0,0003*)	0,4535 (*0,0003*)
INHC04	0,2144 (*0,0002*)	0,8329 (*0,0001*)
INHC05	0,2686 (*0,0000*)	0,9339 (*0,0000*)
INHC06	0,2418 (*0,0001*)	0,8965 (*0,0000*)

Source: own elaboration based on the results of soft model.

The weights and loadings estimates with regard to INHC latent variable external model are presented in Table 5. The results are consistent with expectations. All stimulants received positive weights and loadings. Furthermore all parameters are statistically significant.

The indicator "expenditures on R&D per capita" (INHC05) has the strongest influence on the INHC variable. The indicator "public and private expenditures on health care per capita" (INHC03) is poorly correlated with the INHC variable. Hence, expenditures on the R&D sector are the most significant form of investments in human capital in Polish regions.

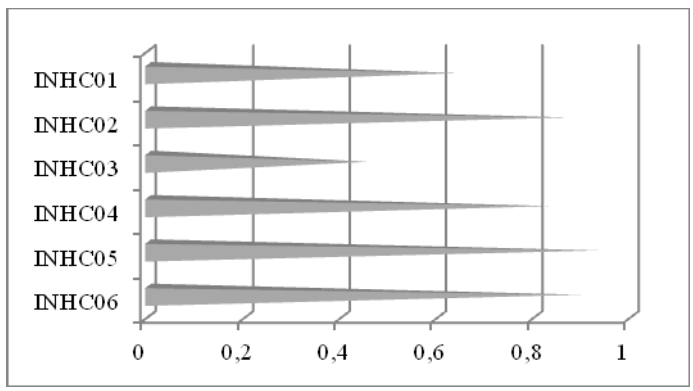

Figure 4. Loadings of INHC variable

Source: own elaboration based on the results of soft model.

Table 6 contains estimates of weights and loadings with regard to the LED latent variable external model. The results are consistent with expectations. Stimulants have positive weights and loadings, destimulants have negatives ones. All parameters are statistically significant ("2s" rule).

Most indicators reflect the LED variable strongly. The indicator "gross value added per employee" (LED02) has the strongest influence on the variable and indicator "percentage of employed in agriculture" (LED03) has the weakest influence on the variable LED.

Equations (3) and (4) present estimations of internal relations. Standard deviations calculated basing on the Tukey cut method are given in brackets.

$$\underset{}{HC_t} = \underset{(0,0661)}{0,6515} + \underset{(0,0003)}{0,9525}INHC_{t-1}, \tag{3}$$

$$\underset{}{LED_t} = \underset{(0,9770)}{1,0385} + \underset{(0,0552)}{0,1448}HC_t + \underset{(0,0318)}{0,7798}PC_t, \tag{4}$$

Table 6. Estimates of LED variable external model

Indicator code	Weight (*error*)	Loading (*error*)
LED01	0,1464 (*0,0587*)	0,9501 (*0,2381*)
LED02	0,1315 (*0,0734*)	0,9544 (*0,2409*)
LED03	–0,0592 (*0,0684*)	–0,6613 (*0,3028*)
LED04	0,0948 (*0,0098*)	0,7675 (*0,0679*)
LED05	0,1364 (*0,0527*)	0,9527 (*0,2359*)
LED06	0,1290 (*0,0747*)	0,9087 (*0,2640*)
LED07	0,1078 (*0,0618)*	0,8887 (*0,2234)*
LED08	0,0694 (*0,0463)*	0,7294 (*0,1941)*
LED09	0,1328 (*0,0356)*	0,7567 (*0,1972)*
LED10	0,1362 (*0,0773)*	0,9496 (*0,2956)*

Source: own elaboration based on the results of soft model.

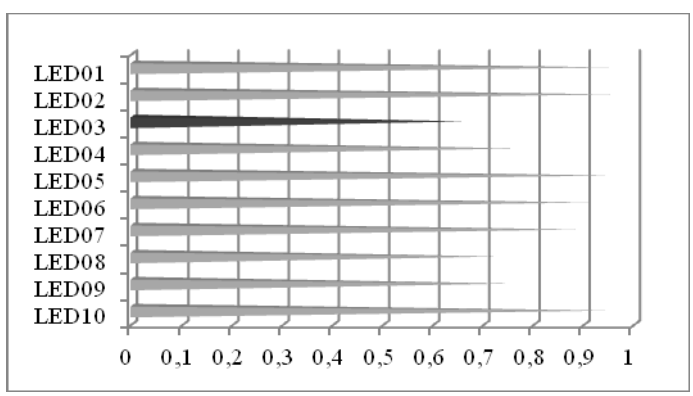

Figure 5. Loadings of LED variable (the absolute value)*

* Darker shade relates to destimulants.

Source: own elaboration based on the results of soft model.

Signs of estimations are consistent with expectations. The relationship between investments in human capital and the stock of human capital is positive. Moreover the correlations between human capital as well as physical capital and the level of economic development are positive. All parameters are statistically significant ("2s"

rule). Coefficient of determination (R^2) have value 0.9 for the equation (3) and value 0.81 for the equation (4). The general Stone-Geisser test is equal to 0.1527[6].

In conclusion, it is possible to claim that investments in human capital borne by regions in 2007 have a very strong and positive influence on the stock of human capital in 2008. The regions which invested more in human capital in 2007 had a higher stock of human capital in 2008.

Equation (4) points out the strong, positive correlation between physical capital and the level of economic development and the weak, positive correlation between human capital and the level of economic development. Hence, regions which had the higher stock of physical capital, had also the higher level of economic development in 2008 and regions which had the higher stock of human capital, also had the higher level of economic development in 2008. The significant conclusion is that the influence of physical capital on the economic development level was much stronger than the influence of human capital.

5. Analysis of latent variables synthetic measures

The Partial Least Square method used to soft model estimation provides calculations of latent variable values. These values can be treated as a synthetic measure and used for comparative analysis.

One of the most important advantages of soft modelling is that the construction of synthetic measure is based not only on latent variable definition but also on relationships among other categories within the model.

Figure 6 presents Polish regions ranked according to the stock of human capital in 2008. The regions were divided into four groups which were constructed based on parameters of synthetic measure (z_i): average $(\bar{z})$ and standard deviation (s_z) [Nowak 1990, pp. 92, 93]:

I group – very high stock of human capital: $z_i \geq \bar{z} + s_z$,

II group – high stock of human capital: $\bar{z} \leq z_i < \bar{z} + s_z$,

III group – low stock of human capital: $\bar{z} - s_z \leq z_i < \bar{z}$,

IV group – very low stock of human capital: $z_i < \bar{z} - s_z$.

The highest stock of human capital is concentrated in the mazowieckie voivodeship. All stimulants reached the highest values for this region. Małopolskie, dolnośląskie, pomorskie, wielkopolskie, zachodniopomorskie, łódzkie and lubelskie are placed in the second group. These regions have achieved high positions in terms of indicators such as: "students of post-graduate studies per 1 thousand population" (HC04), "percentage of employed using computer with Internet access at least once a week" (HC10). The above indicators are strongly correlated with the HC variable.

[6] Stone-Geisser test measures prognostic property of soft model. Positive (negative) value of this test indicates high (poor) quality of model.

Śląskie, podlaskie, kujawsko-pomorskie, opolskie, świętokrzyskie, warmińsko--mazurskie and podkarpackie make up the third group. Lubuskie is in the last position.

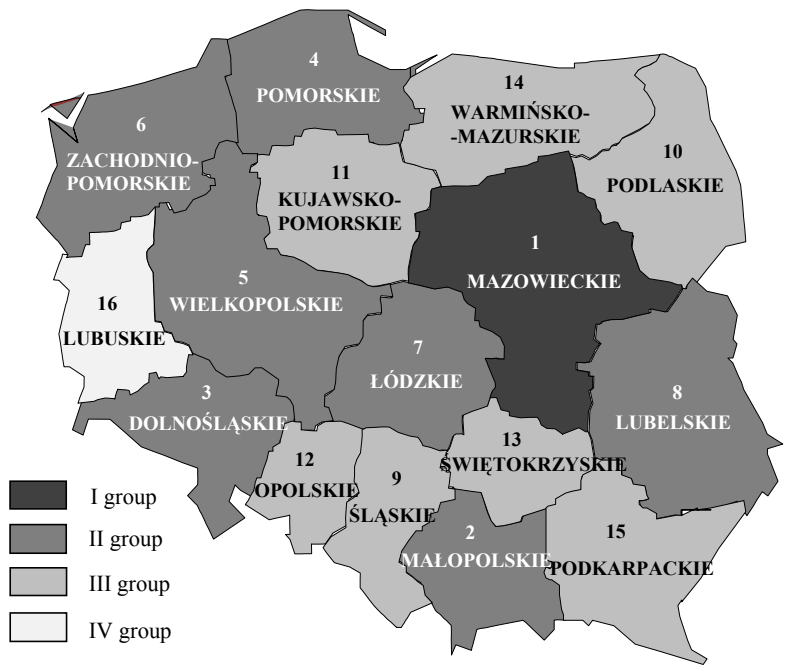

Figure 6. Polish regions ranking in terms of the human capital stock in 2008

Source: own elaboration based on the results of soft model.

The Polish regions ranked in terms of human capital investments in 2007 is shown in Figure 7. The regions were divided into three groups:

I group – very high investments in human capital,

II group – high investments in human capital,

III group – low investments in human capital.

Mazowieckie and dolnośląskie were the biggest investors in human capital in 2007. The indicators which had a strong correlation with the INHC variable achieved the highest value for these regions (INHC05, INHC06). The second group is composed of: małopolskie, pomorskie, łódzkie, wielkopolskie. The rest of the voivodeships was classified to the last group with low investments in human capital.

Figure 8 presents Polish regions ranked in terms of the level of economic development in 2008. The regions were divided into four groups:

I group – very high level of economic development,

II group – high level of economic development,

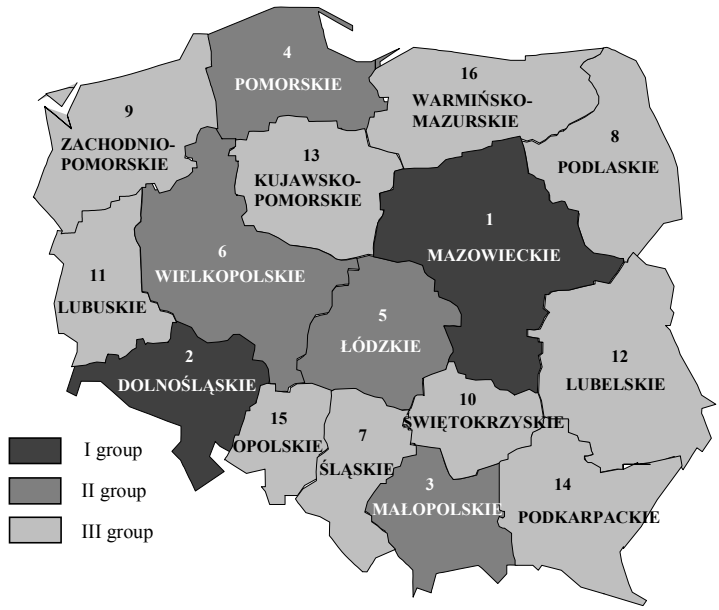

Figure 7. Polish regions ranking in terms of human capital investments in 2007

Source: own elaboration based on the results of soft model.

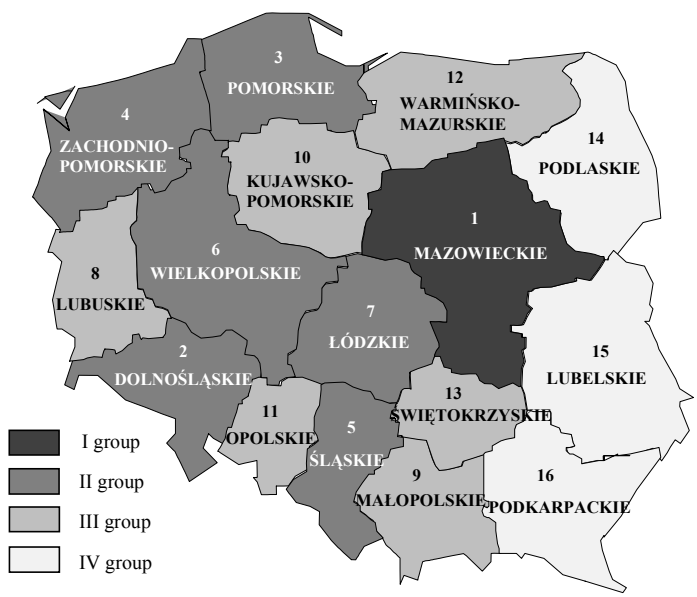

Figure 8. Polish regions ranking in terms of the level of economic development in 2008

Source: own elaboration based on the results of soft model.

III group – low level of economic development.

IV group – very low level of economic development.

The mazowieckie voivodeship is ranked in the first position. The high level of economic development is characteristic of dolnośląskie, pomorskie, zachodniopomorskie, śląskie, wielkopolskie, łódzkie. The regions of east Poland: podlaskie, lubelskie, podkarpackie had the lowest level of economic development in 2008.

6. Conclusions

The presented soft model has enabled us to investigate the relationships among human capital, investments in human capital, physical capital and the economic development level. Some conclusions and remarks can be formulated according to the results of this research:

- education and knowledge are the most significant components of human capital in Polish regions,
- expenditures on the R&D sector are the most significant form of investing in human capital in Polish regions,
- investments in human capital influence positively the human capital stock,
- human capital has a positive, statistically important influence on the level of economic development in Polish regions,
- physical capital is stronger correlated with the level of economic development than human capital,
- the highest stock of human capital is concentrated in the mazowieckie voivodeship, the lowest in lubuskie,
- mazowieckie and dolnośląskie regions are the biggest investors in human capital,
- mazowieckie voivodeship is characterized by the highest level of economic development, the podlaskie, lubelskie and podkarpackie regions have a very low level of economic development.

References

Cichy K., Malaga K. (2007), Kapitał ludzki w modelach i teorii wzrostu gospodarczego, [in:] M. Herbst (ed.), *Kapitał ludzki i kapitał społeczny a rozwój regionalny*, Wydawnictwo Naukowe „Scholar", Warszawa, pp. 20–49.

Dagum C. (2004), Human capital, [in:] *Encyclopedia of Statistical Sciences*, John Wiley & Sons, New York, pp. 1–12.

De la Fuente A., Domenech R. (2002), *Human Capital in Growth Regression: How Much Difference Does Data Quality Make?*, UFAE and IAE Working Papers 537.02.

Domański S.R. (1993), *Kapitał ludzki i wzrost gospodarczy*, Wydawnictwo Naukowe PWN, Warszawa.

Florczak W. (2007), *Kapitał ludzki a rozwój gospodarczy*, [in]: W. Welfe (ed.) *Gospodarka oparta na wiedzy*, PWE, Warszawa, pp. 126–166.

Le T., Gibson J., Oxley L. (2005), *Measures of Human Capital: A Review of the Literature*, Working Paper 05/10, New Zealand Treasury, Wellington.
Łukasiewicz G. (2009), *Kapitał ludzki organizacji. Pomiar i sprawozdawczość*, Wydawnictwo Naukowe PWN, Warszawa.
Marciniak S. (2000), *Innowacje i rozwój gospodarczy*, Kolegium Nauk Społecznych i Administracji Politechniki Warszawskiej, Warszawa.
Mierzyńska D. (1999), *Modele miękkie w analizie porównawczej złożonych zjawisk społeczno--ekonomicznych*, rozprawa doktorska, Uniwersytet w Białymstoku, Białystok.
Nowak E. (1990), *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*, PWE, Warszawa.
Perło D. (2004), *Źródła finansowania rozwoju regionalnego*, Wydawnictwo Wyższej Szkoły Ekonomicznej w Białymstoku, Białystok.
Próchniak M. (2006), Czynniki wzrostu gospodarczego – wnioski z badań empirycznych, *Ekonomista*, nr 3, pp. 320–323.
Rogowski J. (1990), *Modele miękkie. Teoria i zastosowanie w badaniach ekonomicznych*, Wydawnictwo Filii UW w Białymstoku, Białystok.
Wold H. (1980), *Soft Modelling: Intermediate between Traditional Model Building and Data Analysis*, Banach Centre Publication 6, Mathematical Statistics,Warsaw.

ZASTOSOWANIE MODELOWANIA MIĘKKIEGO DO POMIARU ZALEŻNOŚCI MIĘDZY KAPITAŁEM LUDZKIM A POZIOMEM ROZWOJU GOSPODARCZEGO W POLSKICH WOJEWÓDZTWACH

Streszczenie: Artykuł ma na celu przedstawienie wyników badań dotyczących zależności między kapitałem ludzkim a poziomem rozwoju gospodarczego polskich województw. Badania zostały przeprowadzone w oparciu o metodę modelowania miękkiego, która umożliwia badanie powiązań między zmiennymi nieobserwowalnymi bezpośrednio (tzw. zmienne ukryte). Do tego typu kategorii można zakwalifikować zarówno kapitał ludzki, jak i rozwój gospodarczy. Zaprezentowany w artykule model miękki zawierał cztery zmienne ukryte: kapitał ludzki, inwestycje w kapitał ludzki, kapitał rzeczowy oraz poziom rozwoju gospodarczego. Uzyskane rezultaty pozwoliły na zweryfikowanie, która z form kapitału: ludzki czy rzeczowy, siniej wpływała na poziom rozwoju gospodarczego polskich województw oraz na uporządkowanie województw pod względem zasobów kapitału ludzkiego, inwestycji w kapitał ludzki i poziomu rozwoju gospodarczego.

Słowa kluczowe: kapitał ludzki, rozwój gospodarczy, modelowanie miękkie.