Edukacja 2017, 2(141), 113–135
doi: 10.24131/3724.170207
issn: 0239-6858

The significance of student age variance in estimating value added measures for Polish lower secondary schools

Aleksandra Jasińska-Maciążek
Educational Research Institute*

Research shows that there are at least four age related factors correlated with the level of school achievement and improvement: month of birth, early or delayed school entry and grade retention. Educational value added (EVA) measures are used in the evaluation of school effectiveness. They are estimated from relative increases in achievement. However, Polish EVA models do not include student age, which may bias the calculated indicators. This article presents findings from a study estimating the magnitude of this bias for Polish lower secondary schools over four years (2010–2013). The results demonstrated that the inclusion of student age only slightly changed EVA estimates. The bias caused by omitting age in recent years seems, therefore, rather minor.

Keywords: educational value added; EVA validity; birthday effect; age effects; early school entry; delayed school entry; grade retention; gains in school achievements.

The system of external exams has existed in Poland since 2002. This system makes it possible to assess the degree to which curricular requirements are met by students for the needs of selection purposes, to carry out individual assessments, to monitor processes taking place in the education system, as well as to perform evaluations (Dolata and Szaleniec, 2012). However, the use of examination results for internal and external evaluation requires the proper transformation of these results, since the test score depends largely on characteristics beyond the school's control, such as: students' level of intelligence, former achievements and school experiences, the social and financial situation of the family, the aspirations of parents and characteristics of the local environment (Dolata et al., 2013; 2014). One of the methods that makes it possibly to separate the factors that remain under a school's control from examination results is the educational value added (EVA) method.

EVA models are used to estimate the relative progress of students during a given education stage (what is important – progress relating to the school's operation). The basis for its calculation is at least two measurements of school achievement: the first, carried out at the beginning of the analysed teaching period and the second – at its end (OECD, 2008). However, EVA models may differ from one another in the scope of additional variables (the so-called contextual variables), which they use. The simplest EVA models do not include any contextual variables, but use only the results of school achievement measurements. These models assume that the influence of important

* Address: ul. Górczewska 8, 01-180 Warszawa, Poland.
E-mail: ola.jasinska@gmail.com

factors, which are beyond the school's control on the results achieved depends on the inclusion of information about the former achievements of a student. This assumption is correct for factors that are significant most of all for the achievements, not achievement increases, as well as for factors whose influence does not change over time between the measurements of achievement.

The need to include contextual variables in the model results from a conviction that by doing so, the EVA indicator may be less related to those variables that are beyond a school's control. If the value of the EVA indicator, understood as the measure of school effectiveness, was related to factors beyond the school's control, it should be stated that such an indicator is biased, since it is measuring what should not be measured by definition.

Two groups of models that include additional variables can be identified (Raudenbush and Willms, 1995). The first group includes only the individual characteristics of students, such as, for example, sex, age and social background. They indicate in which school the student is most likely to achieve a high test score, irrespective of the extent to which this score is attributable to the school and the favourable environment of the student. Models from the other group additionally include contextual variables, characterising the school's environment (for example, average achievements of students at the school or unemployment in the township). However, the EVA models do not include variables characterising, for example, teachers' efforts or the organisation of the teaching process, since these variables constitute what EVA indicators should measure by definition.

In Polish EVA models[1], in addition to data about previous achievements, information is taken into account on the student's sex, whether he/she is certified as dyslexic and whether he/she completed the programme in the standard three-year period or in four years mainly due to having been retained in a grade (Żółtak, 2013). No information about students' age is included in the models, although studies prove that age is of significance not only for achievement, but also for its increase (Ding and Davison, 2005; Hutchison and Sharp, 1999; Wen, Bulotsky-Shearer, Hahs-Vaughn and Korfmacher, 2012). The need of EVA models to include information about students' age was already emphasised several years ago by an international team of experts developing a report about the EVA method (OECD, 2008). Experiences of certain countries also necessitated the inclusion of this variable in EVA models (Hægeland, Kirkebøen, Raaum and Salvanes, 2005; Ray, 2006). An example of EVA models that include the variable describing students' age are those used in English schools[2] (Ray, McCormack and Evans, 2009).

The main purpose of this article is to analyse whether not including information about students' age in Polish EVA models results in biasing the indicators calculated for schools. If so, what is the scale of the problem and is it a good idea to include this information in EVA models? The presented analyses focus on indicators for lower secondary schools.

The problem analysed in this article is also of more general significance. The presented results show the age effect on relative increases in achievement.

Sources of students' age variance at the level of a given class

The statement that students in the same grade differ from each other in age may seem surprising, since the education law defines

[1] These models are calculated for lower secondary schools, secondary schools and technical colleges, with the exception of, among others, schools for adults, special schools, hospital schools and schools at education centres.

[2] This information, in addition to many other contextual variables, was included in the model as the relative age of a student in months within one's birth year. No additional variables to identify older and younger students from the main cohort were used, since their number was too small.

the age of starting school education as the calendar year in which a child has reached a certain number of years in age. Students learn in classes established on the basis of the birth year criterion. However, there are four main sources of age variance of students from the same class: (a) early school entry; (b) delayed school entry; (c) grade retention; (d) assumed definition of the start of compulsory school education, resulting in a variance of students' age reaching 12 months.

In Poland a child may enter school earlier with the consent of the principal (at the parents' request). This consent depends on an opinion from a psychological and pedagogical counselling centre, issued on the basis of a diagnosis of the emotional and mental maturity of the child (art. 16, § 2 of the Act of 7 September 1991 on the education system). The decision to send a child to school earlier therefore is connected with factors that are of great importance for a child's future results. This positive selection may result in observing, on average, better school achievement in a group of students younger than the main age cohort.

Delayed school entry is an opposite phenomenon. Entry may be delayed in cases that are "justified by important reasons" and is allowed only for one year[3] (art. 16, §3 of the Act). In most cases, these reasons negatively influence the chance of achieving educational success (for example, the lack of school readiness). Thus, a negative selection is observed here. Students for whom school entry has been delayed are older than their classmates, although they have finished the same number of years in school.

Students who were not promoted to the next class during a previous school year will also be older than the main age cohort. It is not only their age that distinguishes them from their classmates, but also the fact that the decision to repeat the same grade has been taken on the basis of poor results or too frequent absences from lessons. This indicates an obvious relation between the fact of being older due to grade retention and school achievement.

The last source of variance of students' age is the differentiation within the main cohort. In the context of analysing the factors relating to achievement, unlike the above-mentioned ones, birth month may be deemed random (Dolata et al., 2013; Dolata and Pokropek, 2012). One's genes, family status, etc. do not influence the month in which a student was born. This fact will be of significance for interpreting the modelled relations. In the case of non-random sources of variance of students' age (the first three mentioned), the models presented in this article do not allow us to interpret these relations in cause and effect categories.

Age effect on school achievement

Many studies show that students who are older in the year group and attend the same class outperform younger students in school achievement (Lee and Fish, 2010; Morrison, Griffith and Alberts, 1997; Sharp, 2002; Sweetland and De Simone, 1987). This effect is observed both with regard to various subject skills (Bell and Daniels, 1990; Hutchison and Sharp, 1999; Konarzewski, 2013; Martin, Foels, Clanton and Moon, 2004; Smith, 2009), as well as broadly understood teaching results, such as motivation, involvement in learning or attitude towards school and teaching (Martin, 2009). Some researchers focus on the situation of the youngest students in a class among children born in the same year and prove that these children more often experience school failures, such as not being promoted to the next class or being assessed as having special education needs (Langer, Kalk and Searls, 1984; Martin et al., 2004; May and Kundert, 1995; Verachtert,

[3] Unless the child has a certificate of special educational needs. If this is the case, postponement of compulsory education to the tenth year of age is possible.

De Fraine, Onghena and Ghesquière, 2010). These results support the hypothesis that students' age is significant for school achievement. However, there exist certain analyses that do not confirm this dependency (Dietz and Wilson, 1985). A review of studies carried out on various teaching levels shows that during consecutive years of school education, the strength of the correlation of achievement with age of the main age cohort decreases (Smith, 2009). This situation is found in many countries (Lee and Fish, 2010). The results of studies quite consistently show that the decrease in the strength of the correlation of achievement and students' age is greater during the first years of education than when students are older (Bell and Daniels, 1990; Hutchison and Sharp, 1999; Langer et al., 1984; Verachtert and others, 2010). These findings differ from each other in determining the moment when this dependency ceases to be significant. Some studies show that the advantage of older students over younger ones from the same class disappears as early as after the second or third grade (Crone and Whitehurst, 1999; Konarzewski, 2013; Stipek and Byler, 2001), whereas others indicate that this effect remains significant at the age of 10 (Smith, 2009), and ceases to be significant not earlier than at the age of 12 (Hutchison and Sharp, 1999) or 17 (Langer et al., 1984).

Analyses carried out by researchers all over the world showed that although the age effect for students from the main cohort is rather small, the age variance resulting from sending a child to school earlier or later and grade retention was clearly correlated with later achievements. This is due to the fact that early or delayed school entry is in most cases directly connected with characteristics of students that are significant for school achievement.

The issues of early and delayed school entry are often considered in the context of school readiness (Carlton and Winsler, 1999; Martin, 2009). The supporters of delaying compulsory school education treat it as an opportunity to achieve the proper level of cognitive and social development by a child, enabling his/her achievement of success in school. On the other hand, results of studies show that children who began school a year later than same-age peers had poorer or comparable school achievement in relation to their classmates (Martin, 2009; May and Kundert, 1995; Morrison et al., 1997). No postponed educational or non-school benefits resulting from delayed school entry were found (Cameron and Wilson, 1990; Lincove and Painter, 2006). Although such students are older than their classmates, they do not outperform them.

The other aspect of this phenomenon is represented by students with early school entry. They usually have significantly better results than their older classmates (Mayer and Knutson, 1999). This is due to the fact that this small group includes students selected with respect to characteristics that are connected with later school success.

The analyses on the issue of grade retention proves the negative relation of this phenomenon with achievement. A meta-analysis of the results of studies from the 1990s carried out by Shane Jimerson (2001) showed that students who repeated a grade at some stage of their education had poorer school achievement than their classmates, although not all results of the analysed studies confirmed this. Moreover, according to other studies, the effect of grade retention is negative, irrespective of the sex, class or age of the student (Martin, 2009).

The said effects may depend on the teaching methods applied and the solutions adopted in a given education system. However, the results of studies carried out in Poland allow us to draw the same conclusions as the studies abroad. Data collected during a study aimed at developing a test of skills at school entry showed that among six- and seven-year-olds, older children gained better results on the skills scales of reading, writing and mathematics than their younger schoolmates at the same educational level (Karwowski and Dziedziewicz, 2012). In standardisation studies of the School Readiness Scale, important effects for age were found, however, only in the case of boys for two of the six scales: independence and school skills (Frydrychowicz, Koźniewska, Matuszewski and Zwierzyńska, 2006). However, the weak relationships may result from the specific character of the tool used, which is designed for teachers, and teachers' assessments are usually less reliable than standardised skills tests.

In Polish nationwide studies carried out with the participation of students starting the 4th grade of primary school, it was also confirmed that students born earlier (up to 12 months) are more likely to gain better results in school achievement tests (Dolata et al., 2014; Jasińska-Maciążek and Modzelewski, 2014). These studies also found a strong negative effect for older students than for the main cohort (results lower by over 2/3 of the standard deviation than in the dominant age group). However, students younger than the main cohort gained comparable results as the main age group (Dolata et al., 2014) or had slightly better achievements in mathematical skills and language awareness, but not in reading skills (Jasińska-Maciążek and Modzelewski, 2014). The described effects were not large, although they were stronger than for the dependencies observed in Poland in the case of data from a test after the sixth grade of primary school or the lower secondary school leaving exam (Dolata and Pokropek, 2012).

The direction of the relation of students' age with teaching results changes if, instead of achievement, we take its increase into consideration. The results of longitudinal studies consistently show that greater increases in school achievement among younger students than older ones are observed (Ding and Davison, 2005; Hutchison and Sharp, 1999; Wei, Blackorby and Schiller, 2011; Wen et al., 2012), and a particularly fast pace of its development is observed in the initial years of education (Ding and Davison, 2005; Wen et al., 2012). This is why younger students catch up in time to older ones and the advantage resulting from the biological age difference loses its significance. These results are consistent with the results of screening studies, in which a decrease in the relation of school achievement to students' age during consecutive years of education was observed. In Polish longitudinal studies carried out on lower secondary school students, a similar regularity was determined. As results of two different (not equated) tests were used in these studies, the models focused on the relative increases in achievement. They showed that among students born in the same year and attending the same class, younger students gained higher relative increases in school achievement during three years of education at a lower secondary school, although the stated effect was not large (Dolata et al., 2013).

Students who began lower secondary school education later than would be assumed from their birth date are older than their schoolmates from the main age cohort. The reasons for delayed school entry or not being promoted to the next class in earlier stages of education may be connected with learning difficulties. Can we expect a different school achievement increase in this group than among the peers of these students? Some studies indicate that special education programmes have a compensating character, enabling students from disadvantaged backgrounds to achieve greater increases of language competence than others do, despite lower initial achievements in this area (Wen et al., 2012). However, other studies show that students who have learning difficulties attain school achievements at a similar pace as their peers who have no such problems (Ding and Davison, 2005). The results of the Polish studies showed that these students are characterised by smaller relative increases in school achievement in the humanities than students from the main age cohort, although this effect is significantly weaker than in the case of the level of achievements (Dolata et al., 2013). The relative increases of school achievement in mathematics were comparable in both groups.

The attempt to analyse or decide which processes are responsible for the observed relation of students' age with their achievement exceeds the framework of this article. The reader who is interested in such analyses can find interesting hypotheses in the studies of other authors (Konarzewski, 2013; Martin et al., 2004; Sykes, Bell and Vidal Rodeiro, 2009; Verachtert et al., 2010).

The results of the studies cited above showed that school achievement increases depending on students' age. From the viewpoint of estimating EVA indicators for lower secondary schools, the most important issue is the strength of the dependency observed at this level of teaching and whether the between-school variance of this characteristic of students is large enough to make significant changes in the EVA indicators (should they then be calculated with the inclusion of information about students' age). These issues are analysed below.

Method

Data used

The analyses were performed using data from the external examination system. The combined results of the test after 6th grade of primary school and the results of the lower secondary school leaving exam for each student were used. In order to examine the stability of the estimated effects, data for four consecutive cohorts of students taking the lower secondary school leaving exam in 2010, 2011, 2012 and 2013 were included in the analyses. The base of the combined exam results included students attending lower secondary school for the standard three years and those who attended for four years.

Exam results were scaled with the two-parameter logistic model in Mplus software using the two-parameter IRT model for dichotomous items and a graded response model (GRM) for polytomous items. The marginal maximum likelihood (MML) method was used. Expected a posteriori (EAP) estimates were used as an indicator of school achievement. In the case of the lower secondary school leaving exam, separate scales of results were developed for the mathematics and science part and for the humanities. The results were reported on a scale with a mean of 0 and standard deviation of 1 for a population of students (see Żółtak, 2013).

Dependent variable: relative increases in school achievement

In order to calculate the relative increases in school achievement, EVA models for the humanities and mathematics and science part for each of the four analysed examination sessions (8 models in total) were first estimated. Two-level random effect regression models with random effect for the intercept connected with the division of students into schools were used. Similar models were calculated for the purpose of publishing the EVA indicators in Poland (Żółtak, 2013)[4]. EVA indicators for a school are calculated from such models as the Bayesian predictions of random effects at the school level and their value is interpreted as the school's EVA. The formula of the very general model is as follows:

$$y_{ij} = w^k(x_{ij}) + \beta Z_{ij} + u_j + r_{ij} \quad (1)$$

The dependent variable (y_{ij}) is the result of the lower secondary school leaving exam in the humanities or in mathematics and science. The main independent variable is the

[4] Provided that publicly available EVA indicators are calculated from the so-called three-cohorts models, in which data from three consecutive examination sessions are jointly used. In the analyses presented in this article, data from one examination session were used in one model.

result of the test at the end of 6th grade (x_{ij}). The relation between them is modelled with the use of a *kth-degree* (w^k) multinomial since it is not linear. Parameters of the multinomial are estimated separately for students attending the standard (three-year) education programme and separately for students whose attendance was extended by a year. The modelling also includes some additional variables originating from the databases of the examinations: sex and having a dyslexia certificate (when taking the 6th grade test and lower secondary school leaving exam and the interactions between them). The vector of control variables describing the i^{th} student at the j^{th} school was described as Z_{ij}, whereas the vector of parameters (fixed effects) connected with these variables – as β. The equation of the model also describes two random effects: the level-2 residual for the j^{th} school (u_j), which is interpreted as the school's EVA and the random error at the individual level (r_{ij}), i.e. the level-1 residual of the i^{th} student at the j^{th} school. Both random effects form this part of the dependent variable's variance that could not be explained by the independent variables included in the model. On the other hand, the sum of these effects, being residuals of the EVA model, is equivalent to the difference between the achieved exam result (value of the dependent variable) and the result expected on the basis of the independent variables included in the model (expectation from the fixed part of the model).

Second, the aforementioned residuals (sum of the level-1 residual and the residual at the school's level) were calculated for each of the eight EVA models, which can be interpreted as the relative increase of students' school achievement (whether students achieved in the lower secondarv school leaving exam better or worse results than the average of other students with the same results during the test and other characteristics). They form the basic dependent variable used in the analyses presented in the article. These analyses were carried out in R software, with the use of the `lmer` function from the *lme4* package.

The decision to model the age effect on relative achievement increases, and not to directly introduce variables describing students' age into the EVA models, was taken for several reasons.

First, for the EVA method, understanding how increases (and not the level) in school achievement depend on the analysed variables is of key importance and such an analysis plan makes it possible to describe this dependency more clearly[5]. Moreover, from the viewpoint of the analysed problem, the most important issue is how many variances of achievement may be explained by including information about students' age in the EVA model, in addition to what is already explained by the variables included in the model (among others, the results of the 6th grade test). The accepted analysis plan made it possible to answer such a question. Even more, this strategy enabled the development of simpler models, which translates into clarity of interpretation.

Independent variables

Delayed school entry to lower secondary school. This variable takes the value of 1 if the student was, at the time of entering a lower secondary school, older by a year or two from the main age cohort of the class. In other cases, it has a value of zero. This group includes students whose school entry was delayed or who were not promoted to the next grade in primary school. The data used do not make it possible to differentiate these two situations.

[5] If students' age influenced only school achievements (not their increase), the results of the 6th grade test suffice to determine its significance (results of the test also depend on age) and no relation with increases in achievement would be observed. However, if it turns out that the aforementioned residuals from the EVA model (results of the lower secondary school leaving exam while controlling for the 6th grade test results) depend on students' age, this will mean that age is important not only for school achievement but also for its increases.

Early school entry to lower secondary school. The variable takes the value of 1 if the student was younger by a year than the main age cohort of the class at the time of entering a lower secondary school. In other cases, it has a value of 0.

Age of student. The age of students was determined on the basis of birth date (from the exam databases). Some data for this variable was missing in the sets combined with examination results. Observations of missing birth date data (0.02–0.28% depending on the cohort) and outliers (about 0.02%) were removed from the analyses. On the basis of students' birth dates, many indicators describing their age were produced, for example:

- Age of student in months – calculated on the basis of birth date, where 0 is the value for students from the main age cohort attending lower secondary school for three years, born in December (i.e. the youngest students of the main cohort).
- Relative age of student in months – calculated on the basis of birth date and transformed separately for the group of students attending for three years and separately for the group of students attending for four years (i.e. older by one year), so that it has the value of 0 for students born in December from the main age cohort attending for three years and four years. Thus, the value of the indicator shows the difference of age expressed in months in relation to students born in December with the same length of attendance in lower secondary school.

Table 1 presents the distribution of the birth year of students taking the lower secondary school leaving exam in 2013 together with information about which age group the students represented when they entered lower secondary school. The groups of students included in the analyses are marked in italics. For the other three cohorts, the distributions are similar, so they will not be presented here.

Table 1

Frequency distribution of students by education cycle among the cohort of students taking a lower secondary school leaving exam in 2013

Group of students by birth year	Birth year	3-year education cycle		Birth year	4-year education cycle		% of population
		n	%		*n*	%	
Delayed school entry to lower secondary school by 4 years	1993	8	0.00	1992	2	0.02	0.00
Delayed school entry to lower secondary school by 3 years	1994	50	0.01	1993	5	0.04	0.02
Delayed school entry to lower secondary school by 2 years	1995	*535*	*0.15*	*1994*	*101*	*0.84*	*0.18*
Delayed school entry to lower secondary school by 1 year	1996	*5 196*	*1.49*	*1995*	*1 211*	*10.07*	*1.78*
Main age cohort	1997	*339 977*	*97.61*	*1996*	*10 680*	*88.81*	*97.32*
Early school entry to lower secondary school by 1 year	1998	*2 517*	*0.72*	*1997*	*27*	*0.22*	*0.71*
Early school entry to lower secondary school by 2 years	1999	5	0.00				0.00
Total		348 288	100.00		12 026	100	100.00

Length of attendance. This variable has the value of 1 for students attending lower secondary school for four years. These are primarily persons who repeated a lower secondary school grade once. For students attending for the standard three years, this variable has the value of 0.

Between-school variance of students' age

If school achievement and its increase depends on students' age, some students may have an advantage over others resulting from age. However, this does not mean that indicators describing teaching results at schools are biased by these dependencies. If schools differed from each other by age of students or number of students with delayed or early school entry at a given school, a bias of such indicators would occur. If the distribution of these variables were the same for all schools, no school would gain an advantage resulting from students' age that is favourable for achievement. Thus, to complete the characteristics of the independent variables, Table 2 presents the results of analyses indicating how much schools differ from each another by students' age. To do so, the variance was decomposed into one part attributed to the division of students into schools and a within-school variance. Hierarchical linear models with a random effect for the intercept connected with the division of students into schools were used (Raudenbush and Bryk, 2002). The age of students expressed in months was the dependent variable in these models. The models did not include any independent variables (they were the null models). The analyses were carried out in Stata 13.1 software (*xtmixed* procedure) using the maximum likelihood estimation method.

The analysis of between-school variance of students' age showed that schools differ significantly from one another in this respect and the division into schools explains 5–7% of the variance of students' age (indicator calculated as the proportion of school effects variance to total variance, expressed as a percentage). This means that between-school variance of students' age at the level of lower secondary school is not presently large, although it is about two times larger than in the case of primary school, where it totals about 1–3%, according to the aforementioned studies (Dolata et al., 2014; Jasińska and Modzelewski, 2013).

Between-school variance of delayed and early school entry is presented with descriptive statistics for variables describing the percentage of such students at schools. The results are presented in Table 3. They indicate the existence of between-school variance, although the average percentage of students with delayed or early school entry is small (particularly in the case of early entry) and there are no such students in at least 25% of the schools.

Table 2

Between-school variance of students' age expressed in months. Random effects from the two-level linear model, random intercept models[*]

	Year of examination session			
Parameter	2013	2012	2011	2010
Variance of schools effects	1.464 (0.047)	1.339 (0.044)	1.001 (0.035)	1.217 (0.040)
Student level variance	19.556 (0.047)	19.152 (0.045)	19.795 (0.046)	20.083 (0.045)
Between-school variance	6.96%	6.53%	4.85%	5.71%
Number of students	360 244	374 285	387 909	408 599
Number of schools	6 415	6 419	6 418	6 362

[*] The values of standard errors are given in parentheses.

The results of the analyses proved the existence of between-school variance of the variables relating to students' age. Although it is not large, it should not be ignored in studies of the optimal form of EVA models.

Analysis plan

The purpose of the analyses performed was to examine whether the unexplained variance – in particular at the level of schools – in EVA models can be partially explained on the basis of information about students' age. If this is the case, EVA indicators would be biased by a partial dependency on students' characteristics that are beyond the school's control. The performed analyses were also performed to show how relative increases in achievement, which are the basis for calculating EVA indicators, depend on students' age, and thus what the bias of the indicators would be if no information on age is included in the models. To do this, the relation between relative increases in school achievement and students' age was studied for each of the four examination sessions and two dependent variables (relative increases in achievement in mathematics and science and the humanities). The analyses were performed separately for each of the four examination sessions to determine whether the observed effects were stable during consecutive years, which would mean that a general regularity, independent of exam or a specific cohort, exists.

These models used the relative age of the student in months, since increases in school achievement were calculated with the controlling variable of length of attendance (three or four years). This is why the relative increases in achievement relate to differences within a given length of attendance, irrespective of whether students who spent four years at a lower secondary school are actually a year older than students attending for the standard three years. Next, the existence of significant effects were verified for the variables of belonging to the group of students with delayed (by a year or two) or early entry. An answer was being sought on whether increases in achievement in these groups differ from the main age cohort's achievement increases, after excluding the significance of age expressed in months. Then, the relation of age on relative achievement increases was examined for the subgroup of students from the same cohort and students with early, one-year delayed and two-year delayed entry to lower secondary school (thus, the interaction of the relative age in months was examined in relation to the dichotomous variables describing membership in individual groups). Moreover, the

Table 3
Descriptive statistics of the proportion of early and delayed school entry students

	% of students with delayed lower secondary school entry				% of students with early lower secondary school entry			
	Year of examination session				Year of examination session			
Statistics	2013	2012	2011	2010	2013	2012	2011	2010
M	2.34	2.16	2.22	2.06	0.77	0.89	0.85	0.79
SD	5.76	5.62	5.81	5.46	2.27	2.70	2.13	1.81
25. percentile	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Me	2.34	2.16	2.22	2.06	0.77	0.89	0.85	0.79
75. percentile	2.70	2.54	2.56	2.47	0.64	0.99	1.00	1.02
99. percentile	26.32	26.67	26.67	22.22	9.09	9.09	9.09	8.82
No. of schools	6 415	6 419	6 418	6 362	6 415	6 419	6 418	6 362

need was verified of whether to include the two variables describing delayed entry to lower secondary school by one year and by two years in the model, or whether it could be simplified by having only one variable described as "starting secondary school later than expected by birth year". The next step was to test whether the dependency of relative increases in achievement on the variables describing students' age was the same for students attending the standard three years or the lengthened four years. To do this, interaction effects for variables describing the age effect on relative achievement increases and the dichotomous variable denoting students with a four year attendance were added to the model.

Testing the need to include dichotomous variables in the model reflecting the selection of students to individual groups (early, delayed school entry to lower secondary school, four-year attendance) and the interaction variables modelling the different age effects on achievement is necessary to correctly estimate the relation between students' age in months and relative increases in achievement, since the membership of students to a particular group is not random and, even more, is by definition related to school achievement. Including variables in the model describing membership in groups (if necessary), as well as the interactions between membership in a group and students' age will make it possible to estimate the selection parameters and various forms of dependencies for the analysed groups of students.

During the analyses, many decisions were needed on whether some of the tested variables should remain in the model. From the viewpoint of the analysed problem, general regularities making it possible to describe the relation between relative increases in achievement and students' age were more interesting than their results for the given cohort (cf. Deming, 1953). The results of individual examination sessions of consecutive class years were treated as realisations of data from the superpopulation of students completing lower secondary school in Poland (Malec, 2008). For this reason, the decision to keep a given variable in the model was taken on the basis of a test of significance. Additional support was provided by an analysis of the pseudo R^2 indicator as a measure of the percentage of the explained variance of the dependent variable by the independent variable added to the model (Domański and Pokropek, 2011). However, it is worth remembering that for multilevel models, this measure manifests a certain instability for estimations (Snijders and Bosker, 2012). This indicator was mainly used to assess which level of the analysis explains the variance after having added the analysed independent variables.

The second stage of the analyses consisted of comparing the EVA indicators calculated from the three-cohorts models with the exclusion of age (u_j from the model described in equation 2) with indicators from models including additional information on age (value of u_j from the model described in equation 3), defined by the functional form that proved the best in the previous level of the analyses (β_2 is the vector of parameters relating to variables describing the age of a student and membership in a group representing delayed or early school entry; AGE_{ij} is the vector of independent variables describing the age of i^{th} student at j^{th} school).

At this stage, three-year models were used, i.e. models including data from three consecutive examination sessions, since the published indicators of schools are calculated on the basis of such models. In equations (2) and (3), this was described with the following variables: the year of taking the 6^{th} grade test ($year\_s_{ij}$) and the year of taking the lower secondary school leaving exam ($year\_g_{ij}$) of *ith* student at j^{th} school. The other denotations are analogous to those used in equation (1).

The purpose of these analyses was to examine how much additional information about students' age would change the

$$y_{ij} = w^k \left(x_{ij}, year_{s_{ij}}, year_{g_{ij}} \right) + \beta Z_{ij} + u_j + r_{ij} \quad (2)$$

$$y_{ij} = w^k \left(x_{ij}, year_{s_{ij}}, year_{g_{ij}} \right) + \beta_1 Z_{ij} + \beta_2 AGE_{ij} + u_j + r_{ij} \quad (3)$$

estimations of the indicators. To estimate EVA indicators, exactly the same functions and estimation methods were applied that are used to calculate the three-year indicators for their publication (Żółtak, 2013)[6].

Results

Effect of age on relative increases of achievement

The first results to be presented are from the analyses in which the variable of relative increases of mathematics and science achievement is explored; the second – the results for the humanities. Because of the limited size of this article, the models of those variables which were found to be significant for predicting relative increases of achievement will be presented. The remaining results will only be discussed.

Mathematics and science

The first question is whether information about age and belonging to groups with early or delayed school entry to lower secondary school makes it possible to predict relative increases of achievement, which are the basis for calculating school effects. It turns out that this information, included in the model through the set of variables listed in Table 4, makes it possible to explain 0.66–0.85% of the variance in the relative increases of achievement in mathematics and science. The pseudo R^2 indicator proves this (Domański and Pokropek, 2011). This is not much; however, it is worth noting that it is the variance in the between-school level that is mainly being explained. The addition of this information to the model makes it possible to explain about 2.5–3% of between-school variance in the relative increases of school achievement. This may already translate into certain differences in estimating the indicators of school effectiveness for schools, depending on whether or not we include information about students' age when calculating these indicators.

The performed analyses confirmed that relative school achievement increases in mathematics and science were negatively related to students' age. The observed effect has a similar size in consecutive years, though it is not strong. Students younger by 12 months gain on average slightly more than one examination point (on a scale with a standard deviation of 15) of greater achievement increase than older students. This dependency has, in principle, the same strength among students from the main age cohort as among students with early or delayed school entry in lower secondary school, as well as among students attending for three or four years. This was examined by testing the significance of the interaction effects of the variable "relative age in months" with the dichotomous variables determining membership in the aforementioned groups[7].

The results of the analyses proved the existence of stable and significant effects for early and delayed school entry to lower secondary school. Students who were one year younger

[6] The author thanks Tomasz Żółtak for providing the scripts to calculate the EVA models. They enabled the application of a comparable methodology to calculate the models used for the publication of the EVA indicators.

[7] These effects were tested by separately including them in models taking into account information about age and membership in the specific groups. They proved statistically insignificant (given an assumed significance level of $p < 0.05$), with the exception of the interaction of "age" with the variable of students with a one year delayed entry to lower secondary school among students attending for three years, but only from 2010, and the same interaction, but among students attending for four years from 2013. The dependency was slightly stronger in these two subgroups. As these results were not confirmed for other years, they were deemed specific for the given examination session, possibly accidental, and thus insignificant in determining the general form of the analysed dependency.

Table 4
Age effect on relative mathematics and science achievement increases. Results of the two-level linear model, random intercept models

Year of examination session	2013		2012		2011		2010	
Fixed effects estimation								
Age (relative)	-0.104*	(0.004)	-0.119*	(0.004)	-0,094*	(0,004)	-0,122*	(0,004)
Delayed school entry	-1.586*	(0.128)	-1.636*	(0.132)	-2,360*	(0,128)	-2,291*	(0,122)
Early school entry	2.354*	(0.179)	2.428*	(0.167)	2,555*	(0,169)	2,178*	(0,157)
4-year attendance	0.049	(0.087)	0.044	(0.090)	0,060	(0,087)	0,139	(0,072)
Interaction: delayed entry x four-year cycle	1.343*	(0.285)	1.773*	(0.291)	1,997*	(0,294)	0,881*	(0,262)
Intercept	0.626*	(0.049)	0.694*	(0.046)	0,572*	(0,049)	0,733*	(0,043)
Random effects estimation								
School level variance	9.491	(0.209)	7.531	(0.173)	9,241	(0,204)	6,605	(0,152)
Student level variance	75.521	(0.180)	77.413	(0.181)	79,580	(0,182)	70,877	(0,158)
Pseudo R^2	0.66%		0.68%		0.67%		0.83%	
Pseudo R^2 (school level)	2.71%		2.74%		2.58%		3.07%	
Pseudo R^2 (student level)	0.40%		0.47%		0.44%		0.61%	
Summary								
Number of students	360 232		374 105		387 738		408 372	
Number of schools	6 415		6 418		6 418		6 362	
Log likelihood	-1 296 090.7		-1 349 821.6		-1 404 697.9		-1 455 028.2	

*The effect is significant at the level of 0.05. Standard deviations are given in parentheses.

than the dominant cohort when starting lower secondary school education gained higher relative increases in achievement in mathematics and science by over two points (about 1/6 of the standard deviation) than students starting lower secondary school education at the typical age. More importantly, this effect does not result from the difference in biological age, which is modelled by the variable *relative age of students in months*. This is due to the fact that the group with early school entry to lower secondary school included students with characteristics that are linked to the pace of increases in achievement. To complete the picture, it also should be added that no significant difference was found between the size of this effect in the groups of students attending for three years or four years.

On the other hand, students with delayed school entry to lower secondary school and educated in the standard three years made less progress by an average of 1.6–2.4 examination points (about 1/8 of the standard deviation) than students from the main cohort over that which resulted from differences in age in months. This difference is smaller for students who attended for four years, indicated by the value of the interaction effect between the variables *delayed entry* and *four-year attendance*. For 2013, 2012 and 2011, the effects eliminate each other, resulting in a linear dependence with no clear abrupt changes in the relative increases between students from the main cohort and those with delayed entry among students attending for four years. In 2010, however, students with delayed entry in the four-year attendance group had about a 1.4 point less (from -2.291 to +0.881) increase in mathematics and science school achievement.

The analyses also examined whether separate effects should be modelled for delayed entry to lower secondary school by one year and for delayed entry by two years. In other words, it was examined whether the dependency in groups of students with entry delayed by one year or two should be modelled by means of a segmented linear regression (with a breakpoint between these two categories) or whether it could be simplified to a linear dependency. The analyses showed that the size of the step between the two groups for the expected values is on average below 0.5 point and that only among students with a four-year attendance who took the exam in 2010 is its value above one point. Because of the small sizes of these effects, it was decided that the simpler model (in which we model the age effect on relative achievement increases for both these groups with a linear function) would be better.

It should also be added that relative increases in the achievement of students with four-year attendance do not differ on average from the relative increases of students who attended for three years. This is shown by the insignificant effects of the *four-year attendance* variable, presented in Table 4. However, this result does not prove that the increase of achievement observed among students with a four-year attendance is the same as increase in achievement observed among students who attended lower secondary school for three years. This is because a variable describing four-year attendance is included in the EVA models, on which the calculations were based for determining the relative increases in students' achievement. Thus, achievement increases of students who attended for a longer period are determined in comparison to other students who attended lower secondary school for four years.

To better describe the results presented in Table 4, a dependency for the example of a student year was also shown in Figure 1. It presents predictions from the fixed part

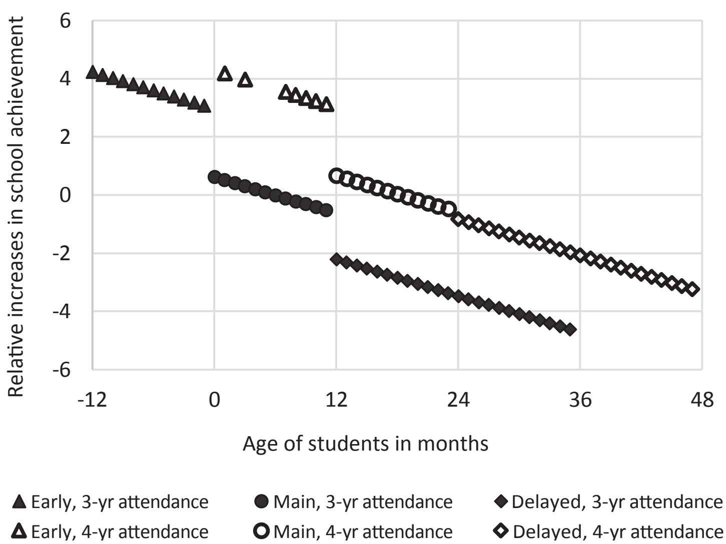

Figure 1. Dependency of relative increases in school achievements in mathematics and science (predictions from the fixed part of the model) on the age of students in months for the cohort taking the lower secondary school leaving exam in 2013.

of the model, in which relative increases in school achievement (the vertical axis) are predicted by the set of independent variables listed in Table 4. The horizontal axis describes the age of students expressed in months, where 0 is the age of the youngest students (i.e. born in December) in the main age cohort who attended for three years. The graph shows the dependency separately for students with the standard three-year attendance and separately for students with the four-year attendance, as well as separately for students with early and delayed school entry and students from the main age cohort. For the needs of the graph, a variable denoting the real age in months was used, so the values for students with four-year attendance are included in the range shifted by 12 months towards the higher values (since these students are 12 months older, spending an additional year in lower secondary school).

The humanities

In the case of increases in school achievement in the humanities, information about students' age explains a similar percentage of their variance as in the case of mathematics and science. The pseudo R^2 for the entire model is about 0.6–0.8%. The variables describing the age of students and membership in groups with delayed or early lower secondary school entry explain, as in the previously described models, most of the between-school variance (the pseudo R^2 at this level varies from 2.5 to 3.6%).

The age effect on relative increases in school achievement in the humanities is less stable in consecutive years than the effect for mathematics and science. Moreover, the completed analyses showed that a slightly more complex model was needed to describe this dependency. Relative increases of school achievement in the humanities are negatively related to students' age, although the effect is slightly weaker than for mathematics and science. Students older by 12 months gain on average relative increases in achievements from almost 0.5 to just under 1 examination point on a scale with the standard deviation equal to 15. No differences in the strength of this dependency were found between groups of students attending for three or four years, for the main cohort of students and for those with early school entry to lower secondary school.

However, students who started lower secondary school education one year earlier made slightly more progress while attending the school – their relative increases in achievement in the humanities are about two points higher from the achievement of students from the main cohort. The effect is, in principle, the same for students attending for both three and four years, since no significant correlation between variables describing early entry and a four-year attendance was found, except for the examination session of 2010. In 2010, the effect of early entry for students having attended for four years was stronger. However, as this was observed only for this particular examination session and, in addition, pertained to just under 0.01% of students, it was considered insignificant for the described effect.

The age effect on relative achievement increases for students starting lower secondary school education later than the main cohort proved more complex. A stronger age effect on achievement increases was found among students who delayed lower secondary school entry by a year, particularly in the group attending for three years. The additional analyses showed that a separate effect for a one year delay in lower secondary school entry and a separate effect for a two year delay in lower secondary school entry should be modelled for this group. A different picture of the effect was observed among students with the four--year attendance who delayed lower secondary school entry. The effects of the delay by two years proved very unstable during consecutive years and the observed differences may very likely be attributed to the small size of this subgroup among the surveyed students (making

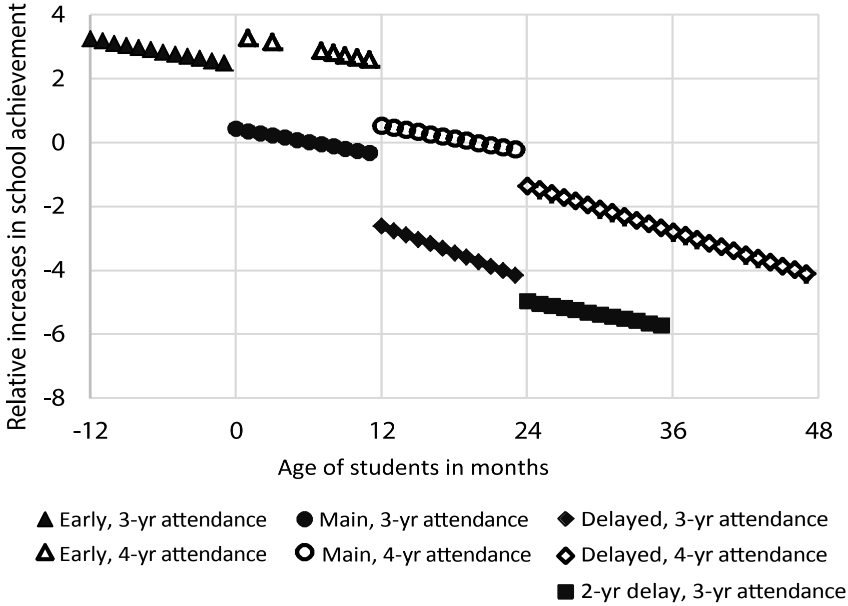

Figure 2. Dependency of relative increases in school achievement in the humanities (predictions from the fixed part of the model) on students' age in months for the cohort taking the lower secondary school leaving exam in 2013.

up 0.03% of the sample). To prevent the influence of random variances, a decision was finally made to model, by linear regression, the age effect on achievement increases for the group of students with a four-year attendance and delayed school entry, without identifying the additional effects of entry delayed by one year and two years[8]. Figure 2, presenting the results of the model for the example of a student year, describes its final form.

Among students with lower secondary school entry delayed by a year and who attended for three years, smaller relative increases in achievements are observed, on average by 2.3–3.3 examination points, compared to students from the main age group[9]. For students with four-year attendance, this difference, as in the case of mathematics and science, proved distinctly smaller (totalled 1.1–2.3). Students attending for three years, who started lower secondary school education two years later, obtained smaller increases in achievement by as much as 0.7–1.1 of a point (in addition to the difference resulting from age) compared to students whose entry was delayed by one year. Moreover, the interaction effects were confirmed between age expressed in months and the dichotomous variables of membership in the student group with a one year entry delay to lower secondary school and three-year attendance (insignificant effect only in

[8] Such a model was compared with the model in which the total effect of entry delayed by two years in the group of students with three- and four-year attendance was estimated. These models turned out to be similarly well matched to the data (on the basis of the *log likelihood* comparison). The selection of a model for students with four-year attendance was therefore dictated by its similarity to the model for the data from mathematics and science.

[9] This is the value of the difference between the oldest students from the main cohort and the youngest students from among those with a one-year delayed school entry, calculated on the basis of the estimated regression parameters.

Table 5

Dependency of relative increases in school achievement in the humanities on students' age. Results of the two-level linear model, random intercepts models

Year of examination session	2013		2012		2011		2010	
Fixed effects estimation								
Age (relative)	-0.068*	(0.005)	-0.074*	(0.005)	-0.047*	(0.004)	-0.040*	(0.004)
Early school entry	1.989*	(0.188)	1.817*	(0.174)	2.174*	(0.175)	1.852*	(0.177)
Delayed by 1 year entry among students with 3-year attendance	-1.362*	(0.625)	-0.832	(0.633)	-1.640*	(0.619)	-1.123	(0.623)
Interaction: delayed by 1 year entry among students with 3-year attendance students x age (relative)	-0.072	(0.037)	-0.129*	(0.038)	-0.135*	(0.037)	-0.163*	(0.037)
Delayed by 2 years entry among students with 3-year attendance	-3.759*	(0.414)	-4.610*	(0.414)	-5.358*	(0.403)	-5.891*	(0.461)
4-year attendance	0.099	(0.092)	0.132	(0.093)	0.172	(0.090)	0.195*	(0.081)
Delayed entry among students with 4-year attendance	-0.459	(1.026)	-1.025	(1.044)	0.061	(1.024)	-0.133	(1.018)
Interaction: delayed entry among students with 4-year attendance x age (relative)	-0.051	(0.057)	-0.030	(0.057)	-0.150*	(0.056)	-0.173*	(0.056)
Intercept	0.441*	(0.050)	0.475*	(0.049)	0.339*	(0.053)	0.294*	(0.050)
Random effects estimation								
School level variance	9.267	(0.209)	8.757	(0.198)	11.697	(0.252)	9.207	(0.206)
Student level variance	83.092	(0.198)	83.270	(0.194)	84.876	(0.194)	89.688	(0.200)
Pseudo R^2	0.61%		0.70%		0.79%		0.59%	
Pseudo R^2 (school level)	3.10%		3.56%		3.09%		2.45%	
Pseudo R^2 (student level)	0.33%		0.39%		0.46%		0.39%	
Summary								
Number of students	360 218		374 254		387 800		408 551	
Number of schools	6 415		6 419		6 418		6 362	
Log likelihood	-1 312 939.2		-1 364 232.2		-1 417 896.5		-1 504 006.9	

*Significant differences at the level of 0.05. Standard deviations are given in parentheses.

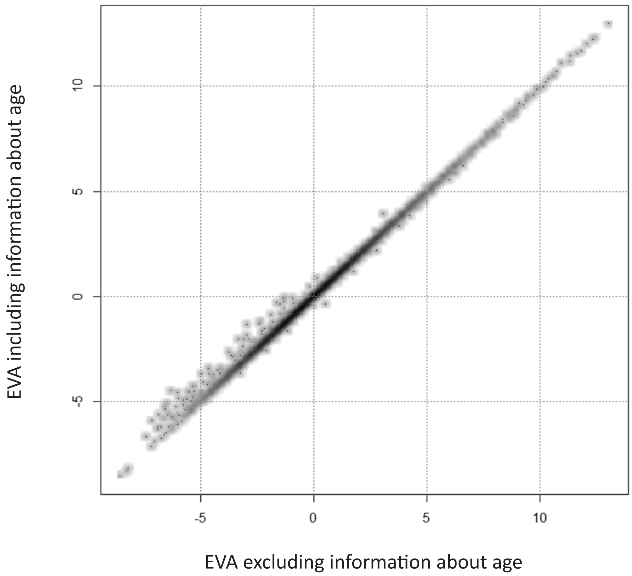

Figure 3. Comparison of three-cohorts EVA indicators for mathematics and science (2010-2012) calculated from the traditional model (excluding age) and the model including information about students' age.

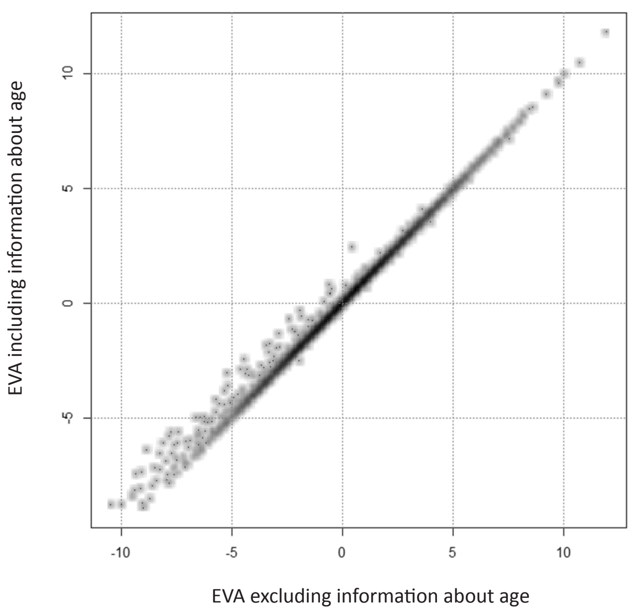

Figure 4. Comparison of three-cohorts EVA indicators for humanities (2010–2012) calculated from the traditional model (excluding age) and the model including information about students' age.

2013) and membership of the student group with delayed entry and four-year attendance (significant effect in 2011 and 2010). The age effect in these groups on relative increases in achievement is stronger.

EVA indicators including and not including information about students' age

The presented results show that information about students' age and membership in a group with delayed or early entry to lower secondary school makes it possible to explain some of the variances in the relative increases of achievement mainly at the between-school level. This may result in changed estimations of the EVA indicators if this information is included in the models. The discussed results also describe the basis of the bias of these indicators, since they showed which students were more likely to increase their achievements. Schools with a larger number of such students may have inflated EVA estimations. To examine whether this is the case, EVA indicators including information about students' age were calculated and compared to the indicators calculated as in the presently published measures.

EVA models were calculated for two three-year periods that included, in addition to the standard set of independent variables (included in models calculated for the purposes of publishing the indicators), the variables listed in Tables 4 and 5 (for mathematics and science models and humanities models respectively). These models assumed that the effects relating to students' age may differ in consecutive years (the interaction effects were added between the year of taking the exam and the variables included in the model) and were estimated on the basis of data from three examination sessions. The results of the comparison of EVA indicators with and without the inclusion of age are presented in Figure 3 (the results are analogous for the indicators for 2011-2013, which is the reason they are not presented).

This comparison shows that adding information about students' age to the model changes the point estimations of the indicators to a very small degree. The Pearson correlation among the indicators published so far and the indicators including information about students' age and students' membership in groups with delayed or early entry equals 0.999. We can also see that insignificant differences in the estimations of the indicators are found primarily in the lower values.

The results are similar for the humanities indicators. In Figure 4, a comparison is shown of indicators for 2010–2012 (the same results were obtained for the indicators for 2011–2013). The indicators calculated from the models including information about students' age and membership in groups with delayed and early entry only slightly differ from the indicators calculated with the model used to date. The correlation between them is only slightly smaller than for mathematics and science, and equals 0.998.

The size of the observed differences among standard three-cohorts EVA indicators and the values of three-cohorts indicators with information about students' age is additionally presented on a box plot graph (Figure 5). These differences are expressed in units of the standard deviation of the EVA indicators.

The box plots clearly show that for about 99% of schools (the area between the whiskers), the values of the indicators from both models do not differ by more than 0.13 of the standard deviation of the EVA indicators used so far. Also, for about 50% of schools, the value of the difference does not exceed 0.04 of the standard deviation. The distribution of these differences has a negative skew, which means that for the majority of schools, the values of the standard EVA indicators are slightly higher than the values of the indicators including information about students' age. However, clear individual differences (outliers on the left side of the distribution) are observed in the situation when the value of

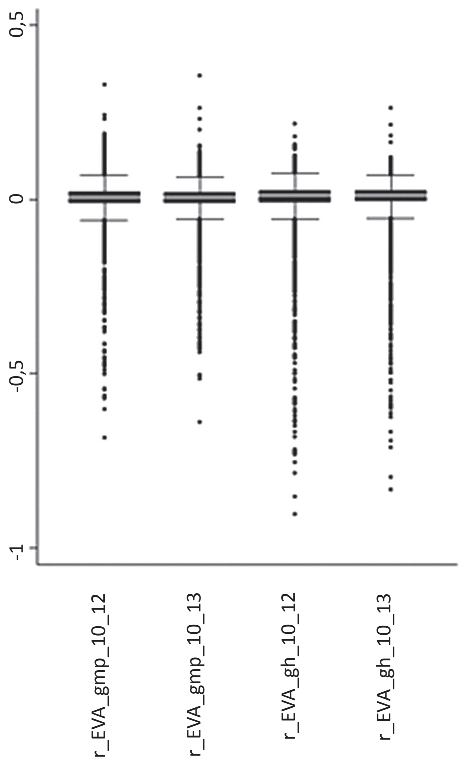

Figure 5. Differences of EVA indicators calculated from the standard models and models including information about students' age.
(gmp – mathematics and science, gh – humanities).

a standard indicator is understated as compared to an indicator that includes information about students' age. However, these differences do not exceed 1 of the standard deviation. These results show that for a great majority of schools (over 99%), including information about students' age in the models would not translate into noticeable differences in the values of the indicators. Significant differences would relate only to individual cases.

Summary

The purpose of the analyses presented in this article was to determine whether or not including information about students' age and their membership in groups with early or delayed lower secondary school entry in the EVA models biases the estimations of indicators on the schools. Inclusion of this information in the models made it possible to reduce a small part of unexplained variance (by about 0.7%). This mainly made it possible to explain between-school variance (by about 3%), which is of greater importance for estimations of EVA indicators. The presentation here of the students (with respect to their age or membership in the aforementioned groups) who are more likely to have larger increases in achievement describes a possible basis for the bias of the indicators. However, schools presently differ only slightly among each other in the age of their students or in having students with early or delayed lower secondary school entry. Thus, although the relative increases in school achievement depend on students' age and early or delayed school entry, the EVA indicators calculated from the models including this information virtually do not differ from those calculated without these data. The correlation between these indicators was 0.999 for mathematics and science and 0.998 for the humanities. The presented results show that for the great majority of schools, including information on students' age in the models would not translate into noticeable differences in the values of these indicators.

This is an important conclusion for the methodology of estimating EVA indicators, since these results show that the bias resulting from not including information about students' age in the currently used EVA models was negligibly small. However, in view of the fact that the addition of these variables to the models does not generate additional costs, but only necessitates the development of more complex models (these data are available in databases used during the calculation of the indicators), including this information in the models could be considered in the future, as it would will slightly improve them.

There is also a need to monitor the processes analysed here, because already next year, the between-school and within-school variance of lower secondary school students' age may become more significant given the manner in which reforms are being introduced that lower the compulsory school starting age[10]. In 2015, the first students from the "transition" period of the reform will have taken the 6th grade test and, three years later, the lower secondary school leaving exam. During the transition period, six- and seven-year-olds are starting their education in 1st grade, which will result in a natural increase in the age variance of students from the same grades. Moreover, the between-school variance of this feature will increase, since the percentage of six-year-olds in the 1st grade depends on, among other things, the activities of local governments and the schools themselves as well as on the disposition of the local community. Before the introduction of the reform in 2008, the average percentage of six-year-olds in the 1st grade was just under 1%. In following years, this percentage grew to reach 18% in 2013[11]. Not only did the share of six-year-olds in 1st grade grow, but the between-school variance of the percentage of six-year-olds in schools also increased. In 2009, the standard deviation of the percentage of six-year-olds in primary school was 3.63 and grew over the following years to reach 19.67 in 2013.

These data show that during the transitional period, student's chances of achieving a good individual result in an exam will depend more on one's age than it has thus far. On the other hand, the increase of between-school variance in age may result in a real bias of the EVA indicators if information about students' age is not included in their calculation. Although presently it may not be necessary to have lower secondary school EVA models include variables describing students' age, this may change in the coming years. The between- and within-school variance of students' age during the transitional period will depend more on activities undertaken by the school itself. As a consequence, including this information in the modelling will be, from a statistical viewpoint, even more problematic.

The last issue is a discussion on the character of the examination tests used in the modelling. The 6th grade test is a relatively easy one, proven by its left skewed distributions of results[12]. The consequence of this is a weaker differentiation of students with better achievements (the ceiling effect), whereas the lower secondary school leaving exam usually distinguishes these students better. This may bias the profile of the analysed dependencies of increases in school achievement by age. As achievements are positively related to students' age, the presence of the ceiling effect with the 6th grade test, as well as its absence (or smaller intensity) with the lower secondary school leaving exam may lead to estimates of larger relative increases in achievement among older students, which would further lead to the weakening of the analysed relation between age and achievement increase. However, the shapes of the distributions of exam results change among years, and the effects for the subsequent one-year models

[10] The Act of 19 March 2009 on the amendment to the act on the education system and amendment to certain other acts (Journal of Laws no. 56, item 458 as amended) introduced compulsory school education for six-year-olds. Previously, school education was compulsory from the age of seven. Further amendments postponed the year in which starting 1st grade was to be compulsory from the age of six and extended the transition period allowing parents to decide whether or not to send their six-year-old to primary school. If no further amendments are introduced, the 2015/2016 school year will be the last one in which both six- and seven-year-olds will start education in the 1st grade.

[11] Own calculations on the basis of Education Information System data from September records.

[12] The description of the distribution of results is available in the reports of the Central Examination Board at http://www.cke.edu.pl/index.php/sprawdzian-left/o-sprawdzianie-do-2014/88-sprawdzian/36-informacje-o-wynikach.

presented in this article are quite stable, which shows that if we experience any bias of this dependency, it will not be large.

Literature

Bell, J. F. and Daniels, S. (1990). Are summer-born children disadvantaged? The birthdate effect in education. *Oxford Review of Education*, *16*(1), 67–80.

Cameron, M. B. and Wilson, B. J. (1990). The effects of chronological age, gender, and delay of entry on academic achievement and retention: implications for academic redshirting. *Psychology in the Schools*, *27*(3), 260–263.

Carlton, M. P. and Winsler, A. (1999). School readiness: the need for a paradigm shift. *School Psychology Review*, *28*(3), 338–352.

Crone, D. A. and Whitehurst, G. J. (1999). Age and schooling effects on emergent literacy and early reading skills. *Journal of Educational Psychology*, *91*(4), 604–614.

Deming, E. W. (1953). On the distinction between enumerative and analytic surveys. *Journal of the American Statistical Association*, 48, 244–255.

Dietz, C. and Wilson, B. J. (1985). Beginning school age and academic achievement. *Psychology in the Schools*, *22*(1), 93–94.

Ding, C. S. and Davison, M. L. (2005). A longitudinal study of math achievement gains for initially low achieving students. *Contemporary Educational Psychology*, *30*(1), 81–95.

Dolata, R. and Pokropek, A. (2012). Czy warto urodzić się w styczniu? Wiek biologiczny a wyniki egzaminacyjne. In B. Niemierko and M. K. Szmigel (eds.), *Regionalne i lokalne diagnozy edukacyjne* (pp. 52–58). Kraków: Polskie Towarzystwo Diagnostyki Edukacyjnej.

Dolata, R. and Szaleniec, H. (2012). Funkcje krajowych egzaminów w systemie edukacji. *Polityka Społeczna*, 1(Numer tematyczny), 37–41.

Dolata, R., Hawrot, A., Humenny, G., Jasińska, A., Koniewski, M., Majkut, P. and Żółtak, T. (2013). *Trafność metody edukacyjnej wartości dodanej dla gimnazjów*. Warszawa: Instytut Badań Edukacyjnych.

Dolata, R., Hawrot, A., Humenny, G., Jasińska-Maciążek, A., Koniewski, M. and Majkut, P. (2014). *Kontekstowy model oceny efektywności nauczania po pierwszym etapie edukacyjnym*. Warszawa: Instytut Badań Edukacyjnych.

Domanski, H. and Pokropek, A. (2011). *Podziały terytorialne, globalizacja a nierówności społeczne: wprowadzenie do modeli wielopoziomowych*. Warszawa: Wydawnictwo IFiS PAN.

Frydrychowicz, A., Koźniewska, E., Matuszewski, A. and Zwierzyńska, E. (2006). *Skala gotowości szkolnej. Podręcznik*. Warszawa: Centrum Metodyczne Pomocy Psychologiczno-Pedagogicznej.

Hægeland, T., Kirkebøen, L., Raaum, O. and Salvanes, K. (2005). *School performance indicators for Oslo, Reports 2005/36*. Oslo: Statistics Norway.

Hutchison, D. and Sharp, C. (1999). *A lasting legacy? The persistence of season of birth effects*. Paper presented at The British Educational Research Association Conference, University of Sussex, Brighton. Retrieved https://www.nfer. ac.uk/publications/44411/44411.pdf

Jasińska, A. and Modzelewski, M. (2013). Międzyszkolne zróżnicowanie wyników nauczania po pierwszym etapie kształcenia. In B. Niemierko and M. K. Szmigel (eds.), *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych* (pp. 165–178). Kraków: Polskie Towarzystwo Diagnostyki Edukacyjnej.

Jasińska-Maciążek, A. and Modzelewski, M. (2014). Modele analizy zróżnicowania wyników nauczania. In R. Dolata (ed.), *Czy szkoła ma znaczenie? Analiza zróżnicowania efektywności nauczania na pierwszym etapie edukacyjnym* (pp. 65–92). Warszawa: Instytut Badań Edukacyjnych.

Jimerson, S. R. (2001). Meta-analysis of grade retention research: implications for practice in the 21st century. *School Psychology Review*, *30*(3), 420–437.

Karwowski, M. and Dziedziewicz, D. (2012). *Test umiejętności na starcie szkolnym. Podręcznik*. Warszawa: Instytut Badań Edukacyjnych.

Konarzewski, K. (2013). Wiek startu szkolnego a osiągnięcia w nauce w okresie wczesnoszkolnym. *Edukacja*, *124*(4), 5–19.

Langer, P., Kalk, J. M. and Searls, D. (1984). Age of admission and trends in achievement: a comparison of Blacks and Caucasians. *American Educational Research Journal*, *21*(1), 61–78.

Lee, J. and Fish, R. M. (2010). International and interstate gaps in value-added math achievement: multilevel instrumental variable analysis of age effect and grade effect. *American Journal of Education*, *117*(1), 109–137.

Lincove, J. A. and Painter, G. (2006). Does the age that children start kindergarten matter? Evidence of long-term educational and social outcomes. *Educational Evaluation and Policy Analysis*, *28*(2), 153–179. DOI: 10.3102/01623737028002153

Malec, D. J. (2008). Superpopulation. In P. J. Lavrakas (ed.), *Encyclopedia of survey research meth-*

ods (vol. 2, pp. 855–856). Thousand Oaks: Sage Publications.

Martin, A. J. (2009). Age appropriateness and motivation, engagement, and performance in high school: effects of age within cohort, grade retention, and delayed school entry. *Journal of Educational Psychology*, *101*(1), 101–114.

Martin, R. P., Foels, P., Clanton, G. and Moon, K. (2004). Season of birth is related to child retention rates, achievement, and rate of diagnosis of specific LD. *Journal of Learning Disabilities*, *37*(4), 307–317.

May, D. C. and Kundert, D. K. (1995). Does delayed school entry reduce later grade retentions and use of special education services? *Remedial and Special Education*, *16*(5), 288–94.

Mayer, S. E. and Knutson, D. (1999). Does the timing of school affect how much children learn? In S. E. Mayer and P. E. Peterson (eds.), *Earning and learning: how schools matter* (pp. 79–102). Washington: Brookings Institution Press.

Morrison, F. J. F. J., Griffith, E. M. and Alberts, D. M. (1997). Nature-nurture in the classroom: entrance age, school readiness, and learning in children. *Developmental Psychology*, *33*(2), 254–262.

OECD (2008). *Measuring improvements in learning outcomes: best practices to assess the value-added of schools*. Paris: OECD Publishing.

Raudenbush, S. W. and Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods*. Thousand Oaks: Sage Publications.

Raudenbush, S. W. and Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, *20*(4), 307–335.

Ray, A. (2006). *School value added measures in England.* Retrieved from http://webarchive.nationalarchives.gov.uk/20130401151715/http://www.education.gov.uk/publications/eOrderingDownload/RW85.pdf

Ray, A., McCormack, T. and Evans, H. (2009). Value added in English schools. *Education Finance and Policy*, *4*(4), 415–438.

Sharp, C. (2002). *School starting age: European policy and recent research*. Paper presented at LGA Seminarium "When Should Our Children Start School?", London.

Smith, J. (2009). Can regression discontinuity help answer an age-old question in education? The effect of age on elementary and secondary school achievement. *The B.E. Journal of Economic Analysis and Policy*, *9*(1). DOI: 10.2202/1935-1682.2221

Snijders, T. A. B. and Bosker, R. J. (2012). *Multilevel analysis: an introduction to basic and advanced multilevel modeling* (2nd ed.). Los Angeles: Sage Publications.

Stipek, D. and Byler, P. (2001). Academic achievement and social behaviors associated with age of entry into kindergarten. *Applied Developmental Psychology*, *22*(2), 175–189.

Sweetland, J. D. and De Simone, P. A. (1987). Age of entry, sex, and academic achievement in elementary school children. *Psychology in the Schools*, *24*(4), 406–412.

Sykes, E., Bell, J. and Vidal Rodeiro, C. (2009). *Birthdate effects: a review of the literature from 1990-on*. Retrieved from http://www.cambridgeassessment.org. uk/Images/109784-birthdate-effects-a-review-of-the-literature-from-1990-on.pdf

Verachtert, P., De Fraine, B., Onghena, P. and Ghesquière, P. (2010). Season of birth and school success in the early years of primary education. *Oxford Review of Education*, *36*(3), 285–306.

Wei, X., Blackorby, J. and Schiller, E. (2011). Growth in reading achievement of students with disabilities, ages 7 to 17. *Exceptional Children*, *78*(1), 89–106.

Wen, X., Bulotsky-Shearer, R. J., Hahs-Vaughn, D. L. and Korfmacher, J. (2012). Head Start program quality: examination of classroom quality and parent involvement in predicting children's vocabulary, literacy, and mathematics achievement trajectories. *Early Childhood Research Quarterly*, *27*(4), 640–653.

Żółtak, T. (2013). *Statystyczne modelowanie wskaźników edukacyjnej wartości dodanej – podsumowanie polskich doświadczeń*. Warszawa: Instytut Badań Edukacyjnych.

This article is based on data gathered in research carried out within the systemic project "The development of educational value-added method for enhancing the evaluation function of national examinations" implemented by the Educational Research Institute and co-financed from the European Social Fund (Human Capital Operational Programme, Priority III High quality of the education system). A preliminary version of this article was published in Polish in *Edukacja*, *133*(2), 2015.