Socio-economic Status and its Effect on Value-added[1]

Juraj FALATH – Zuzana JUŠČÁKOVÁ***

Abstract

The Lisbon strategy requires European education systems to produce applicable graduates in present environment of the increasing competitiveness and social cohesion. Inclusive growth starts with providing effective education to all children regardless of sex, disabilities or socio-economic status (SES). We present the methodology of identifying the value-added as one of the indicators of school effectiveness. In the sample of 26 schools and 1 229 pupils we observe their results in nationwide cognitive testing, information about family background and attitudes. We aim to explain disproportions between schools in the context of equal access to education.

Keywords: *value-added, latent class analysis, socio-economic status (SES)*

JEL Classification: I21, I24

Introduction

In the middle of 1980´s "… there has been an increasing interest in the problem of measuring the performance of teachers, schools, and districts independently of factors such as school composition that are related to student achievement, but cannot be easily manipulated" (Hibpshman, 2004, p. 4). "Although in many countries, performance of educational institutions have mainly focused on student attainment measures, such as the average score on standardised test" (OECD, 2008), this is not an optimal way to measure school effectiveness.

* Juraj FALATH, Comenius University in Bratislava, Facutly of Mathematics, Physics and Informatics, Department of Applied Mathemathics and Statistics, Mlynská dolina, 842 48 Bratislava, Slovak Republic; e-mail: juraj.falath@gmail.com

** Zuzana JUŠČÁKOVÁ, The National Institute for Certified Educational Measurements, Žehrianska 9, 851 07 Bratislava, Slovak Republic; e-mail: zuzana.juscakova@nucem.sk

[1] This methodology and the whole paper were prepared during the project *of Increasing educational quality on primary and secondary schools with usage of electronic testing*.

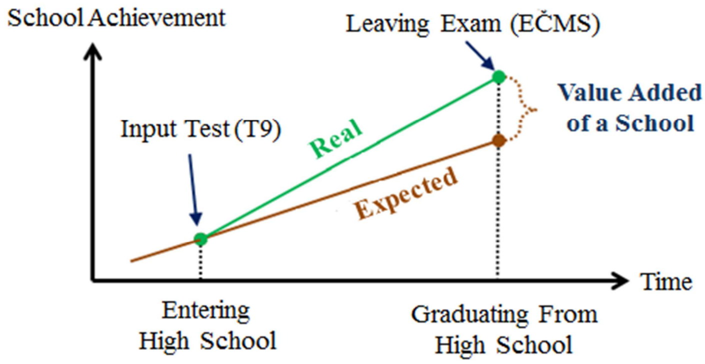
One of the main reasons is the fact that "it makes no sense to hold schools accountable for mean achievment levels when students enter those schools with large mean differences in achievement" (Raudenbush, 2004).

Our goal is to present a methodology of computing the value-added which is considered a much better indicator of school effectiveness than standardised test scores. We also aim to improve the Value-added model by accounting for socio-economic background. Chosing correct method to measure and process such a complex term as socio-economic status is not straightforward. In this article we attempt to present reasonable and very intuitive method of measuring this construct – the latent class analysis. This article presents the methodology and first results of modelling value added while accounting for socio-economic status (SES).

1. Theoretical Framework

In line with the idea of subtracting the initial skill of a pupil from school evaluation, the so-called Value-added scores have often been considered as much better indicators of school effectiveness for policy purposes (OECD, 2008). Value-added is generally considered a measure of progress in pupils´ school performance that can be attributed to the influence of a particular school. It can be calculated as a difference between actual result of a pupil at a standardised test and an expected performance of the same pupil based on the evaluation of a Value-added model:

G r a p h 1
Schematic Graph of Value-added



Source: Pavelka (2014), p. 108.

1.1. Economic Effects of Higher Quality of Education

Although there has been lack of relevant literature on the topic of economic value of higher teacher/school quality, this aspect might have significant implications when evaluating profitability of some policies to the future economic growth. As Hanushek (2011) claims, "high quality teachers are the most important asset of schools". The debate about economic effects of increased quality is mostly present in the context of teacher quality, but it can be easily transformed into the metrics of school quality.[2] Considering economic value of higher teacher or school quality is especially relevant in discussions about teacher salaries or school financing.

Several studies have shown that increased performance at standardized output tests lead to higher lifetime earnings of individuals. Murnane et al. (2000) estimate that one standard deviation increase in score at standardized test leads to 15% increase for males and 10% increase for females in annual earnings. Uppon slightly varied frameworks, Lazear (2003) estimates this effect to be 12% and Mulligan (1999) estimated it to be 11%. All mentioned studies relied on early--carrer incomes. Hanushek and Zhang (2009) claim that this number might even rise to 20% increase in lifetime earnings. In absolute values, one standard deviation increase in output test might amount to USD 150 000 increase in lifetime earnings of average individual in the USA. We stress the fact that this number is further multiplied by the amount of pupils in teachers´ class. These effects were further subject to several cost-benefit analyses which try to evaluate also costs of higher teacher performance (see e.g. Belfield and Levin, 2007). If there would be a proven straightforward influence of higher teacher salary on higher teacher performance, these economic effects could be a strong argument in the campaign for increasing teacher salaries.

1.2. Practical Use of Value-added Models in Policy

International experience shows that measuring value-added might have numerous policy implications which can improve quality of education by addressing and identifying issues rather than just putting the blame on particular schools. As OECD (2008) suggests, one possible policy implication is to identify particular schools which stand out both positively and negatively in terms of addedtheir contribution to the progress of pupils. Consequently, the educational methods and particular specifics of those excellent schools can be adapted to those struggling schools. On the other hand, identifying problematic schools is

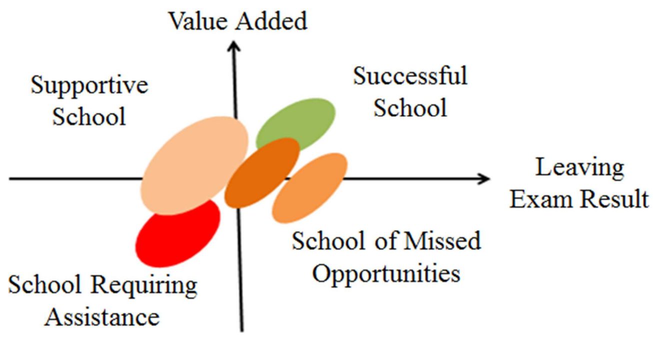
[2] If we evaluate economic effects of teacher quality, this can be easily transformed to school quality metrics because both teacher and school quality debate is based on value added.

also important, because attention and resources of policymakers can be drawn towards those struggling educational institutions. In general, the idea of this identifying process is following:[3]

An example of practical policy implication is the evaluation of teacher effectiveness in the USA. In some states in the USA, the value-added plays role in general teacher evaluations which might consequently affect their career progress and wage. Another policy in the USA called *No Child Left Behind*, obliges all states to implement school evaluation system based on the value-added (Hull, 2013). This policy uses the same principle as described in the Graph 2 and possible government interventions for schools requiring assistance for more than three years in a row might include changing teachers or curriculum, sending external experts to the school or even prolonging the current school year (Jorgensen and Hoffman, 2003). In some countries, pupils and their parents find value-added measures useful when deciding which school to attend (Raudenbush, 2004). Another examples of successful implementation of policies based on the value-added can be found in Poland (Żółtak, 2013) or in the United Kingdom (Kelly and Downey, 2010).

G r a p h 2

Distinguishing School Success



Source:Available on internet: <http://2013.ewd.edu.pl/educational-value-added-in-poland/> .

1.3. Limitations

Results of value-added shouldn´t be regarded as a perfect proxy of school effectiveness for several reasons. Firstly, this type of dataset is very likely to suffer from imperfections caused by missing data or insufficient representativeness

[3] The origin of the axis graph represents an average school both in terms of school achievement and value added. Excellent schools are those with values of both indicators above average and struggling schools are those with values below average in both indicators.

of the dataset.[4] Secondly, there is a serious concern that some pupils tend to cheat at those test, or some might provide false information about their personal characteristics, such as the socio-economic status. Thirdly, a significant influence of contextual factors might not be controled for by the model of value-added thus possibly leading to biased estimates of value-added. Fourthly, standardised tests[5] don´t reflect cognitive skills perfectly. Some teachers might also train their students to perform well at the particular test rather than to train their overall cognitive skills. Another problem of value-added models is that they don´t tell us the particular reasons of excellent or poor school performance. It is thus crucial to include more methods than just value-added alone. And finally, it is important to realize that linear regresson is based on comparing, which means that high quality of one subject is caused by low quality of another subject. Hence this doesn´t reflect a direct personal credit of a subject. This is a serious risk of interpreting the results when the model of value-added is misused.

1.4. Value-added Modelling in Slovakia

In 2015, for the first time in Slovakia, the National Institute for Certified Measurements in Education (NÚCEM) has provided to the principals information about the value-added of their school for last three cohorts of pupils. It was evaluated by a two-level hierarchical linear model in four year time frame (Kaclík et al., 2015). The model included basic variables of school context and individual characteristics and the results were not reported to the public.

1.5. Contextual Variables as an Improvement of Standard Models

"When attempting to evaluate schools work it is important to consider factors, which cannot be controlled or affected by school personnel" (Juščáková, 2014). The role of this article is to propose an improved methodology of value-added modelling in Slovakia which controls for contextual factors in calculating the value-added thus providing more precise and fair measure of school effects. Three contextual variables are regarded as having the most influence: socio-economic status, motivation and intellectual abilities.

[4] In both cases there is a risk of violating basic statistical assumptions such as e.g. the normality of our data. Another possibly violated assumption is independent observations, since teachers and students are obviously not independently distributed into schools.

[5] Nowadays, Norm Referenced tests (i.e. tests which are designed to sort students according to their performance) are being used in Slovakia. These tests are, however, not able to capture the real knowledge level of students. Furthemore, tests are not criterial hence the relationship between input and output tests is unknown.

2. Data Description

2.1. Descriptive Statistics

To model the value-added we have used results of the *T*9[6] test in Slovak language and literature (*T*9) as an input test, and results of the external part of exit examination in the Slovak language and literature as an output test (further we will refer to it as EČMS.[7]

Our sample comprises of 26 schools with 1 229 pupils altogether. The size of our sample is limited by a small amount of observations of the socio-economic status. We can, however, consider this satisfactory according to the "rule of thumb" of 30/30 (number of schools/number of pupils per school) (Kreft, 1996). Some studies claim more observations are necessary, for example "the results show that a small sample size at the second level (meaning a sample of 50 or less) leads to biased estimates of the second-level standard errors" (Maas and Hox, 2005, p. 86). Based on this reasoning we should aim to significantly increase the amount of observations in the future, which mainly concerns the measuring of the socio-economic status. Another reason for inaccuracies might be a weaker correlation between the input and output test, because "with a weaker correlation between the input and output tests, a large gap for inaccuracies is created, while those inaccuracies are absorbed by the residuals", meaning by the value-added (Ivica, 2013, p. 29). In our data set the correlation equals to r = 0.542, which is a satisfactory result. In plenty other countries, however, they have found higher values, e.g. in the United Kingdom between 0.52 and 0.87 (Fitz-Gibbon, 1997, p. 35). From the initial description we observe that in our data set, grammar schools (GRAM) consist on average of more pupils per school and higher share of girls per in school than vocational schools (VOC).

T a b l e 1

School Size and Share of Girls

	# of pupils	**# of schools**	**# of pupils per school**				**% of girls in school**			
			average	st. D.	min	max	average	st. D.	min	max
OVERALL	1 229	26	47.27	20.87	17	93	53.85	25.74	0	97.7
GRAM	691	13	**53.15**	21.38	24	93	**63.78**	11.77	41.7	80.6
VOC	538	13	**41.38**	19.37	17	89	**43.92**	32.06	0	97.7

Source: Own calculations.

[6] *T*9 (Testovanie deviatakov) – External testing of pupils of 9th grade at primary schools. Test comprises of 20 questions (always one correct answer among four possible choices).

[7] EČ MS (Externá Časť Maturitnej Skúšky) – External Part of Graduation Exam, high-stakes test of high schools. Test comprises of 64 questions. Questions 1 – 40 have always one correct answer among four possible choices. Questions 41 – 64 demand short written answers.

The results in EČMS are not only influenced by input results from *T*9. In this initial description we are looking for the variables that might explain the differences in achievement in EČMS. We observe that GRAM pupils achieved significantly better results than VOC students (p-value < 0.001) and also the fact that girls attained significantly better results than boys (p-value < 0.001).

This suggests that sex and school type might be suitable predictors for explaining the differences in EČMS scores. Another possible factor is a different power of a linear relationship between EČMS and *T*9 for GRAM/VOC schools and for boys/girls. We have compared average slopes of this linear relationship for different types of schools and different genders:

T a b l e 2a

School Achievement across School Types and Genders

		Average score	St. D.
GRAM	OVERALL	**76.06**	10.28
	girls	77.41	9.51
	boys	73.63	11.15
VOC	OVERALL	**59.82**	12.96
	girls	63.12	12.73
	boys	57.34	12.59

Source: Own calculations.

T a b l e 2b

School Achievement of Girls and Boys

	Average score	St. D.
girls	**72.51**	12.68
boys	**64.62**	14.45

Source: Own calculations.

T a b l e 3

Comparing Slopes of *T*9 – EČMS Relationship across School Types and Genders

		Average slope	St.D.	Min	Max
GRAM	girls	0.2598	0.2553	–0.04	0.72
	boys	0.2455	0.3116	–0.10	0.86
VOC	girls	0.3672	0.3036	0.01	0.93
	boys	0.3015	0.2556	0.03	0.78

Source: Own calculations.

Table 3 depicts a substantially higher average slope for VOC compared to GRAM and higher average slope for girls compared to boys. This might indicate that for certain groups (e.g. girls from VOC) the result in *T*9 is a stronger predictor of a success in EČMS. The interaction terms of *T*9 with sex and school type might be necessarry in estimating the value-added.

2.2. Adjusting Variables

Including unadjusted explanatory variables into the value-added model might cause problems with interpreting the numerical results of a regression. Therefore "it is necessarry to consider including relevant explanatory variables and correctly decide the form in which we include them into the model (a problem of centering, a level, a random intercept, a random slope, etc.)" (Kaclík et al., 2015, p. 12). The modification of variables is performed according to this cited manual of modeling the value-added.

The major exogenous variable in the Value-added model is an achievement of a pupil in the input test *T*9. If we aim to interpret the intercept as reasonable, average value, we must implement the group mean centering:

$$T9\_group_{ij} = T9_{ij} - \overline{T9}_{.j} \tag{1a}$$

where

$$\overline{T9}_{.j} = \frac{1}{n_j}\sum_{i=1}^{n_j} T9_{ij} \tag{1b}$$

and n_j represents the number of pupils in the *j*-th school. After this adjustment the intercept of a simple model consisting of the only exogenous variable – *T9_group*, can be interpreted as an average result of a pupil in EČMS within the particular school. According to our statistical description, GRAM and VOC recorded on average different results in EČMS.

We have to account for these differences by including the variable *School_Type* which we recode as follows:

$$School\_Type = 1 - r, \qquad for\ GRAM \tag{2a}$$

$$School\_Type = -r, \qquad for\ VOC \tag{2b}$$

Where *r* represents the relative share of GRAM in the data set. "A thorough look at mutual interactions of pupils within the school system shows that a group of excellent schoolmates can escalate the performance of previously poorly achieving student – so called ’peer effect‘" (Juščáková and Falath, 2015, p. 7). We again use centering to create a proxy for this effect, this time we use the grand mean centering:

$$\overline{MT9.} = \frac{1}{j}\sum_{j=1}^{J} \overline{T9}_{.j} \tag{3a}$$

$$T9\_peer_j = \overline{T9}_{.j} - \overline{MT9.} \tag{3b}$$

This variable describes to what extent an average result in $T9$ within the particular school differs from the mean value of all averages of schools in $T9$. In the model containing $T9\_peer$ as the only explanatory variable, the intercept can be interpreted as the average score in EČMS of pupils from an average school. We have also found out that girls performed significantly better in EČMS than boys, therefore using the same logic we should account for an equivalent of peer effect represented by the percentage share of girls in a particular school:

$$Girls\_Ratio_j = RGirls_j - 0.5 \tag{4}$$

where *RGirls* is the share of girls in ***j***-th school and *Girls_Ratio* is the gap of ***j***-th school from equilibrium ratio of girls and boys (0.5). With *Girls_Ratio* being the only explanatory variable of EČMS, the intercept can be explained as an average result in EČMS of pupil coming from the school with equal gender shares.

3. Measuring Socio-economic Status

"A general trend, observed in the recent literature on measuring SES and its effects, is that of moving from a conception of SES as a single composition toward a multidimensional normally distributed continuous latent construct that imposes its effects differently at different levels of observation, for example, students and schools" (Munck and Hansen, 2012, p. 51; Sirin, 2005). To put it simply, there might exist unobserved groupings of subjects (latent classes), which are linked within groups and which differ across groups by characteristical system of answers or features (latent profiles).

We have performed a latent class analysis using a library "poLCA" in a statistical program R (Linzer and Lewis, 2011). After determining the number of latent classes, the program assigns each pupil into one of the latent classes. Afterwards it displays latent profiles based on conditional probabilities which describe probabilities with which members of a particular latent class record high values of the SES indicators.

For example, for a hypothetical latent class we can obtain its latent profile by depicting the probabilities with which its members have the index of family wealth higher than 15,[8] have more than 200 books at home, at least one parent has finished university etc.

[8] The thresholds determining the categories of high, average and low attainment in the indexes of SES were chosen by us based on our perception of how affluent and disadvantaged person looks like within our Slovak population.

3.1. Indicators of the Socio-economic Status

Latent classes are formed based on the fact that pupils appear to be similar within classes and to differ across classes in the latent profiles described by the indicators of the SES.[9] We have recoded values of those indicators into three qualitative levels (1, 2, 3) as follows:

T a b l e 4

Recoded Values of SES Indicators

Highest education level of parents (HISCED)	1 – Without secondary school graduation (9%) 2 – Graduated secondary school (61.5%) 3 – University education and higher (29.5%)
Highest occupational status of parents (HISEI)	1 – Occupational status coding HISEI, score 0 – 46 (38.6%) 2 – Occupational status coding HISEI, score 47 – 61 (41.2%) 3 – Occupational status coding HISEI, score 62 and more (20.2%)
Home educational resources (HEDRES)	(number of answers "yes" to the list of educational items 1 – score HEDRES 0 – 6 (52.6%) 2 – score HEDRES 7 – 10 (47.4%)
Number of books at home (BOOKS)	1 – 0 – 25 books (0%) 2 – 26 – 200 books (76.1%) 3 – viac ako 200 books (23.9%)
The index of cultural possessions (CULTPOSS)	(within objects: literature, poetry, work of art) 1 – has 0 objects at home (12.1%) 2 – has 1 – 2 objects at home (42.6%) 3 – has all objects at home (45.2%)
The index of family wealth (WEALTH)	1 – Score 0 – 11 (4.3%) 2 – Score 12 – 15 (26.2%) 3 – Score 16+ (69.5%)

Source: Own calculations.

3.2. Criteria for Chosing the Final Model

Our task was to chose the optimal amount of latent classes which will describe the data properly and which will have latent profiles that can be well interpreted and will correspond with theory. The theory expects three factors of SES to have the main influence: social (education, occupation of parents), cultural (number of books at home, home educational resources, cultural possessions) and economic status (family wealth). Bearing in mind that "among the factors of SES, we can expect the influence of the cultural capital for GRAM students, for students of VOC the influences of social and economic status add up but with small power." (Juščáková and Falath, 2015), the experience from factor analysis therefore suggests relatively weak influence of all factors with cultural capital being the strongest.

[9] The formation of all subindexes of SES with detailed description of scoring of items can be found in the SES manual (Kudácseková and Juščáková, 2012, pp. 16 – 26).

For comparing model quality we will use the Bayesian information criterion (BIC) (Schwartz, 1978) as the main indicator as it has proven to be one of the most reliable criteria for determining the optimal number of latent classes (Nylund, Asparouhov and Muthén, 2007, pp. 556 – 559). We also employ the Akaike information criterion (AIC) and the value of log-likelihood function as auxiliary indicators. The main reason the BIC is favored lies behind the fact that it penalizes more for an excessive amount of parameters and classes which is suitable for our case.[9] We will try to minimize the value of BIC and AIC and to maximize the value of log-likelihood.

3.3. Results of Latent Class Analysis[10]

Factor analysis in (Juščáková and Falath, 2015) suggests that components of the SES aggregate into up to three latent factors – social, cultural and economic. Meaningful model should therefore consist of up to $2^3 = 8$ latent classes.[11] Opposing to the theoretical claims, SES has not proven to be the stable construct in our data set, partly because of the fact that our sample does not perfectly represent the whole population. Alarmed by this fact, we have tested whether pupils from GRAM and VOC have the same latent class decomposition. The hypothesis of two school types having different latent class decompositions was rejected with both VOC and GRAM pupils having almost identical latent profiles of SES.[12]

We have computed the criteria for all possible amounts of latent classes resulting in the following devlopment of the BIC criterion.

T a b l e 5

Values of BIC for Latent Class Models

2 classes	3 classes	**4 classes**	5 classes	6 classes	7 classes	8 classes
11 586	11 529	**11 514**	11 561	11 601	11 663	11 721

*Source:*Own calculations.

Values of AIC and log-likelihood, stating of which we omit, are constantly improving with increasing amount of latent classes proving our suspicion of their insufficient penalizing for excessive amount of classes. Our main indicator (BIC) reaches its minimum for the model with 4 latent clases and it rises significantly

[10] Model with too many classes can in our case lead to unclear interpretation of classes which we need to be able to make conclusions and collect observations about the behaviour of our population.

[11] The number of possible combinations of latent profiles if we evaluate each factor as either disadvantaged or not.

[12] We omit results and the proof for this part to make the argument line clear and not to focus on too many small problems. Latent class analysis assigned VOC and GRAM almost identical set of latent classes with almost the same latent profiles.

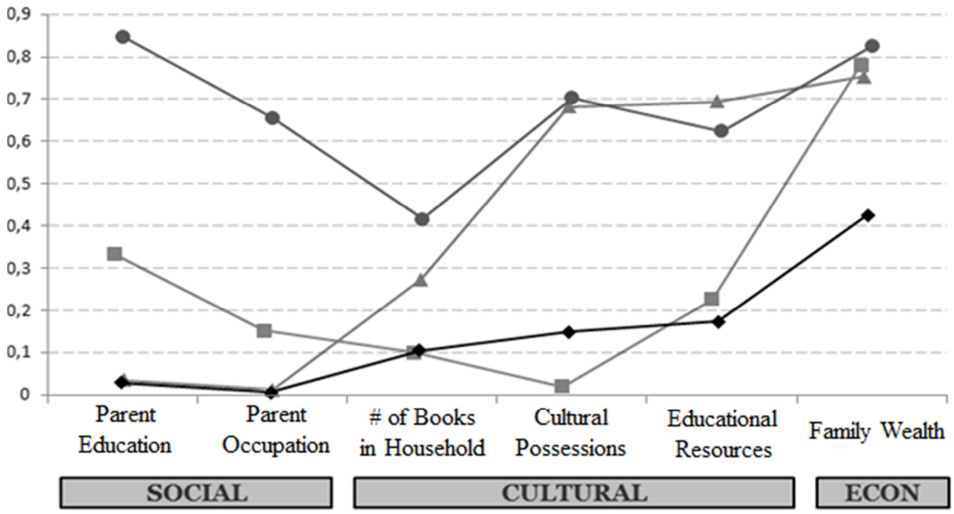
afterwards.[13] The model with four latent classes therefore turns out to be optimal and it has following graphic interpretation of latent profiles.

The interpretation of results is now straightforward:

- The first latent class, designated by the ◆ symbol, has almost zero probability of the highest educiation or occupation level of parents and also of all three cultural variables. This latent class can therefore be described as *socially and culturally disadvantaged group*. This is not, however, the case of an economic handicap as we will explain in the next paragraph.
- The second class, designated by the ■ symbol, has an obvious cultural handicap, high values of family wealth and relatively low values in terms of the social status. After looking more thoroughly one can find out that most of the members of this class have an average value (2) for both parent education and parent occupation thus this class should be evaluated as *culturally disadvantaged* with no social handicap.
- The third class, designated by the ▲ symbol, has an obvious social handicap while having substantially high values of the other two factors. We label this class *socially disadvantaged otherwise affluent.*[14]
- The fourth latent class, designated by the ● symbol, is simply *affluent* in all factors of the SES.

G r a p h 3

Latent Profiles of the Four-class Model



Source: Own calculations.

[13] The "rule of thumb" for significantly worse model based on BIC is the increase of this criterion by 10 or more.

[14] We emphasize the affluence of this class because except for the social disadvantage its latent profile copies the profile of the affluent group.

Remarkable is the fact that no latent class can be described as economically disadvantaged. Despite the fact that the *Class*1 has a low probability of having the highest level of an economic status, almost all other pupils within the class have an average level of economic status thus we cannot characterise this class as aconomically disadvantaged.[15] This observation compliments the fact that the economic factor has proven to be the factor with the weakest influence on the school achievement among all the SES factors. We consider this model the final one and it will enter the Value-added model as dummy variables of the latent class membership: *Class*1, *Class*2, *Class*3, *Class*4.

3.4. Socio-economic Status at the School Level

Dummy variables of a latent class membership are supposed to explain the variability which is caused by an individual socio-economic status. From our experience, an average socio-economic status of a particular school can also have an impact on pupil as a form of a "peer effect". To test this hypothesis we have created the *Af fluent_ratio* variable which describes the share of pupils belonging to the affluent class within the particular school. Values of this ratio span from 0% to 35% and the average value-equals 17.15% of affluent pupils within a school. Another possible influencing factor might be the share of the most disadvantaged pupils within a school. We have therefore created the *Disadvantated_ratio* variable as some form of a "negative peer effect". Values of this ratio span from 4.4% to 50% and the average value-equals to 21.31%.

4. Value-added Model

4.1. General Value-added Model

As we have already stated, modelling the value-added is a linear hierarchic problem because of the clustering effect at the school level. Gradual creation of the model and its variants, adding and removing of the variables, or determining the randomness of intercepts and slopes, is a statistical exercise based on the empirical findings from the past, statistical analyses of the particular data set and on understanding theoretical reccomendations. A general form of the value-added model is:

[15] An attentive reader might notice that only about 4% of pupils have the lowest rating of family wealth. It might therefore be hard for any latent class to record economic disadvantage. However, even after recoding the variable in order to increase the size of disadvantaged group, no economically disadvantaged latent class was found.

$$Y_{ij} = \beta_{0j} + \beta_{1j} * X_{ij} + \beta_{2j} * X'_{ij} + r_{ij} \tag{5a}$$

$$\beta_{0j} = \gamma_{00} + \sum_{k=1}^{p} \gamma_{0k} W_{kj} + u_{0j} \tag{5b}$$

$$\beta_{1j} = \gamma_{10} + \sum_{k=1}^{q} \gamma_{1k} W_{kj} \tag{5c}$$

$$\beta_{2j} = \gamma_{20} + \sum_{k=1}^{s} \gamma_{2k} W_{kj} \tag{5d}$$

where

Y_{ij} – the result of the i-th pupil within the j-th school in the output test (*EČMS*),

X_{ij} – the result of the i-th pupil within the j-th school in the input test (*T*9),

X'_{ij} – the first (individual) level characteristics (e.g. *sex, SES*),

W_{kj} – the second (school) level characteristics (e.g. *School_Type, Girls_ratio*),

r_{ij} – the residual at the individual level,

$\boldsymbol{u_{oj}}$ – the residual at the school level – thus being the value-added of a particular school.

This model can be rewritten into more intuitive form (Kaclík et al., 2015, p. 13):

$$Y_{ij} - \left(\gamma_{00} + \sum_{k=1}^{p} \gamma_{0k} W_{kj} + \beta_{1j} * X_{ij} + \beta_{2j} * X'_{ij} \right) = \boldsymbol{u_{0j}} + r_{ij} \tag{6}$$

$$\textit{actual achievement} - \textit{expected achievement} = \boldsymbol{VA} + \textit{residual}$$

4.2. The Null Model

The extent to which the hierarchic character of the data influences school achievement can be determined by using the null model which is the basic model of multilevel modelling. The result in output test (EČMS) will be in our case explained only by a constant and random terms on both the individual and school level:

$$E\check{C}MS_{ij} = \beta_{0j} + r_{ij} \tag{7a}$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \tag{7b}$$

where γ_{00} is the average score within the whole data set. Schools, however, differ in average scores thus they vary from the total average score by the random term $\boldsymbol{u_{oj}}$. After summing these two terms we obtain β_{0j}, the intercept that equals the mean score of a particular school. Pupils, obviously, differ in scores within a particular school from a school mean by a random term r_{ij}.

The null model divides the total variability into the variability within schools and the variability across schools. This enables us to infer how much of the variability can be ascribed to differences across schools:

$$Var\left(E\check{C}MS_{ij}\right) = Var\left(\gamma_{00} + u_{0j} + r_{ij}\right) = Var\left(u_{0j} + r_{ij}\right) = \sigma^2 + \tau^2 \qquad (8)$$

To validate whether the data set has a hierarchical nature we calculate the ICC coefficient (Intra class correlation coefficient) which can be calculated as the ratio of the deviance at the school level to the total deviance (McCoach and Adelson, 2010, pp. 152 – 155):

$$ICC = 105.52/(111.71 + 105.52) = 0.4857$$

The ICC suggests that 48.57% of the total variability is caused by the hierarchical nature of our data set hence the hierarchical model is needed.

4.3. Criteria for Chosing the Final Model

The creation of the Value-added model will consist of sequential addition and exclusion of variables at both the school level and the individual level. The suitability of a model will be evaluated by four indicators:

- ICC – a ratio of the variability unexplained at the school level to the total variability unexplained
- PVE_1 – a ratio of the variability explained at the individual level to the initial variability at the individual level
- PVE_2 – a ratio of the variability explained at the school level to the initial variability at the school level
- Deviance – an indicator of a model fit, equals BIC – k*ln(n) (= –2*Restricted log--likelihood)

Obviously, we will attempt to minimize the share of the variability at the school level (ICC) and at the same time to maximize the share of variability explained at both levels. It is not, however, necessarry to apply this strategy in all circumstances because in some cases the PVE can increase even by adding inappropriate predictors.[16]

A deviance is the indicator of a model fit and we will try to minimize it. The final model doesn´t ultimately have to be the one having the best statistical parametres, because "the aim of model creation is to find a model which explains the data properly and has simple and reasonable interpretation of parameters" (Kaclík, et al., 2015).

[16] It is also worth noting that part of the unexplained variability will show in the value added itself thus there is no need to explain all the variability.

4.4. Variables at the School Level

Thanks to the null model, we have been able to identify a substantial amount of the variability at the school level. For a correct inference of scores in EČMS we have to account for the differences between schools. We include several combinations of variables at the school level and identify an optimal model.[17]

The addition of any of the three variables decreases the share of variability unexplained at the school level to the total variability and explains a substantial part of the differences between schools. The best predictors are variables *School_Type* (VOC/GRAM) and *T*9*_peer* (peer effect) and if we include them both we achieve the best model so far in all indicators. We have also tried the model containing all three variables at the school level and it appears to be optimal. The ratio of the variability unexplained at the school level to the total variability unexplained decreased to 9.62% and we were successful in explaining 88.73% of variability between schools.

T a b l e 6

Adding Variables at the School Level

	Intercept	School_Type	*T*9_peer	Ratio_Girls	ICC	PVE_1	PVE_2	Dev.
0	■				48.6	–	–	9 376
1.1	■	■			18.87	0	75.38	9 339
1.2	■		■		14.71	0.03	81.75	9 338
1.3	■			■	39.1	–0.01	32.02	9 361
1.4	■	■	■		12.35	0.01	85.09	9 329
1.5	■	■	■	■	**9.62**	**0.03**	**88.73**	**9 318**

Source: Own calculations**.**

4.5. Variables at the Individual Level

The main variable in each Value-added model at the individual level is the score in the input test, in our case it is represented by the *T*9*_group*. The statistical description has shown significant differences in the results between girls and boys thus we also employ the variable *Sex*. Because the final model will have to explain as much variability at the school level as possible, we will keep all second-level variables and try to further improve the model 1.5 by including combinations of first-level variables.[18] The description of our data set suggests that slopes of the relationship between EČMS and *T*9 might differ between school

[17] To illustrate how the table works, we have also included the null model into the table. Its ratios of variability explained are obviously zero since it has no explanatory variables.

[18] We have tested all the possible combinations of models including all the stated variables. For the simplicity we are displaying mostly those steps which lead to the model improvement.

types and genders. It is therefore suitable to test the significance of interactions *School_Type*T9_group* and *Sex*T9_group*.

Table 7

Adding Variables at the Individual Level

		*T*9_group	Sex	*T*9_group*School_Type	ICC	PVE_1	PVE_2	Dev.
2.1	*	■			11.54	15.74	88.37	9 117
2.2			■		9.79	0.016	88.69	9 296
2.3		■	■		11.63	16.40	88.35	9 105
2.4		■	■	■	**11.66**	**16.60**	**88.34**	**9 106**

Note: * All models include variables *School_Type*, *T9_peer* and *Ratio_Girls* in the random intercept. In the table we display only variables that we have added to the model 1.5.

Source: Own calculations.

After including the interaction of school type and *T*9 the model appears to be statistically more suitable unlike the model containing the interaction of sex and *T*9 which is not fitting the data better. The decision lies between models 2.3 and 2.4. The interaction term of school type and *T*9 is significant which supports the initial finding of different slopes between GRAM and VOC. Therefore we chose the model 2.4 as optimal while it explains 16,6% of the variability at the individual level with the share of variability explained at the school level remaining approximately the same. This model provides us another finding that success in the input test is more important for GRAM pupils than for VOC pupils when it comes to determining success in the EČMS.

4.6. Extending the Model with the Socio-economic Status

In our last step we will attempt to improve the best model so far (2.4) by adding another variable at the individual level – dummy variables of membership into latent classes of the SES. We will not include the dummy variable *Class*1 membership into which will show in the intercept.[19]

Table 8

Adding SES to the Value-added Model

		SES	Affluent_ratio	Disadvantaged_ratio	ICC	PVE_1	PVE_2	Dev.
3.1	*	■			**11.35**	**16.54**	**88.68**	**9 100**
3.2		■	■		11.46	16.54	88.56	9 092
3.3		■		■	11.17	16.56	88.89	9 101
3.4		■	■	■	11.71	16.55	88.29	9 095

Note: * All models include every variable from the model 2.4. The table states those variables which are added to the model 2.4 in order to improve it.

Source: Own calculations.

[19] This allows us to interpret the intercept as the score of an average pupil belonging to the most disadvantaged class.

The model 3.2 has decreased the deviance but the variable *Af fluent_ratio* was not significant (p-value ≥ 0.383) hence we will not include it. The same problem applied to our two variables in models 3.3. and 3.4. Considering very wide confidence intervals of both variables we think that by including them we would reduce the model accuracy. We have also tested the presence of different slopes of SES in explaining EČMS. The tested interactions *SES * School_Type*, SES * Sex, SES * cRGirls a SES * T9_peer were all insignificant thus we will not include them in the model. The only improvement compared to the model 2.4 can in some cases be the inclusion of the SES as a variable at the first level. The SES was, however, insignificant in all dummy variables thus we don´t reccomend using this variable in the current situation of having too few observations.

4.7. The Final Model

We present two final models of the value-added. The first one contains the SES as an explanatory variable as an extension of recently published models (Kaclík et al., 2015).

We emphasize that this model is an example of a methodology of modelling the value-added on bigger, more representative dataset. In our case we suggest not to include SES into the equation because it is insignificant in all dummy variables. We also present more simple version of the Value-added model, without SES. This model meets all criteria for significance.

The First Model: with the SES

$$EČMS_{ij} = \beta_{0j} + \beta_{1j}T9\_group_{ij} + \beta_{2j}Sex_{ij} + \beta_{3j}Class2_{ij} + \\ + \beta_{4j}Class3_{ij} + \beta_{5j}Class4_{ij} + r_{ij} \tag{9a}$$

$$\beta_{oj} = \gamma_{00} + \gamma_{01}School\_Type_j + \gamma_{02}Girls\_ratio_j + \gamma_{03}T9\_peer_j + u_{0j} \tag{9b}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}School\_Type_j \tag{9c}$$

$$\beta_{2j} = \gamma_{20} \tag{9d}$$

$$\beta_{3j} = \gamma_{30} \tag{9e}$$

$$\beta_{4j} = \gamma_{40} \tag{9f}$$

$$\beta_{5j} = \gamma_{50} \tag{9g}$$

Resulting coefficients are:

T a b l e 9

Results of the Value-added Model with SES as a Regressor

Parameter	Estimate	Std. Error	df	t	Sig.	Low 95%	Upper 95%
Intercept	65.8863	1.0038	62.377	65.639	0.000	63.8800	67.8925
*T*9_peer	0.6754	0.1868	22.077	3.616	0.002	0.2881	1.0627
Girls_Ratio	6.8947	3.3536	24.985	2.056	0.050	–0.0123	13.8018
School_Type	6.7859	2.7722	20.918	2.448	0.023	1.0194	12.5525
Sex	2.1938	0.6376	1 197.641	3.441	0.001	0.9429	3.4446
*T*9_group	0.3040	0.0224	1 197.511	13.594	0.000	0.2601	0.3478
School_Type**T*9_group	–0.0841	0.0434	1 197.484	–1.939	0.053	–0.1691	0.0010
Class2	0.7423	0.9336	1 212.315	0.795	0.427	–1.0894	2.5740
Class3	0.0563	0.7693	1 208.279	0.073	0.942	–1.4530	1.5656
Class4	1.0941	0.8383	1 216.223	1.305	0.192	–0.5507	2.7389

Source: Own calculations.

All dummy variables of latent class membership are insignificant thus the influence of the SES on the leaving exam achievement is very weak. This might be partially ascribed to the small amount of recorded SES questionnaires. Another reason might be the fact that the effect of the SES is already contained in the input performance of a pupil, in the *T*9*_group* variable. Results support the opinion that "in general, slovak schools base education on meritocratic principals and eliminate the reproduction of social inequalities" (Juščáková and Falath, 2015).

The Second Model: without the SES

$$EČMS_{ij} = \beta_{0j} + \beta_{1j} T9_{groupij} + \beta_{2j} Sex_{ij} + r_{ij} \quad (10a)$$

$$\beta_{oj} = \gamma_{00} + \gamma_{01} School\_Type_j + \gamma_{02} Girls\_ratio_j + \gamma_{03} T9\_peer_j + u_{0j} \quad (10b)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} School\_Type_j \quad (10c)$$

$$\beta_{2j} = \gamma_{20} \quad (10d)$$

Resulting coefficients are:

T a b l e 10

Results of the Value-added Model without SES as a Regressor

Parameter	Estimate	Std. Error	df	t	Sig.	Low 95%	Upper 95%
Intercept	66.3575	0.8449	30.218	78.534	0.000	64.6324	68.0826
*T*9_peer	0.6832	0.1890	22.046	3.615	0.002	0.2913	1.0751
Girls_Ratio	6.8791	3.3910	24.872	2.029	0.053	–0.1065	13.8648
School_Type	6.8617	2.8064	20.918	2.445	0.023	1.0241	12.6992
Sex	2.1066	0.6345	1 200.384	3.320	0.010	0.8618	3.3515
*T*9_group	0.3052	0.0223	1 200.384	13.701	0.000	0.2615	0.3489
School_Type**T*9_group	–0.0854	0.0433	1 200.384	–1.973	0.049	–0.1703	–0.0005

Source: Own calculations.

Each variable is now significant.[20] The intercept has a value of 66.36 which is a score in output test of an average boy within average school with equal proportion of boys and girls. Furthemore, we can make following interpretations:

- with every 10% increase in the share of girls at *j*-th school, students of this school scored on average 0.68617 points more in the output test;
- with each one-point improvement against the average score of his/her school, students achieved on average:[22]
- 0.25396 points more in the output test if he/she comes from GYM, [21]
- 0.33936 points more in the output test if he/she comes from VOC;
- with each one-point improvement in average score of *j*-th school, students of this school achieved on average 0.6832 points more in the output test.

Conclusion

We have successfuly performed a latent class analysis that suggests that we can observe four different types of socio-economic background of pupils – socially and culturally disadvantaged, culturally disadvantaged, socially disadvantaged otherwise affluent and affluent in all fields. This simplifies the complicated construct of plenty socio-economic subindexes and variables into small number of easy-to-use dummy variables which can be easily implemented into the equation. Remarkable is our finding that no latent class is economically disadvantaged thus proving that family wealth plays little role in high school achievement of children in Slovakia.

In the Value-added model we have found two significant regressors at the individual level (*T*9, sex) and several significant regressors at the school level (peer effect, school type, girls ratio, intercept of school type and *T*9). We have shown that the SES has very weak predictive power which led into exclusion of this variable. The final model without the SES explained 16.6% of the variability at the individual level and 88.34% of variability at the school level. Our experience has shown a weak impact of the SES on the value-added and this might have several reasons. Firstly, the effect is probably hidden to the large extent in the results of input test because analyses show that SES has a significant impact on test scores. Secondly, despite the fact that low SES might lower chances of pupils to achieve good scores, this appears to be no longer true in the terms of the value-added. This finding would imply that Slovak schools don´t further

[20] The significance of variable Girls_Ratio is questionable (p-value = 0.053). We are keeping it in the model because of its very convincing confidence interval.

[21] The discrepancy betwen slopes is here caused by the presence of the interaction term *School _Type*T9_group*.

deepen inequalities caused by socio-economic status during the observed time and they appear to be providing the value-added "fairly" regardless of the socio--economic status. We would like to stress the fact, that this work proposes the methodology of implementing improved model of value-added in Slovakia and its main purpose is to present methods and ways to account for contextual factors when evaluating school performance. After developing the methodology for implementing other contextual factors, future research will aim to apply this complete methodology on more complete, robust and reliable dataset.

Special attention needs to be drawn towards the limitations of our study. The main problem in modelling the value-added was insufficient amount of the SES observations which might have led to insignificant coefficients and exclusion of the SES from the model. In spite of this fact, we have presented the methodology that can be used in the future to account for the socio-economic status in explaining value-added. Future development of the methodology should also include another contextual variables, namely pupil motivation and test of general abilities. Schools have no impact on these variables thus their effect should be accounted for in explaining school achievement. In further analyses, we also suggest investigating whether adding hierarchical structure to the latent class analysis can improve the quality of analysis.

References

BELFIELD, C. R. – LEVIN, H. M. (eds) (2007): The Price we Pay: Economic and Social Consequences of Inadequate Education. Washington, DC: Brookings Institution Press.

FITZ-GIBBON, C. T. (1997): The Value Added National Project: Final Report: Feasibility Studies for a National System of Value-added Indicators. London: SCAA.

HANUSHEK, E. A. – ZHANG, L. (2009): Quality-consistent Estimates of International Schooling and Skill Gradients. Journal of Human Capital, *3*, No. 2, pp. 107 – 143.

HANUSHEK, E. A. (2011): The economic Value of Higher Teacher Quality. Economics of Education Review, *30*, No. 3, pp. 466 – 479.

HIBPSHMAN, T. (2004): A Review of Value-added Models. Frankfort, KY: Kentucky Education Professional Standards Board.

HULL, J. (2013): Trends in Teacher Evaluation. How States are Measuring Teacher Performance. Alexandria: National Schools Board Association. Center for Public Education.

IVICA, L. (2013): Hodnotenie kvality školy. Bratislava: Fakulta matematiky, fyziky a informatiky Univerzity Komenského v Bratislave.

JORGENSEN, M. A. – HOFFMAN, J. (2003): History of the No Child Left Behind Act of 2001 (NCLB). San Antonio: Pearson Inc.

JUŠČÁKOVÁ, Z. (2014): The Contextual Variables of Academic Achievement of Pupil. Forum Statisticum Slovakum, *2014*, No. 2, pp. 57 – 63.

JUŠČÁKOVÁ, Z. – FALATH, J. (2015): Socioeconomic and Cultural Status versus Value Added.. in Education. [Conference Paper, Volume 23.] In: Conference: Etické a sociální aspekty v oblasti vzdělávání a pedagogickém výzkumu. At Czech Republic. Plzeň.

KACLÍK, P. et al. (2015): Technická správa: Pridaná hodnota školy vo vzdelávaní v predmete slovenský jazyk a literatúra. Bratislava: NÚCEM.

KELLY, A. – DOWNEY, Ch. (2010): Value-added Measures for Schools in England Looking Inside the "Black Box" of Complex Metrics. Educational Assessment, Evaluation and Accountability, *22*, No. 3, pp. 181 – 198.

KREFT, I. G. G. (1996): Are Multilevel Techniques Necessary? An Overview, Including Simulation Studies. Los Angeles: California State University.

KUDÁCSEKOVÁ, M. – JUŠČÁKOVÁ, Z. (2012): Správa zo štatistického spracovania dotazníka ekonomického, sociálneho a kultúrneho status žiaka. Bratislava: NÚCEM.

LAZEAR, E. P. (2003): Teacher Incentives. Swedish Economic Policy Review, *10*, No. 3, pp. 179 – 214.

LINZER, D. A. – LEWIS, J. B. (2011): An R Package for Polytomous Variable Latent Class Analysis. Available at: <http://www.jstatsoft.org/v42/i10/paper>.

MAAS, C. J. M. – HOX, J. J. (2005): Sufficient Sample Sizes for Multilevel Modeling. Utrecht University. Methodology, *1*, No. 3, pp. 86 – 92.

McCOACH, D. B. – ADELSON, J. L. (2010): Dealing with Dependence (Part I): Understanding the Effects of Cclustered Data. Gifted Child Quarterly, *54*, No. 2, pp. 152 – 155.

MULLIGAN, C. B. (1999): Galton versus the Human Capital Approach to Inheritance. Journal of Political Economy, *107*, No. 6, pt. 2, pp. S184 – S224.

MUNCK, I. – HANSEN, K. Y. (2012): Exploring the Measurement Profiles of Socioeconomic Background Indicators and their Differences in Reading Achievement: A Two-level Latent Class Analysis. Available at: <http://www.ierinstitute.org/fileadmin/Documents/IERI_Monograph/IERI_Monograph_Volume_05_Chapter_3.pdf>.

MURNANE, R. J. – WILLETT, J. B. – DUHALDEBORDE, Y. – TYLER, J. H. (2000): How Important are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings? Journal of Policy Analysis and Management, *19*, No. 4, pp. 547 – 568.

NYLUND, K. L. – ASPAROUHOV, T. – MUTHÉN, B. O. (2007): Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. Structural Equation Modeling, *14*, No. 4, pp. 535 – 569.

OECD (2008): Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value Added of Schools. Paris: Organisation for Economic Cooperation and Development.

PAVELKA, R. (2014): Výzkum a modelování přidané hodnoty škol jako náhodných efektů škol. Forum Statisticum Slovacum, 2014, No. 2, pp. 107 – 112.

RAUDENBUSH, S. W. (2004): What are Value-added Models Estimating and what does this Imply for Statistical Practice? Journal of Educational and Behavioral Statistics, *29*, No. 1, pp. 121 – 129.

SCHWARTZ, G. (1978): Estimating the Dimension of a Model. The Annals of Statistics, *6*, No 2., pp. 461 – 464.

SIRIN, S. R. (2005): Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. Review of Educational Research, *75*, No.3, pp. 417 – 453.

ŻÓŁTAK, T. (2013): Value-added Indicators for Polish Schools: An Overview. In: PICHANIČOVÁ, I., KUBIŠ, T. a POKORNÝ, M. (zostav.): Hodnotenie kvality vzdelávania – súčasný stav a perspektívy. [Zborník príspevkov z medzinárodnej konferencie.] Bratislava: NÚCEM. Dostupné na internete (12. 5. 2016): <http://www.nucem.sk/sk/projekt_esf/project/13>.