Concept for Measuring the Efficiency of Public Goods Provision Based on the Education Sector in Poland[1]

Bazyli CZYŻEWSKI – Jan POLCYN** – Anna HNATYSZYN-DZIKOWSKA****

Abstract

The problem of quantifying public goods is one of the most complex problems related to public choice theory. We argue that the public goods constitute an isomorphic, socio-economic system which is not a "black box". The central goal of the article is to develop a universal methodology for measuring the quantity and quality of public goods, and the efficiency of their provision in different sectors of the economy. The authors have applied the developed methodology to a study of the process of public goods provision in the Polish education sector. The empirical research bears out the theory that the supply of public goods in the education sector is determined by the structure of the local budgetary funds used for that purpose, and not only by the total amount of public spending.

Keywords: *public goods; education sector; taxonomic analysis*

JEL Classification: H41, I25, I28, I29

Introduction

The problem of quantifying public goods (PGs), and in particular measuring the efficiency of their provision, is one of the most complex problems related to public choice theory. The term "public goods" is a generalisation. Economic theory distinguishes four types of goods: private, common, club, and public. The criteria of taxonomy include four traits: "rivalry", "non-rivalry", "excludability" and "non-excludability" (Klimowicz and Bokajało, 2012, p. 98). In a narrow

* Bazyli CZYŻEWSKI, Poznań University of Economics and Business, Al. Niepodległości 10, 61-875 Poznań, Poland; e-mail: b.czyzewski@ue.poznan.pl

** Jan POLCYN, Stanisław Staszic University of Applied Sciences in Piła, ul. Podchorążych 10, 64-920 Piła, Poland; e-mail: Jan.Polcyn@pwsz.pila.pl

*** Anna HNATYSZYN-DZIKOWSKA, Collegium Medicum at the Nicolaus Copernicus University in Toruń ul., Jagiellońska 13-15, 85-067 Bydgoszcz, Poland; e-mail: a.hnatyszyn@cm.umk.pl

[1] The article was financed as a part of a grant from the Polish National Science Centre's resources awarded on the basis of the decision No. DEC-2013/11/B/HS4/00572.

perspective, pure PGs are considered to meet two conditions: “non-rivalry” and “non-excludability” (Ulbrich, 2003, p. 67). In practice, however, such goods are scarce in the economy (examples include national services, national defence, order and security). In our considerations, we therefore extend the definition to include common goods (characterised by “rivalry” and “non-excludability”) and the so-called merit goods, which may be private goods in terms of their physical traits, but, as a result of social doctrine and the social policy implemented by public authorities, are provided to citizens even without their acceptance. They include most goods financed by the public sector, particularly in the field of education, in healthcare and, in accordance with the latest concepts, in agriculture.

The absolute value of a public good, depends on individual utility functions, and hence an objective determination of such a value is difficult, as well as approximation for individually experienced welfare. However, a vast literature exists, reflecting attempts at estimating utility functions for public goods. Essentially, three avenues have been pursued: revealed preference methods (i.e. the hedonic method and the defense expenditure approach), stated preference methods (e.g. the contingent valuation method) and the Life Satisfaction Approach (i.e. the method to value the psychic costs of public bads) (Levinson, 2012; Luechinger, 2009; Kahneman and Thaler, 2006; Gruber and Sendhil, 2005; Van Praag and Baarsma, 2005; Di Tella, MacCulloch and Oswald, 2001; Alesina, Di Tella and MacCulloch, 2004; Kahneman, Wakker and Sarin, 1997). A common point for those approaches is a need of microeconomic data revealing a demand for public goods which is always a very debatable issue. Thus, a relative value of a public good is not as much questionable as its absolute value. It is objectively possible to identify which public goods are more or less valuable but there is not a generally accepted methodology for doing this.

Usually, economists assumes that public spending should translate into the highest performance of public sector. The performance measures are perceived as outputs of the public spending. In this way the efficiency indicators for a public sector are calculated, using parametric or nonparametric analyses. In the authors opinion it is very simplified approach, because the public spending grants a package of complementary public goods which finally contributes to public sector performance but do not ensure its definitive quality. We argue that this package constitutes in each sector of economy an isomorphic, socio-economic system which is not a “black box”. There is a missing element in the frontiers analyses of public sector efficiency – a specification of a basket of PGs, and the research problem is to explore it. If we consider public spending on the one hand and the performance indicators on the other, we assume that a set of public goods with its attributes – a quantity, a sequence of provision and complementarities –

doesn't matter, but it does. The authors make attempts at filling this gap. A different approach to estimating efficiency of PGs provision has been adopted. It distinguishes three dimensions of this process: a public spending, a given quantity of public goods and the measures of public goods quality.

In order to specify the level of possible shortage of PGs, their available quantity and quality should first be measured, which at the same time raises the question of the efficiency of provision of these goods. A commonly accepted research methodology in this field has unfortunately not yet been developed, and there do not exist any universal methods for the quantification and valuation of public goods.

Therefore, the central goal of this article is to develop a universal methodology filling aforementioned gap, for measuring both the quantity and quality of PGs, and the efficiency of their provision in different sectors of the economy, which at the next stage would enable the performance of taxonomic analyses and identification of possible ways to increase the efficiency. The subject matter of the article, however, is not merely methodological. The authors have applied the developed methodology to a study of the process of PGs provision in the Polish education sector, choosing the North-Western Region in Poland, according to NTS – 1, and a representative sample of its poviats as a case study. This case study tests the following hypothesis**:** PGs supply in different sectors in Poland is determined by the structure of local allocation of expenditure, and not only by the total amount of public spending.

Motivation for the Research Problem

It is obvious that "no decentralised price system will allow the determination of an optimum level of public goods – a solution exists, but the problem is how to find it" (Samuelson, 1967). Thus a market system does not automatically lead to optimum allocation of the goods, either in theory or in practice, as is the case with private goods. In Pareto's optimality theory, the marginal rate of PG substitution with private goods is lower than in the case of an individual optimum; thus each person consumes more public than private goods. Although the difference in quantity depends on the shape of individual utility functions, voluntary market exchange will always lead to PG shortage as compared with a socially optimum level (Osiatyński, 2006, p. 55). It is commonly known that individuals have no incentive to disclose their true demand for non-excludable goods (Frey, Luechinger and Stutzer, 2009). For that reason public goods must be funded by state but there is still the question in which quantities?

The most popular approach to measure the public sector efficiency is an input-oriented DEA model. In this model the inputs (i.e. public spending) are minimized and the outputs are held at their current levels (e.g. Afonso, Schuknecht and Tanzi, 2005; Afonso and Aubyn, 2005; Antonis and Manthos, 2011). The DEA nonparametric method allows only for grouping analyzed units by efficiency based on the estimated value in the range from 0 to 1. It has the fundamental disadvantage: data envelopment analysis (DEA) determines the most effective units (value 1) even from the set of non-effective units (i.e. the highest level of analyzed efficiency of the non-effective unit collection). There is the assumption that effective units are the best ones in the examined set and thus the determined efficiency in the examined set is relative and can be bias.

There are different performance indicators engaged as the outputs. In order to capture any qualitative differences among the educational systems, Hanushek and Kimko (2000) have constructed the public goods' quality indicator. Afonso, Schuknecht and Tanzi (2005) has proposed a set of composite indicators of public sector performance defined as the outcome in relation to the resources employed. Most studies conclude that public spending could be much smaller and, assuming the output remains constant, more efficient than today (e.g. Joumard et al., 2004; Tanzi and Schuknecht, 1997). These conclusions could be bias to some extent. The cited authors have assumed that chosen performance indicators are (or should be) a function of public spending. In fact, the public funds provide, as we said before, a package of goods and services which determines the output. Thus, both quantity and quality of public goods should be analyzed to assess the public sector efficiency. The quantity of PG is a very sensitive variable since it correlates directly to the life satisfaction of a society much more that the performance indicators do. Voters do not care as much about the educational value added (EVA) indicator, as about present schools and teachers availability. Thus, policymakers consider first of all a broadly understood quantity of public goods which can be delivered than its overall performance. For that reason, the synthetic measures of PG's quantity should be well examined, not omitted, in the public sector efficiency analyses. However, research so far carried out to evaluate the efficiency of education systems has been also based on the DEA method, e.g. the efficiency of universities was examined by Nazarko, and Šaparauskas (2014) and secondary schools Aristovnik and Obadić (2014).

Education Sector as a Provider of Public Goods and Classification of their Measuresa

As has been stated above, PGs in Poland are provided chiefly by three sectors of the economy: education, healthcare and agriculture (including rural areas, which offer natural resources). The education sector as a PGs provider is described below.

In the healthcare sector, most medical goods and services are not considered purely as PGs, since there occur rivalry in their consumption and the possibility of exclusion from consumption (Czyżewski, Hnatyszyn, Polcyn, 2016; Mucha, 2006, p. 11; Hsiao, 1995, pp. 127 – 128; Laskowska, 2012). The third area of PG supply (agriculture) is related to the natural environment and its resources. In this case, agriculture and rural areas are key sectors generating PGs (Baldock, Hart and Scheele, 2014).

In the education sector, a group of PGs defined as merit goods (also called "social goods") has been identified. These goods share the traits of private goods but, mainly because of national social policy, they are available to every citizen, and their financing is based on funding from the national (and also local government) budget (Shaw, 2010). With regard to the above, initial quantitative and qualitative classification of the goods may be carried out. Quantitative measures include such variables as number and profile of education institutions, number of places offered at particular stages of education, equipment of educational institutions, and number of teachers according to career stage. The following measures are suggested for evaluating the quality of PGs: number of graduates at particular stages of education, secondary school-leaving examination pass rate, number of students obtaining a matriculation certificate, and educational value added (EVA). EVA is defined as improvement in students' knowledge as a result of a specific educational process. It measures students' progress over a specific research period (Ballou, Sanders, Wright, 2004). In Polish conditions, EVA is measured in a modified form (Dolata, 2007, p. 9). PGs in the education sector are financed partly by a mechanism of general subsidy (redistribution of funds from the central budget via the Ministry of National Education) and local government budgets. Research into the efficiency of the functioning of education models with reference to the size of a school, method of organisation and volume of expenditure has been carried out by, among others, Deller and Rudnicki (1993). That research, however, was of a different nature: above all, they did not analyse the context of PGs sufficiently explicitly.

General Concept for the Quantification of Public Goods and Efficiency of their Provision

The authors have developed the following research procedure, serving to analyse the process of PG provision in different sectors of the economy (for the healthcare sector see Czyżewski, Hnatyszyn and Polcyn, 2016):

1. Estimate synthetic measures of PG quantity in a specific sector on the basis of the phenomena described in the previous section (e.g. Hellwig's measure). In this case there is a matrix in a local arrangement based on the division into poviats.

2. Estimate synthetic measures of PG quality in accordance with the above remarks.

3. Normalise the synthetic measures, e.g. with the use of the zero unitarisation method (ZUM), for comparison purposes, with non-negative values of normed features retained.

4. Estimate a measure giving PG provision efficiency in terms of relations of normed synthetic quantity- and quality-related measures from the steps 1 and 2.

5. Identify the structure of PG financing in a specific sector – matrix of structure measures in a local arrangement.

6. Cluster analysis of territorial units according to the PG financing structure criterion, with the aim of identifying similar models of institutional valuation of PGs.

7. Estimate descriptive statistics: among others, mean values of quantity, quality and efficiency measures (from stages 1, 2 and 4) in obtained clusters (classes), assuming that these clusters (classes) are institutional predictors of the PG provision process.

8. Carry out a multifactorial ANOVA/MANOVA. In these analyses, the classes (clusters) from stage 6 are a qualitative predictor, while the measures from stages 1, 2 and 4 are dependent variables. The aim of the analyses is, firstly, to determine statistically significant relations between the goods-financing structure and PG quantity, quality and provision efficiency, and, secondly, to answer the question to what degree individual financing models are responsible for the variation in the measures describing the process of PG generation (contrast analysis)?

9. Identify the optimum PG financing models in a local or regional arrangement from the perspective of quantitative, qualitative or efficiency-related criteria.

As every method, the proposed approach has its weaknesses. Determining synthetic measures (steps 1 and 2) is the most problematic. Since variables expressed in different units have to be added up somehow (in order to calculate then the efficiency measure), the problem of attributing weights appears. In the carried out case study we assumed that weights of variables that create synthetic measures of public goods quantity and quality are the same, however, here the researcher must be extremely careful. A solution would be to determine linear functions for every type of effect (where a given effect would be a function of respective public goods quantity) and in this way to determine average weights based on regression coefficients of all the functions. In our case study it would mean evaluation of 25 functions and that considerably would increase the workload of the analysis. The procedure was piloted in a study to evaluate the process of PGs creation in the education sector in the North-Western Region in Poland, and its results are described in the sections below.

Methodology of the Case Study

The study was carried out according to the algorithm described above, on the basis of data concerning secondary education facilities (general secondary and vocational) for the representative sample of 31 poviats (counties) of the North-Western Region in Poland (covering 3 of 16 provinces: zachodniopomorskie, wielkopolskie, lubuskie). Data obtained from the Regional Examinations Board in Poznań, the Regional Chamber of Accounts in Poznań and the Ministry of National Education were analysed.

Three groups of variables described in Tables 1, 2 and 3 were applied. Synthetic measure of public goods quantity (tab. 1) shows what public funds were allocated for in education and what the scale of financial activity was. The side of effects in the carried out analyses (Table 2) is represented by the education quality synthetic measure. The measure is based on the educational value added (EVA) method and reflects the input of education into creating human capital. It is possible to determine EVA if at least two results of educational achievements measurements are available, i.e. at the beginning of education in a given school and at the end of it. A score that student may achieve at the end of an educational period is predicted based upon the initial score they achieve. The educational added value is the difference between the estimated examination score and the actual score a student achieved. The EVA may have a positive value if the score is higher than the predicted one, or negative, when the score value is lower than the predicted score (Ballou, 2005; Ballou, Hart and Scheele, 2004; Jakubowski, 2008). Regional models of funding education are presented in tab. 3 as scores of cluster analysis. According to the assumptions these models (or indicated clusters) are the qualitative (institutional) determinants of quantity, quality and efficiency of supplying public goods.

Data for analyses was obtained from governmental agencies. It can be assumed that it is very reliable. Original data was just raw data created based on exam scores and information about the number of students was obtained from financial reporting systems of the Ministry of Education. Regional Examination Boards operating in Poland execute standardized exams for the whole youth population that undergoes the examination process. Educational added value which is a significant part of these analyses is defined for the whole population by the Educational Research Institute of the Ministry of Education which guarantees appropriate quality of the analyzed data. The study covered the year 2013 for the synthetic PGs measures and the averages of 2007 – 2013 if considering a structure of the public spending, as the qualitative predictor. As a result, a set of variables allowing the definition of a synthetic measure of PG quantity (Table 1), a set of variables allowing the definition of a synthetic measure of quality (Table 2) and a set of variables reflecting the structure of expenditure (Table 3) were obtained.

T a b l e 1

Diagnostic Variables Allowing the Definition of a Synthetic Measure of PG Quantity

Symbol	Diagnostic variable	Type
Synthetic measure of PG quantity		
X_1	Number of secondary school students	stimulus
X_2	Number of technical and vocational school students	stimulus
X_3	Number of probationary teachers counted in full-time equivalents	stimulus
X_4	Number of contract teachers counted in full-time equivalents	stimulus
X_5	Number of nominated teachers counted in full-time equivalents	stimulus
X_6	Number of certified teachers counted in full-time equivalents	stimulus

Source: Own study based on data from the Regional Examinations Board in Poznań, the Regional Chamber of Accounts in Poznań and the Polish Ministry of National Education.

T a b l e 2

Diagnostic Variables Allowing the Definition of a Synthetic Measure of PG Quality

Symbol	Diagnostic variable	Type
Synthetic measure of effects		
General secondary schools		
X_1	Secondary school-leaving examination pass rate	stimulus
X_2	Number of students obtaining a matriculation certificate	stimulus
X_3	Educational value added by the humanities group	stimulus
X_4	Number of schools with positive EVA measure for the humanities	stimulus
X_5	Number of schools with negative EVA measure for the humanities	inhibitor
X_6	Educational value added by Polish language	stimulus
X_7	Number of schools with positive EVA measure for Polish language	stimulus
X_8	Number of schools with negative EVA measure for Polish language	inhibitor
X_9	Educational value added by the mathematics and natural science group	stimulus
X_{10}	Number of schools with positive EVA measure for mathematics and natural science	stimulus
X_{11}	Number of schools with negative EVA measure for mathematics and natural science	inhibitor
X_{12}	Educational value added by mathematics	stimulus
X_{13}	Number of schools with positive EVA measure for mathematics	stimulus
X_{14}	Number of schools with negative EVA measure for mathematics	inhibitor
Technical schools		
X_{15}	Secondary school-leaving examination pass rate	stimulus
X_{16}	Number of students obtaining a matriculation certificate	stimulus
X_{17}	Educational value added by the humanities group	stimulus
X_{18}	Number of schools with positive EVA measure for the humanities	stimulus
X_{19}	Number of schools with negative EVA measure for the humanities	inhibitor
X_{20}	Educational value added by Polish language	stimulus
X_{21}	Number of schools with positive EVA measure for Polish language	stimulus
X_{22}	Number of schools with negative EVA measure for Polish language	inhibitor
X_{23}	Educational value added by the mathematics and natural science group	stimulus
X_{24}	Number of schools with positive EVA measure for mathematics and natural science	stimulus
X_{25}	Number of schools with negative EVA measure for mathematics and natural science	inhibitor
X_{26}	Educational value added by mathematics	stimulus
X_{27}	Number of schools with positive EVA measure for mathematics	stimulus
X_{28}	Number of schools with negative EVA measure for mathematics	inhibitor

Source: Own study based on data from the Regional Examinations Board in Poznań, the Regional Chamber of Accounts in Poznań and the Polish Ministry of National Education.

Synthetic measures of PG quality and quantity were determined by Hellwig's method, according to the following procedure:

1. *Determining coefficient of variation of studied traits*

Initial analysis of empirical data included determining the coefficient of variation for each jth variable. The coefficient is a relative measure of dispersion and it allows the elimination of quasi-steady variables. The coefficient was calculated using formula (1) (Borkowski, Dudek and Szczęsny, 2003):

$$V_j = \frac{S_j}{\overline{x}_j} \tag{1}$$

where

V_j – the coefficient of variation for the *j*-th variable;

$S_{(xj)}$ – the standard deviation for the *j*-th variable, determined according to formula (2):

$$S_j = \sqrt{n^{n-1}\sum_{i=1}^{n}(x_{ij} - \overline{x})^2} \tag{2}$$

$\overline{x}_i$ – the arithmetic mean for the *j*-th variable, determined according to formula (3):

$$\overline{x}_j = n^{-1}\sum_{i=1}^{n} x_{ij}, \ (i = 1, \ ..., \ n) \tag{3}$$

Based on the result, the dispersion force is usually evaluated in the following manner:

0 – 0.20 – variation, permanent traits;

0.21 – 0.40 – variation, moderate traits;

0.41 – 0.60 – variation, strong traits;

0.61 and above – variation, very strong traits.

Traits satisfying the inequality, $\left|V_j\right| \le V^*$ where denotes the V^* critical value of the coefficient of variation, are eliminated from the set of analysed variables. V^* = 0.10 was taken as the critical value for the analysed set of variables.

2. *Standardisation according to formula* (4) (Gatnar and Walesiak, 2004; Walesiak, 2003):

$$t_{ij} = \frac{x_{ij} - \overline{x}_j}{S_j} \tag{4}$$

where

t_{ij} – the standardised value of the *j*-th trait in the *i*-th poviat;

x_{ij} – the empirical values in the *i*-th poviat;

$\overline{x}$ – the arithmetical mean of the *j*-th trait;

S_j – the standard deviation of the *j*-th trait.

3. *Division of studied traits into stimuli and inhibitors*

The gross scholarisation coefficient and number of students were qualified as stimuli. Suggested variables which are qualified as inhibitors can be transformed into stimuli using formula (5):

$$x_{ij} = \frac{1}{x_{ij}} \tag{5}$$

However, the lowest value of an inhibitor was taken when determining the pattern object, ignoring the process of transforming inhibitors into stimuli.

4. *Determining the development pattern*

The standardised matrix of variables is a basis for determining the so-called development pattern, i.e. an abstract object (poviat) P_0 with standardised coordinates z_{01}, z_{02}, …, z_{0j}, where z_{0j} = max $\{z_{ij}\}$ when z_j is a stimulus, and z_{0j} = min $\{z_{ij}\}$ when z_j is an inhibitor. This pattern is seen to represent a hypothetical poviat with the best observed variable values.

5. *Calculating Hellwig's synthetic measure*

At the next stage, the distance from the pattern was determined for each object P_i (poviat) according to formula (6):

$$d_i = 1 - \frac{D_{i0}}{D_0} (i = 1,\ 2,\ ...,\ n) \tag{6}$$

where

D_{i0} – the distance of the i-th object from P_0

$$D_{i0} = \sqrt{\sum_{j=1}^{m} (z_{ij} - z_{oj})^2} \tag{7}$$

$$\bar{D}_0 = n^{-1} \sum_{i=1}^{n} D_{i0} \tag{8}$$

$$S_0 = \sqrt{n^{-1} \sum_{i=1}^{n} (D_{i0} - \bar{D}_0)^2} \tag{9}$$

$$D_0 = \bar{D}_0 + 2S_0 \tag{10}$$

6. *Normalising Hellwig's synthetic measure*

The results obtained as a result of the calculations in stage 5 (Hellwig's synthetic measure) were subjected to normalisation using the (ZUM) given by formula (11):

$$z_{ij} = \frac{x_{ij} - \min_{x_{ij}}}{\max_{x_{ij}} - \min_{x_{ij}}} \tag{11}$$

7. *Poviat classification according to Hellwig's normalised taxonomic measure*

Classification of poviats according to Hellwig's normalised taxonomic measure was carried out with the use of the arithmetic mean of Hellwig's measure and the standard deviation of that measure. Based on the above-mentioned quantities, the following classes of poviats were obtained (Pomianek, 2010, p. 233):

$$\text{class A: } d_i > \bar{d}_i + s_{d_i}$$

$$\text{class B: } \bar{d}_i - s_{d_i} < d_i \le \bar{d}_i + s_{d_i}$$

$$\text{class C: } d_i \le \bar{d}_i - s_{d_i}$$

where

d_i – the value of a synthetic measure calculated by Hellwig's development pattern method,

$\bar{d}_i$ – the arithmetic mean of the synthetic measure d_i,

s_{d_i} – the standard deviation of the synthetic measure d_i.

At the next stage, we verify the hypothesis that the structure of expenditure (budgetary subsidies) on a specific public good is the qualitative predictor which determines the quantity/quality of provided PGs and the efficiency of the process, taking account of the fact that the absolute volume of expenditure is limited. To this end, a cluster analysis using Ward's method was carried out, which enabled the identification of clusters of poviats sharing similar traits.

A multifactorial ANOVA was carried out, where the structure of expenditure on education was taken as a qualitative predictor, and measures of PG quantity, quality and provision efficiency were dependent variables. Multidimensional Wilks', Pillai's, Hotelling's and Roy's significance tests were applied to accept or reject the zero hypothesis asserting the equality of vectors of mean measures related to the quantity, quality and efficiency of PG provision, as opposed to the alternative hypothesis that they differ significantly (which would bear out the hypothesis put forward in the introduction). Tests verifying the fulfilment of the assumptions of variance analysis were performed, such as Box's M test to check the assumption of the homogeneity of covariances in multidimensional space, and Hartley's, Cochran's and Barlett's variance homogeneity tests. The next stage involved carrying out a so-called post-hoc analysis, i.e. Tukey's HSD tests (for unequal N) for significant dependent variables (according to the unidimensional results), allowing evaluation of which classes of a qualitative predictor significantly determine the variation of the studied variables. Contrasts for the predictor classes identified in the post-hoc analysis were then calculated, to answer the question to what degree the contrast coefficients enable prediction of the means of the groups. In other words, it evaluated what part of the variation (total variation of means for a given variable in all classes) may be assigned to a specific contrast. The sum of squares (SS), i.e. the variation for which a contrast is responsible, was calculated according to formula (12) and divided by the SS for a specific dependent variable in all predictor classes (Stanisz, 2007, p. 367):

$$SS_L = \frac{\overline{L}^2}{\frac{1}{n}\sum_{i=1}^{k} c_i^2} \tag{12}$$

where

$\overline{L}$ – the contrast evaluation value;

n – the number of replications (measures in a group);

c_i – the weights describing the contrast.

The contrast evaluation values were determined using formula (13):

$$\overline{L} = \sum_{i=1}^{k} c_i \overline{x}_i \text{ , where } \overline{x}_1, \dots \overline{x}_k \tag{13}$$

At the last stage of analysis, the so-called ω-measure was calculated, estimating the variance of a dependent variable explained by an independent variable in the entire population for the quantity of PG (as significant dependent variables). The ω-measure was determined according to formula (14) (Stanisz, 2007, p. 367ff):

$$\omega = \frac{(SS_{effect} - p.MS_{error})}{SS_{effect} + SS_{error} + MS_{error}} \tag{14}$$

where

$SS\,effect$ – an intergroup SS of differences between mean values of variables for particular predictor classes and their global mean (measure of total variation of means);

p – the number of degrees of freedom of a qualitative predictor;

SS_{error} – a measure of incidental variation, i.e. the SS of differences between the result of observation and the mean of a class;

MS_{error} – the mean square error.

The ω-measure enables evaluation of what percentage of the variation of individual dependent variables (quantity of PGs and efficiency of PG provision) in the entire population can be attributed to the qualitative predictor, in this case the structure of expenditure on education.

Results

The cluster analysis carried out using Ward's method, taking into account the synthetic measure of PG quantity, the synthetic measure of PG quality and the efficiency of PG provision, led to the identification of three clusters among the studied poviats. Basic descriptive statistics for models of education financing, determined on the basis of the cluster analysis, are given in Table 3.

Table 3

Characteristics of Clusters of Similar Poviats According to the Criterion of Expenditure on Education (mean values of traits)

No.	Class	Structure of expenditure on education*						
		Percentage						General subsidy measure
		X_1	X_2	X_3	$X_1 + X_2 + X_3$	X_4	X_5	X_6
1.	A	**1.49↑**	8.01	21.39	**30.89↑**	**23.86↓**	45.25	97.18
2.	B	**1.02↓**	8.82	17.95	27.79	27.13	45.08	**105.10↑**
3.	C	1.38	7.07	14.70	**23.15↓**	**31.07↑**	45.77	100.18
4.	*Total*	*1.38*	*7.69*	*17.93*	*27.00*	*27.54*	*45.47*	*99.55*

* X_1 – expenditure on remuneration of probationary teachers per student; X_2 – expenditure on remuneration of contract teachers per student; X_3 – expenditure on remuneration of nominated teachers per student; X_4 – expenditure on remuneration of certified teachers per student; X_5 – remainder of the subsidy calculated per student; X_6 – correction coefficient used for calculation of the general subsidy.

Source: Own study using the *Statistica* package, based on source data as in Table 1.

As a result of the analysis, three clusters of poviats were identified. Each cluster is characterised by the following descriptive statistics related to the PG provision process (Table 4).

Table 4

Descriptive Statistics Related to the PG Provision Process

Factor level	N	Synthetic measure of PG quantity			Synthetic measure of PG quality			PG provision efficiency ratio		
		Mean	Standard deviation	Standard error	Mean	Standard deviation	Standard error	Mean	Standard deviation	Standard error
A	13	0.3456	0.1584	0.0439	**0.6312↑**	0.2236	0.0620	2.1724	1.1073	0.3071
B	4	**0.1315↓**	0.1242	0.0621	**0.3219↓**	0.3319	0.1659	**5.2060↑**	4.7014	2.3507
C	14	**0.4598↑**	0.2637	0.0704	0.5597	0.2712	0.0724	**1.5726↓**	0.9115	0.2436

Source: Own study using the *Statistica* package, based on source data as in Table 1.

In all, 13 poviats, with 46 technical schools and 41 general secondary schools, were assigned to ***class A.*** Here, 26 of the technical schools obtained a satisfactory educational result in the humanities as measured by EVA (56.52%), 25 technical schools obtained a positive EVA for Polish language (54.35%), 18 obtained a good result in mathematics and natural sciences (39.13%), and 16 technical schools were positively evaluated for mathematics (34.78%). General secondary schools in the analysed poviats produced positive EVA values as follows: for humanities, 19 schools (46.34%); for Polish language, 23 schools (56.10%); for mathematics and natural sciences, 19 schools (46.34%); and for mathematics, 16 schools (39.02%). The value of the Di correction coefficient for each poviat was lower than 100, which means that each of them received less than 100% of the general educational subsidy, which takes into account such elements as number of students and teacher employment structure. At the same time, class A has the highest synthetic measure of PG quality (see Table 4) and the second highest measures of PG quantity and PG provision efficiency.

The poviats assigned to ***class B*** had 9 technical schools, three of which (33.33%) obtained positive EVA values. In none of the analysed poviats was there recorded a positive EVA value in the humanities group and separately in the Polish language group. The situation was even worse in general secondary education, where only one of eight schools obtained a positive EVA value (12.5%). Among general secondary schools in these poviats, no positive values were obtained in any of the analysed groups of subjects, i.e. humanities, Polish language, mathematics and natural sciences, and mathematics. The value of the D_i correction coefficient for each analysed poviat was greater than 1, indicating a clear difference from the poviats in class A. Class B is characterised by the lowest number of probationary teachers and a relatively high proportion of certified teachers (X4), i.e. those at the last stage of their professional career, much higher than the proportion in class A. This provides evidence of the low rotation of teachers under this model of financing.

Class C consists of 14 poviats, where secondary education is provided by 69 technical schools and 53 general secondary schools. Technical schools obtained positive EVA values as follows: humanities, 36 schools (52.17%); Polish language, 37 schools (53.62%); mathematics and natural sciences, 24 schools (34.78%); and mathematics, 23 schools (33.33%). General secondary schools had positive EVA values as follows: humanities, 16 schools (44.44%); Polish language, 19 schools (52.78%); mathematics and natural sciences, 20 schools (55.56%); mathematics, 18 schools (50.00%). The correction coefficient (X_6) was lower than 1 in exactly 50% of the poviats, and higher than 1 in the other 50% (7 poviats). Class C is characterised by the highest measure of PG quantity (see Table 4) and the highest share of so-called non-developmental teacher posts (X4). This is reflected in a slightly lower quality of PGs than in class A, and the lowest efficiency of PG provision (see Table 4).

Multidimensional significance tests lead to the rejection of the zero hypothesis of the equality of vectors of mean measures related to PG quantity, quality and efficiency, in favour of the alternative hypothesis that they differ significantly, which provides a ground for confirmation of the hypothesis put forward in the introduction and the correctness of the above considerations (see Table 5).

Table 5

Multidimensional Significance Tests

	Test	Value	F	Df effect	Df error	p
Classes for the structure of expenditure as a qualitative predictor	Wilks	0.512467	3.43985	6	52	**0.006175**
	Pillai	0.517944	3.14529	6	54	**0.010199**
	Hotelling	0.892003	3.71668	6	50	**0.003925**
	Roy	0.819599	7.37639	3	27	**0.000920**

Source: Own study using the *Statistica* package, based on source data as in Table 1.

The assumption of the homogeneity of covariance in multidimensional space raises certain doubts – Box's M test points to grounds for rejecting such 0H in favour of the hypothesis that covariances are not homogenous. On the other hand, Hartley's Cochran's and Bartlett's tests of homogeneity of covariance confirmed the homogeneity.

Unidimensional results prove the significance of variation in two variables: PG quantity and efficiency of PG provision (Table 6).

Table 6

Unidimensional Results for Dependent Variables

	Degrees of freedom	**Synthetic measure of PG quantity**				**PG provision efficiency ratio**			
		SS	MS	F	p	SS	MS	F	p
Structure of expenditure on education	2	0.348262	0.17413	3.8950	**0.03217**	41.397	20.699	6.311	**0.0054**
Error	28	1.251773	0.04470	–	–	91.827	3.279	–	–
Total	30	1.600035	–	–	–	133.225	–	–	–

Source: Own study using the *Statistica* package, based on source data as in Table 1.

Post-hoc tests (see Tables 7 and 8) proved the significance of the change in the education financing structure from model B to C in terms of the quantity of PGs. The suggested change in financing from model B to C is justified by the fact that in model B, average non-financial outlays on education per student were four times higher. Despite the considerably higher expenditure, a much lower quantity of PGs was obtained as compared with class C. The data presented show that an increase in non-financial outlays per student did not ensure either higher or even comparable effects with respect to class C. It should be concluded that a higher quality of education might be obtained in class B by changing the structure of financing to allocate a larger part of the funds to cheaper posts filled by probationary, contract and nominated teachers.

T a b l e 7

HSD Test (for unequal N); "Synthetic Measure of PG Quantity" Variable; Approximate Probability for Post-hoc Test; Error: Intergroup MS = 0.04471, df = 28.000

	Classes for structure of expenditure as a qualitative predictor	A	B	C
1	A		0.197752	0.353310
2	B	0.197752		**0.027840**
3	C	0.353310	**0.027840**	

Source: Own study using the *Statistica* package, based on source data as in Table 1.

T a b l e 8

HSD Test (for unequal N); "PG Provision Efficiency Measure" Variable; Approximate Probability for Post-hoc Test; Error: Intergroup MS = 3.2796, df = 28.000

	Classes for structure of expenditure as a qualitative predictor	A	B	C
1	A		**0.017863**	0.669558
2	B	**0.017863**		**0.004040**
3	C	0.669558	**0.004040**	

Source: Own study using the *Statistica* package, based on source data as in Table 1.

T a b l e 9

Evaluation of Contrasts for the PG Quantity Synthetic Measure

	Synthetic measure of PG quantity	**Synthetic measure of PG quantity**	**Synthetic measure of PG quantity**	**Synthetic measure of PG quantity**	**–95.00%**	**+95.00%**
	Evaluation	Standard error	t	p	Confidence limits	Confidence limits
CONTR.2 (B vs C, i.e. 0; 1; –1)	–0.328344	0.119874	–2.73907	**0.010596**	–0.5738	–0.0827
$SS_{contrast}/SS_{effect}$	**0.62**					

Source: Own study using the *Statistica* package, based on source data as in Table 1.

T a b l e 10

Evaluation of Contrasts for the PG Provision Efficiency Measure

	Synthetic measure of PG quantity	**Synthetic measure of PG quantity**	**Synthetic measure of PG quantity**	**Synthetic measure of PG quantity**	**–95.00%**	**+95.00%**
	Evaluation	Standard error	t	p	Confidence limits	Confidence limits
CONTR.1 (B vs A, i.e. 1; –1; 0)	–3.03361	1.035455	–2.92974	**0.006679**	–5.15465	–0.9125
$SS_{contrast}/SS_{effect}$	**0.44**					
CONTR.2 (B vs C, i.e. 0; 1; –1)	3.63344	1.026716	3.53890	**0.001425**	1.5303	5.7365
$SS_{contrast}/SS_{effect}$	**0.65**					

Source: Own study using the *Statistica* package, based on source data as in Table 1.

In the case of the efficiency of the PG provision process, changing the model of financing from B to A or from B to C is of great significance. The analysis of contrasts carried out for statistically significant dependent variables shows the relative weight of the discussed change in the PG financing structure (see Tables 9 and 10).

This shows that the change in the financing structure from model B to C explains 62% of the growth in the synthetic measure of PG quantity. On the other hand, in terms of the efficiency of the PG provision process, a change from model C to A rather than B seems a better solution, since the resulting drop in efficiency will be less severe.

Finally, reference should be made to the ω coefficient. In the case of PG provision efficiency, ω = 0.25, and in the case of PG quantity, ω = 0.16. This means that the structure of expenditure only explains respectively 25% and 16% of the variation in the synthetic measures of quantity and efficiency in the studied population. It leads to the conclusion that other variables (analyses within poviats), which determine the process of PG creation in the education sector, should also be considered. These variables certainly include demographic processes and the broadly understood social and economic development of regions, but this is a subject for further analysis.

The Contribution of the Proposed Method and Its Practical Implications

The applied method allow to observe the underlying reasons for differences in the performance of different financing models of education. Since both the quantity and quality of PGs are disclosed, we can point out why in some cases bigger funds do not translate into higher EVA. There is not such possibility in the ordinary DEA approach where we can only asses a relation of public funds to education effects and rank it. One has any idea, why sometimes bigger money doesn't result in EVA improvement?

Discussing the best performance of the ***model A***, it may be considered whether this is the result of its having the highest share of so-called "developmental teacher posts" – see the column "$X_1 + X_2 + X_3$" in Table 3. It is worth noting that the salary of certified teachers (X_4) is, in practice, the highest of all salaries available to teachers on their career path; thus teachers in this group may lack the motivation for self-improvement. Teachers in groups X_1, X_2 and X_3, on the other hand, are better motivated for self-improvement, because they are seeking promotion. Staffing issues have also been identified as a source of inefficiency in Slovenia and Croatia. Analyses of teaching efficiency in secondary schools in those countries show that there is a problem of excessive employment of teachers.

The main conclusions from those studies, however, suggest lowering the number of teachers employed by retiring those teachers who meet the criteria, and employing a proportionally lower number of new teachers, who will be subject to a strict selection process. Significant reasons for the inefficiency of education systems in Slovenia and Croatia include poor education results (in the last two quartiles among 31 studied OECD countries) and relatively high costs of education, especially in Slovenia. In the case of Slovenia, higher costs contribute to improved education results, but in proportions which deviate considerably from those expected (Aristovnik and Obadic, 2014).

In ***cluster B***, paradoxically, the accumulation of experience was not reflected in either the quantity or quality of PGs, although it is associated with the highest efficiency of PG provision (Table 4). It shows that the efficiency of education can be a "tricky issue". The level of efficiency, however, results from the relatively small number of schools and students (low synthetic measure of PG quantity), which leads to a high value of the basic part of the general educational subsidy per student (the highest correction coefficient X), which allows the maintenance of a high percentage of certified teachers. Thus it may be stated that the efficiency of the education system is improved here by minimising expenditure, i.e. the quantity of PGs, while accepting their relatively low quality. But is this a desired model?

The results for the ***cluster C*** may suggest that the general educational subsidy, which decreases as the number of students grows, should rather be spent on employing/promoting teachers in the so-called developmental posts. These posts are not only cheaper, but also, in view of the greater motivation for self-improvement, may lead to improved quality of PGs and higher efficiency of their provision. On the other hand, this points to the inappropriate construction of teachers' career paths and of the systems used to evaluate the effects of work in their profession.

Developing tools that enable improvement of teachers' remuneration system is the practical implication of the proposed method. The applied method indicated that the present teachers' remuneration system does not contribute significantly to the increase of quality of public goods in education. There are even symptoms indicating that comparable education scores, and in some cases even higher, are achieved by students taught by teachers of lower level of professional development (in our analysis referred to as 'developmental' teacher posts). Modified teachers' remuneration system should contribute to creating human capital of higher quality. Such an effect will be achievable through the change of the teachers' remuneration structure, without a need of any extra expenditures on education.

Conclusion

The authors have achieved the goals of the article both in the methodological dimension, and in terms of verifying the hypothesis put forward in the introduction. As a methodology, they suggest a research procedure serving the quantification of PG quantity and quality, and the evaluation of the efficiency of the process, followed by the identification of its financial determinants, assuming that PGs are of the nature of merit goods funded from the state budget. It should be emphasised that the methodology is universal, i.e. it may be applied to analyse the process of PG generation in different sectors of the economy.

The empirical research has borne out the theory that the supply of public goods in the education sector is determined by the structure of local budgetary funds used for that purpose, and not only by the total amount of public spending. The authors have identified three types of financing systems for secondary education. Paradoxically, it turns out that a petrified structure dominated by certified teachers does not guarantee the highest quality of education, and is associated with high efficiency of PG provision only in the case where there are few schools and students (class B). A model dominated by "developmental teacher posts", characterised by a relatively high rotation of staff (significant proportion of probationary teachers) seems a much better solution. Discussion concerning a motivational scheme to be applied to secondary school teachers' promotion paths is therefore justified. A concerning fact is that it is more advantageous to spend the general educational subsidy on employing/promoting teachers in lower posts, as these are not only less costly, but may also result in improvement in both the quality of PGs and the efficiency of their provision.

References

AFONSO, A. – SCHUKNECHT, L. – TANZI, V. (2005): Public Sector Efficiency: An International Comparison. Public Choice, *123*, No. 3, pp. 321 – 347.

AFONSO, A. – AUBYN, S. M. (2005): Non-parametric Approaches to Education and Health: Expenditure Efficiency in OECD Countries. Journal of Applied Economics, *8*, No. 2, pp. 227 – 246.

ANTONIS, A. – MANTHOS, D. – PANTELIS, K. (2011): Public Sector Efficiency: Leaving the Playing Field between OECD Countries. Public Choice, *146*, No. 1, pp. 163 – 184.

ALESINA, A. – TELLA DI, R. – MacCULLOCH, R. (2004): Inequality and Happiness: Are Europeans and Americans Different? Journal of Public Economics, *88*, No. 9, pp. 2009 – 2042.

ARISTOVNIK, A. – OBADIĆ, A. (2014): Measuring Relative Efficiency of Secondary Education in Selected EU and OECD Countries: The Case of Slovenia and Croatia. Technological and Economic Development of Economy, *20*, No. 3, pp. 419 – 433.

BALDOCK, D. K. – HART, K. – SCHEELE, M. (2014): Dobra publiczne i interwencja publiczna w rolnictwie. Available from: <https://webgate.ec.europa.eu/myenrd/app_templates/filedownload.cfm?id=45227AED-EB65-0E88-C0FF-9D706AF6572C>. [Accessed: 07/03/2015.]

BALLOU, D. (2005): Value-added Assessment: Lessons from Tennessee. In: LISSITZ, R. W. (ed.): Value-added Models in Education: Theory and Application. Maple Grove, MN: JAM Press, pp. 272 – 297.

BALLOU, D. – SANDERS, W. – WRIGHT, P. (2004): Controlling for Student Background in Value-added Assessment of Teachers. Journal of Educational and Behavioral Statistics, *29*, No. 1, pp. 37 – 66.

BORKOWSKI, B. – DUDEK, H. – SZCZESNY, W. (2003): Ekonometria. Wybrane zagadnienia. Warszawa: Wydawnictwo Naukowe PWN, pp. 63 – 64.

CZYŻEWSKI, B. – HNATYSZYN–DZIKOWSKA, A. – POLCYN, J. (2016): Problems of Quantifying Public Goods in the Healthcare Sector. Gospodarka Narodowa, *3*, No. 283, pp. 105 – 125.

DELLER, S. – RUDNICKI, E. (1993): Production Efficiency in Elementary Education: The Case of Maine Public Schools. Economics of Education Review, *12*, No. 3, pp. 45 – 57.

Di TELLA, R. – MacCULLOCH, R. – OSWALD, A. (2001): Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness. American Economic Review, *91*, No. 1, pp. 335 – 341.

DOLATA, R. (2007): Edukacyjna wartość dodana jako metoda oceny efektywności kształcenia na podstawie wyników egzaminów zewnętrznych. Warszawa: Centralna Komisja Egzaminacyjna.

FREY, B. S. – LUECHINGER, S. – STUTZER, A. (2009): The Life Satisfaction Approach to Valuing Public Goods: The Case of Terrorism. Public Choice, *138*, No. 3, pp. 317 – 345.

GATNAR, E. – WALESIAK, M. (eds) (2004): Metody statystyczne analizy wielowymiarowej w badaniach marketingowych. Wrocław: Wydawnictwo Akademii Ekonomicznej im. O. Langego.

GRUBER, J. – SENDHIL, M. (2005): Do Cigarette Taxes Make Smokers Happier? Advances in Economic Analysis and Policy, *5*, No. 1, pp. 1 – 43.

HANUSHEK, E. A. – KIMKO, D. D. (2000): Schooling, Labor-force Quality, and the Growth of Nations. American Economic Review, *90*, No. 5, pp. 1184 – 1208.

HSIAO, W. C. (1995): Abnormal Economics in the Health Sector. Health Policy, *32*, No. 1 – 3, pp. 125 – 139.

JAKUBOWSKI, M. (2008): Implementing Value-added Models of School Assessment. [EUI Working Papers RSCAS 2008/06.] Badia Fiesolana: European University Institute.

JOUMARD, I. – ANDRÉ, C. – NICQ, C. – CHANTAL, O. (2008): Health Status Determinants: Lifestyle, Environment, Health Care Resources and Efficiency. [OECD Economics Department Working Papers, No. 627.] Paris: OECD.

KAHNEMAN, D. – WAKKER, P. P. – SARIN, R. (1997): Back to Bentham? Explorations of Experienced Utility. Quarterly Journal of Economics, *112*, No. 2, pp. 375 – 406.

KAHNEMAN, D. – THALER, R. H. (2006): Anomalies: Utility Maximization and Experienced Utility. Journal of Economic Perspectives, *20*, No. 1, pp. 221 – 234.

KLIMOWICZ, M. – BOJKAŁO, W. (2012): Kapitał społeczny – interpretacje, impresje, operacjonalizacja. Warszawa: CeDeWu.

LASKOWSKA, I. (2012): Zdrowie i nierówności w zdrowiu – determinanty i implikacje ekonomiczno-społeczne. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

LEVINSON, A. (2012): Valuing Public Goods Using Happiness Data: The Case of Air Quality. Journal of Public Economics, *96*, No. 9 – 10, pp. 869 – 880.

LUECHINGER, S. (2009): Valuing Air Quality Using the Life Satisfaction Approach. Economic Journal, *119*, No. 536, pp. 482 – 515.

MUCHA, J. (2006): Interwencje państwa w system opieki zdrowotnej. Racje ekonomiczne. Polityka Społeczna, *7*, pp. 11 – 15.

NAZARKO, J. – ŠAPARAUSKAS, J. (2014): Application of DEA Method in Efficiency Evaluation of Public Higher Education Institutions. Technological and Economic Development of Economy, *20*, No. 1, pp. 25 – 44.

OSIATYŃSKI, J. (2006): Finanse publiczne. Ekonomia i polityka. Warszawa: Wydawnictwo Naukowe PWN.

POMIANEK, I. (2010): Poziom rozwoju społeczno-gospodarczego obszarów wiejskich województwa warmińsko-mazurskiego. Oeconomia, *9*, No. 3, pp. 227 – 239.

SAMUELSON, P. A. (1967): Fundations of Economic Analysis. New York: McGraw-Hill.

SHAW, J. S. (2010): Education – A Bad Public Good? The Independent Review, *15*, No. 2, pp. 241 – 256.

STANISZ, A. (2007): Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny, Vol. t.2. Kraków: StatSoft Polska.

TANZI, V. – SCHUKNECHT, L. (1997): Reconsidering the Fiscal Role of Government: The International Perspective. American Economic Review, *87*, No. 2, pp. 164 – 168.

ULBRICH, H. (2003): Public Finance in Theory and Practice. Cincinnati, OH: South-Western Educational Publishing.

Van PRAAG, B. M. S. – BAARSMA, B. E. (2005): Using Happiness Surveys to Value Intangibles: The Case of Airport Noise. The Economic Journal, *115*, No. 500, pp. 224 – 246.

WALESIAK, M. (2003): Uogólniona miara odległości GDM jako syntetyczny miernik rozwoju w metodach porządkowania liniowego. In: JAJUGA, K. and WALESIAK, M. (eds): Prace Naukowe Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Taksonomia 10, Klasyfikacja i analiza danych – teoria i zastosowania, Vol. 988, pp. 124 – 133. Wrocław: Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu.