Studies in Depth

Effects of age and school year

An overall picture and basic review of methods

Skolverket

The publication can be downloaded from
www.skolverket.se/publikationer

ISBN: 978-91-86529-53-6

Production: AB Typoform
Stockholm 2011

Effects of age and school year

An overall picture and basic review of methods

Contents

Foreword

The target group in International studies can be defined in different ways. One approach is to start from school year, as the IEA studies (International Association for Evaluation of Educational Achievement) does and then focus, for instance, on school year 4 or 8. Another approach is to start from age as in the OECD PISA study (Programme for International Student Assessment), where 15-year-olds make up the target group. In Sweden school starts at the age of 7, whilst students in many other countries start at the age of 6 or even earlier. This means that in some international studies Sweden has students participating who are somewhat older than the average for other participating countries (IEA studies), but who have attended just as many years in school. Sweden participates in other studies such as PISA where students are just as old – 15 years – but have one school year less than students in most other countries.

Since the issue is highly relevant for Sweden, as well as other countries, it is important to determine as fully as possible the importance of both effects on student results in different subjects.

In the current study, the effects of age and school year have been investigated in a reading comprehension study from 2001 with Swedish students, a reading comprehension study from 2006 using data from Norway and Iceland, a study of Swedish students' knowledge in mathematics and physics in upper secondary school from 2008, as well as a study of democracy, society and citizenship from 2009 with Swedish students.

The analyses show that the effects of age and school year on student results are dependent on a number of factors, such as the specific age and school years studied, the subjects and countries studied. This means that it is not possible to simply correct for the effect of age and school year by adding or removing a constant. The effects vary substantially between age/school year, subject and country.

Anders Auer is responsible for the analyses and has written the report.

Skolverket, August 2011

Sverker Härd
Head of unit

Anita Wester
Project leader

Introduction

The aim of this report is to provide the National Agency for Education with a rapid and simple overview of the problem complex concerning student age and school year in international studies of student performance. An additional aim has been to report on the approaches that can be taken for estimating the effects of age and school year on test results, and the type of data required for such analyses to be carried out. Assessing the relative effects of age and school year is crucial for developing an understanding of the extent to which these factors may affect the relative position of Sweden in different international studies of student performance.

The report contains a summary and presentation of research published earlier in the area, as well as replications of some earlier studies in order to explain the method. The review of the research can be regarded as an introduction to the area and is by no means comprehensive.

Chapter 1 provides an introduction to the problem area, as well as a brief review of the research. Chapter 2 describes the approach taken to estimate the effects of age and school year. Chapter 3 presents results regarding estimates of data from Swedish students in PIRLS[1] 2001 (replication of Gustafsson 2009), data from Norway and Iceland from PIRLS 2006, Swedish data from TIMSS 1995, school years 6, 7 and 8 (replication of Cliffordson 2008), Swedish data from TIMSS advanced 2008, Swedish students from school years 2 and 3 at upper secondary school, as well as data from ICCS 2009, Swedish students in school years 8 and 9. Chapter 4 discusses the problems inherent in different methodological approaches, as well as the ways in which methods could be refined and developed based on already published research, and Chapter 5 concludes with a discussion on how the results can be evaluated. The Appendix contains more detailed information about the different replications.

In all international measures of knowledge, it is important to try to make the results between students from different countries as comparable as possible, e.g. by giving the same test to all students, irrespective of country (even though not all students in the same country necessarily answer the same questions) and that they have equivalent time at their disposal etc. Another way of maximising comparability is to fix either the school year students attend, or their age. Both TIMSS and PIRLS have chosen to fix school year, which means, for instance, that all students participating in TIMSS attend school year 4 or 8 (one test for school year 4, and another for school year 8). But since the age at which school starts varies in different countries, this means that with such an approach, students in different countries may have different average ages. As regards Sweden, this means that Swedish students participating in e.g. TIMSS school year 4 are half a year older than the average age internationally.[2] In PISA, however, the choice was made to fix age so that all participating students in PISA were aged 15. For Sweden with its late

1. The document uses the abbreviations generally used in different international studies. For a more detailed presentation of the International studies, together with explanations of the abbreviations, see National Agency for Education 2004.
2. The average Swedish age was 10.8 years in TIMSS 2007, whilst the average age for EU/OECD was 10.3 years, a difference of half a year.

school start this means that about 95 percent of all Swedish students participating in PISA attend school year 9. In many other countries the majority of students on average attend year 10, and in some countries even year 11. [3]

The consequences for Sweden of fixing school year (as in TIMSS and PIRLS) is that Swedish students enjoy a relative advantage as they are somewhat older and thus probably more mature. However, when age is fixed, as in the case of PISA, the consequence is that Swedish students have not had the opportunity to learn as much as students in those countries participating in PISA where the majority of students are attending school year 10. This puts Swedish PISA students at a relative disadvantage.[4]

The problem complex mentioned above means that comparisons between two countries with different ages for school start do not give a fully accurate picture, irrespective of whether it is TIMSS, PIRLS or PISA. Countries with a late school start are favoured in TIMSS and PIRLS, whilst countries with an early school start are favoured in PISA. This brings us to the question of how large the effect is on students' academic skills, as measured by test results, of an additional year in school, as well as the effect of being one year older. Obtaining an understanding of the relative size of these effects is important when making comparisons between different countries or comparing Sweden's relative position in PISA to that in e.g. TIMSS.

3. Examples of such countries are Australia, Canada and the UK.
4. It is worth noting that Finland, just like Sweden, has a relatively late school start and that the majority of Finnish students also attend school year 9 when participating in PISA. Despite this, Finland is the country which has the best performance in the three earlier PISA studies.

What does the research tell us?

The dominant method used in the research for focusing on comparisons between the effect of age and school year, is "between grade regression discontinuity design". This report focuses explicitly on this method. Cahan & Davis, 1987, and Cahan & Cohen, 1989 are regarded as the originators of this method, and it will be explained in greater detail in the next chapter. Cahan & Cohen estimated the school year effect as approximately twice as great as the age effect, and even though the estimates vary depending on the data used, essentially, the school year effect seems to be greater than the age effect. But exceptions exist. Martin, Mullis & Foy presented during the IEA conference in Taipei 2008 analyses of the effects of age and school year based on data from PIRLS 2006 (reading comprehension) for Norway and Iceland, and found that in both countries the age effect was twice as large as the school year effect for school years 4 and 5.

Cliffordson also reported results for Swedish TIMSS 1995 data (mathematics and natural sciences), school years 6, 7 and 8 at the same conference. [5] She finds that the school year effect is approximately twice as large as the age effect, even though the school year effect is less between school years 7 and 8 compared with that between school years 6 and 7. Cliffordson's results will be presented and replicated in Chapter 3. Luyten (2006) who also analysed a number of countries using TIMSS 1995 (but not Sweden), for school years 3–4 in mathematics and natural sciences, also finds that the school year effect is generally stronger than the age effect, although their relative sizes vary between countries.[6]

Gustafsson (2009) has also presented results based on Swedish PIRLS 2001 data for reading comprehension, where Sweden participated with students from both school years 3 and 4. Gustafsson also finds that the school year effect is basically twice as large as the age effect.

Even though extensive research indicates that the effect of school year is greater than that of age, it is probable that this relationship is not necessarily constant over all school years, or indeed constant between different skills/school subjects. Moreover, it is also clear (see Luyten 2006) that these effects can vary between different countries. The fact that the relative sizes of the effects appear to vary between different countries will be discussed later, after the method has been explained in more detail. In Chapter 4 additional research in this area which can be regarded as an extension of the basic model is discussed.

5. This research paper has now been published, see Cliffordson, 2010.
6. The eight countries Luyten analysed are Cyprus, England, Greece, Iceland, Japan, Norway, Scotland and Singapore.

Method

The dominant method according to the research for separating *the age effect and school year effect* is, as mentioned the "*Between grade regression discontinuity design*". The method is based on estimating a regression line for performance based on month of birth in a given year. By estimating this regression line for two consecutive years with a dichotomous (0 or 1) variable for school year, the difference that occurs between the years (effect of the dichotomous variable) is interpreted as the school year effect, whilst the gradient of the coefficient of the regression line can be interpreted as the age effect.

The most important underlying assumptions that have to be satisfied for the method to work are:

1. That students start school in accordance with a strict date rule, for instance as in Sweden where all children born in the same calendar year usually start school in the same year.
2. That the relationship between performance and age is approximately linear and thus equally large for the two consecutive school years.

If an increasing proportion of children start school on the basis of how mature or "gifted" they are, the relationship between age and performance will become somewhat blurred, as many relatively more "gifted" students start earlier, and thus appear as young in their school year, whilst those who are less "gifted" start later (or repeat a year) appear to be older in their school year. In order to study the effects of age and school year optimally, all children would be allocated to a school year strictly on the basis of their year of birth/date.

In the case of Sweden, it is, of course, possible that children start earlier or later, but as this is relatively uncommon (approximately 95–96 percent start in accordance with their calendar year), these students, who generally are unusually strong/weak, do not affect the estimates in any significant way.[7]

The assumption that the relationship is linear has been tested by Cliffordson (2008) and based on TIMSS 1995 data, she finds that the assumption is valid for Sweden. From a theoretical perspective, there is good reason to question the assumption of a linear relationship between performance and age. Particularly when studying development of knowledge over a number of years. However, in this context, what is important is that the relationship is essentially linear across two consecutive school years.

The assumption that at least 95 percent of an annual cohort follow a specific date rule is only true for some countries, Sweden being one of these.[8] Martin,

7. A rule of thumb proposed by Luyten (2006) is that at least 95 percent of the students follow the rule of starting school based on their age.
8. If school start is strictly in accordance with age, the age difference amounts to 1 year within the cohort. We take this up in the analyses later on.

Mullis and Foy (Martin et al 2008) studied all countries that participated in PIRLS 2006, and show that out of the 40 participating countries and provinces, 10 in fact fulfil this requirement. [9]

An additional requirement that must be fulfilled for analysing a country in terms of the effects of age and school year, is of course that such data must exist for two consecutive school years. This serves to limit the number of countries still further, and as regards PIRLS 2006, only Iceland and Norway participated in both school years 4 and 5. Sweden took part in PIRLS 2001 for both school years 3 and 4.

9. The countries which according to PIRLS 2006 have at least 95 percent of the students following a strict date rule are in descending order: Iceland, Norway, England, Poland, Taiwan, British Columbia (Canada), Singapore, Ontario (Canada), Slovenia and Sweden.

Results

PIRLS 2001, Sweden Years 3 and 4

To illustrate the method, all students in school years 3 and 4 who participated in PIRLS 2001 were grouped by date of birth (month). The students who are "under-age" and "over-age" have been assigned to their own groups. Figures 1 and 2 show students' average performance on the PIRLS test by month of birth. Younger students are on the left on the horizontal axis, and older ones on the right.

It can be seen from the figures that within each cohort there is a relationship between age in months and average performance, older students tend to perform better on average. But this relationship doesn't hold for the under-age and over-age students, as the former perform significantly better than other young students in their school year, and "over-age" students perform significantly worse than all other groups in the school year.

These figures show that if under-age and over-age students are not excluded when the age effect in a school year is estimated, the relationship between age and results will be more or less blurred or underestimated. The first step in the following analysis is thus to exclude both under-age and over-age students, which is also the praxis in other studies. Excluding these students is also entirely logical as they do not belong to the cohort being investigated i.e. the population we wish to study.

Note that even if over-age and under-age students are excluded in these analyses, the remaining students are also affected as students who belong to the cohort are missing, and are in either a higher or lower school year. The students who belong to the cohort but attend a higher school year (under-age) tend to be born earlier in the year, and the students attending a lower school year (over-age) tend to be born later in the year. This in combination with the fact that under-age students tend to perform somewhat better than other students born in the same month, and that over-age students tend to perform somewhat worse than others born in the same month, means that these students are not representative of the rest of the cohort. The consequence is that the average result for normal-age students in the early birth months is somewhat underestimated, and that the average result for students born in the later months is overestimated, compared to all students from the same cohort in the same school year.

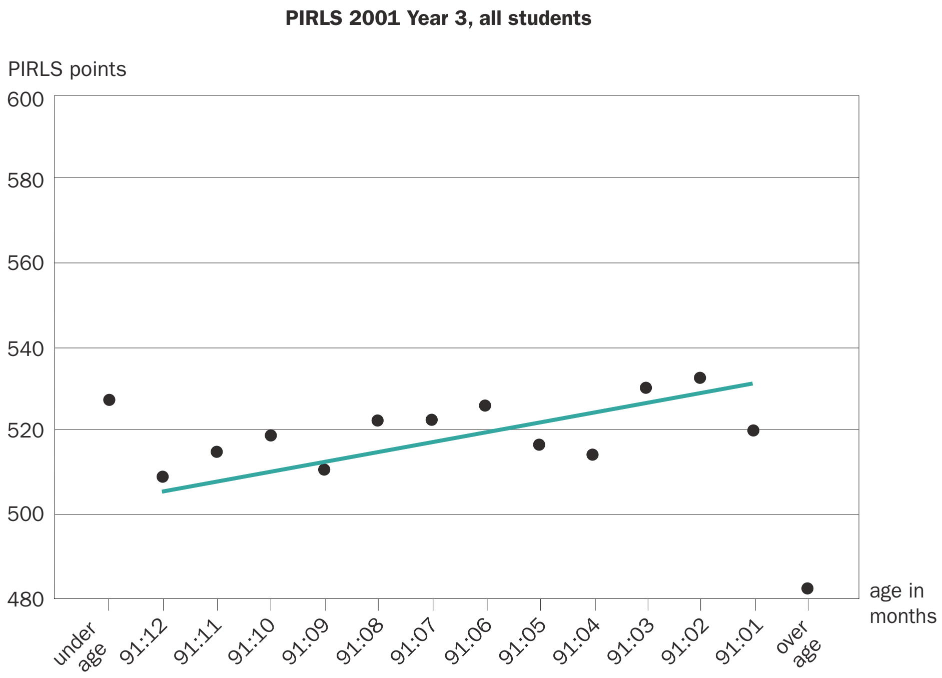

Figure 1. The relationship between average results and age in months as per PIRLS 2001, school year 3, Sweden. Each point represents the average for students born in the same month, apart from those below and above the standard age. The regression line (red) is an approximation and only based on “normal-age” students, i.e. no account is taken of “under-age” and “over-age” students.

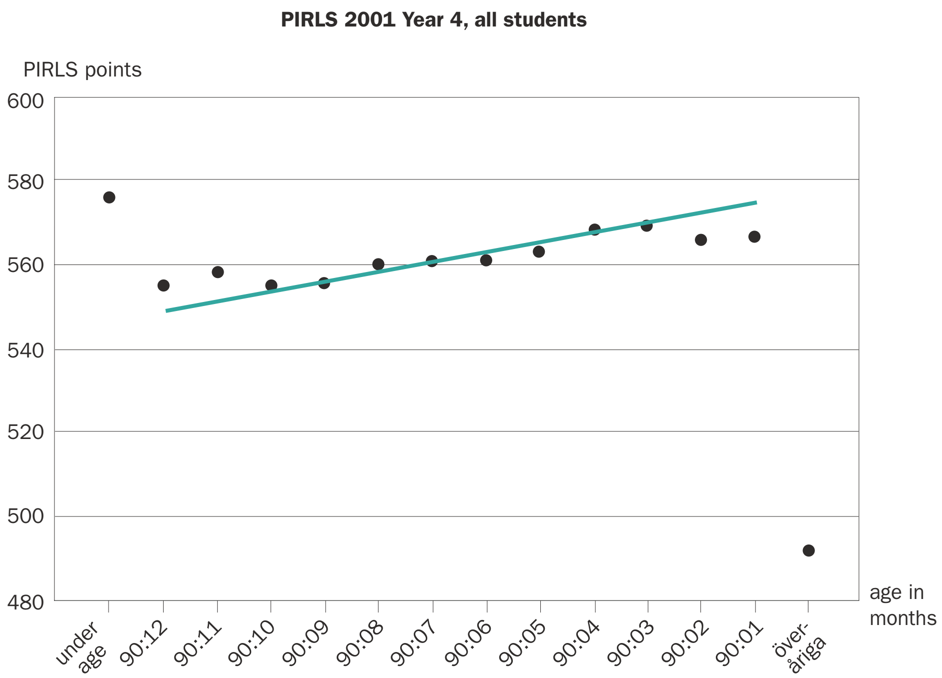

Figure 2. The relationship between average results and age in months in PIRLS 2001, school year 4, Sweden. Each point represents the average for students born in the same month, apart from those above and below the standard age. The regression line (red) is an approximation and only based on “normal-age” students, i.e. no account taken of “under-age” and “over-age” students.

This is shown in Figure 3. Normal-age students born in 1991 (light green ellipsis) attend school year 3 (rectangle). School year 3 also includes students born in 1992 and 1990, which can be seen in the school year 3 rectangle in the green and black segments. These students were excluded when the method was applied. But students born in 1991, who for some reason started school earlier or later are in school year 2 and school year 4 rectangles (light green segments).

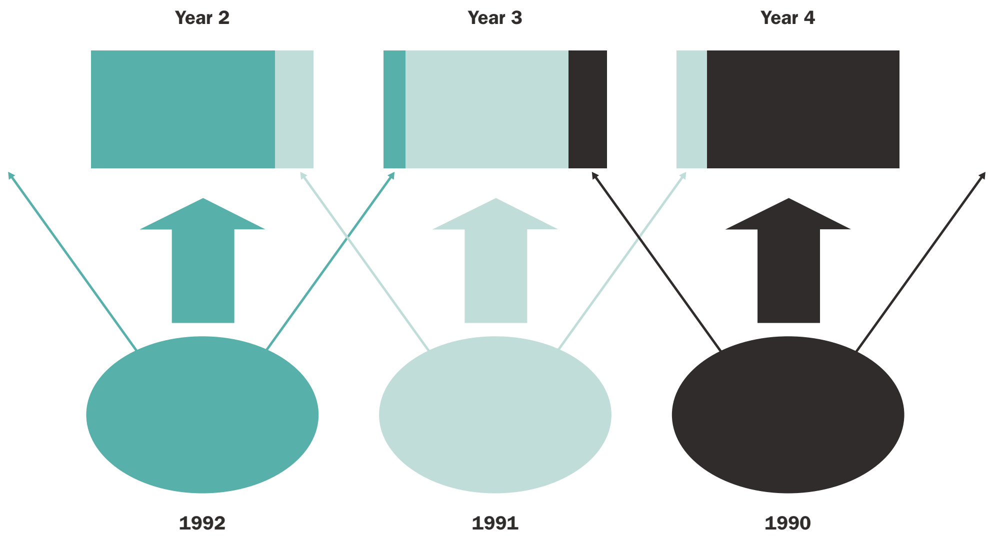

Figure 3. Distribution of birth cohorts between different school years. The circles and colours represent all students born in the same year, whilst the rectangles represent all students attending the same school year.

The consequence of this is that the relationships of the age effect which we estimate using students of normal age will be somewhat underestimated. For this reason, it is important that the proportion of under and over-age students is not excessively large, hence the rule of thumb that they should not exceed 5 percent of all students.[10]

Table 1 shows average results for different school years by normal-age, under-age, and over-age students. It can also be seen here that level of performance is generally higher by approximately 42 points for normal-age school year 4 students compared to normal-age school year 3 students..

10. Note that a rule of thumb of 5 percent refers to the proportion of students excluded from different school years. These students were not members of the cohort (students born same year). On the other hand, there are other students who are members of the cohort but attend other school years. Optimally, it is the size of this latter group, as a proportion of the whole cohort, which represents the true degree of exclusion. However, since data from other school years prior to and following the actual school year is rarely available, the relative size of this group is unknown.

Table 1. Descriptive results for PIRLS 2001 students in school years 3 and 4.

Grade	Student type	Number students	Total houwgt*	Proportion %	Mean PVread**	Standard deviation
3	Under-age	181	139	2.6 %	527.3	71.3
	Normal-age	4 987	5 040	95.6 %	520.2	76.9
	Over-age	103	92	1.7 %	482.0	88.0
4	Under-age	170	117	1.9 %	576.0	63.7
	Normal-age	5 778	5 830	96.5 %	561.9	64.9
	Over-age	95	94	1.6 %	491.6	79.8

* houwgt = houseweight and is a weight available in the international databases. "Houseweight" corresponds to the most common weight "total student weight" with the difference that it does not "inflate" the number of observations to the population size (as total student weight does) but is multiplied by a constant so that the original sample size is maintained.

** PVread = measure of results in reading comprehension obtained as plausible values. Based on student responses to the test, their skills were estimated in terms of a probability distribution from which 5 plausible values were obtained.

The fact that performance level is generally higher for school year 4 can also be seen in Figure 4 where all normal-age students from both school years 3 and 4 are brought together.

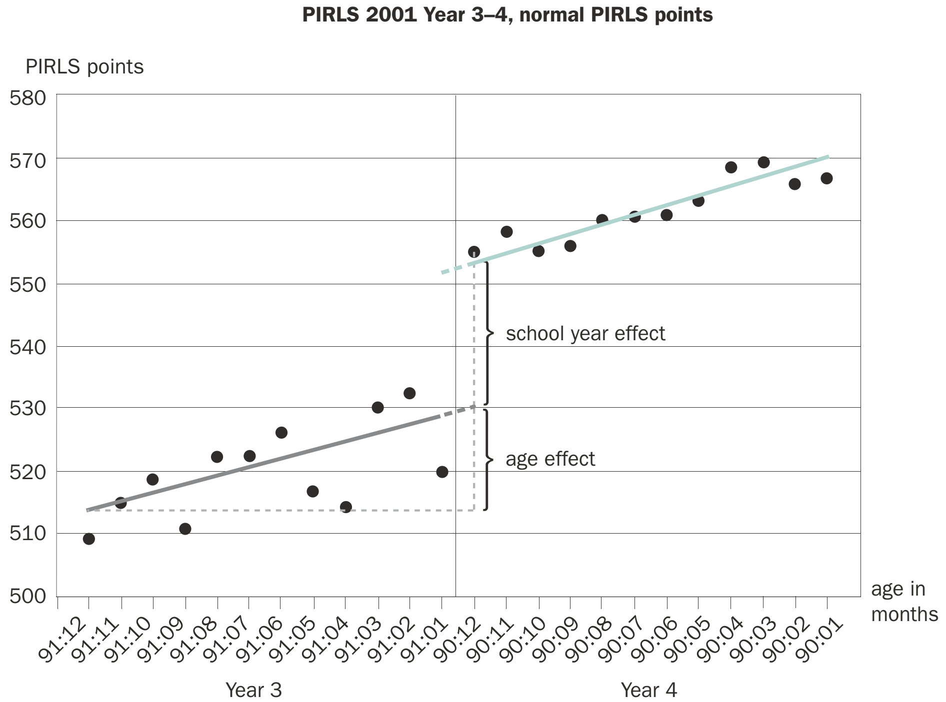

Figure 4. School years 3 and 4, average performance of students by month of birth (only normal-age students). The regression lines in the figure are approximations of the estimates and only intended to illustrate the concept.

Figure 4 clearly shows that there is a difference in average levels between school year 3 and 4 students. But if we only calculate the average difference between the two school years which we could define as the *total effect* and which

amounts to about 42 points, then we do not know how much of this consists of the *age effect* and the *school year effect*. [11]

In order to separate the effects of age and school year, a regression line was estimated with the same slope for both school years, but where the intercept (intersection with the y axis in the figure or the constant in the regression equation) is permitted to be different for the two school years. Pedagogically the school year effect can be observed by extending the regression line for school year 3 to the first month for school year 4, and then calculating the vertical distance between the lines (see figure). Mathematically the school year effect can be calculated as the difference between the intercepts for the two regression lines. Statistically the school year effect is estimated by introducing into the regression model a dichotomous variable for school year, where the effect of the dichotomous variable represents the school year effect.

The effect of age is shown in the figure by the coefficient for the regression line, which is estimated by calculating the change in points (the vertical change in the line) over a whole year, i.e. 12 months (see figure).

In reality the effects of age and school year are not calculated based on average values in different months of birth, but directly in a regression model at the student level, i.e. by using data at the student level for students' age (usually in years and months) and also school year. The regression model thus looks as follows:

PVread = b0 + b1 * (*age*) + b2 * (*school year*)

Where *PVread* = result in reading comprehension in terms of PIRLS (expressed as a plausible value)

age = variable for age in terms of months, and has been recoded so that the youngest students have value 1.

school year = dichotomous variable for school year where students from the lower school year have the value 0, and students from the higher year the value 1.

b0 = constant (intercept for the lower school year) to be estimated by the model.

b1 = age effect expressed as an effect per month. This effect was multiplied by 12 in order to derive an age effect for a whole year.

b2 = school year effect (expressed as an effect per school year)

Note that all under- and over-age students in both school years 3 and 4 were excluded before the estimates were made. Before the consolidated model is estimated, simple separate regressions should be estimated for the two school years for age and results. Thereafter the estimated regression coefficients are compared (effects of age in different years) with each other, to see if they are significantly different. This has been done and the age effect in different school years shows no

11. The total effect corresponds to the average differences between the two "normal-age" cohorts, which means that the older cohort is on average one calendar year older, and has also been attending school for a year longer. The total effect thus includes both the effect of students in the higher school year having attended school one additional year (school year effect), and in addition the effect of having an additional year of maturity (age effect).

significant differences, and this means that the model can be estimated by using a common age effect for both school years.[12]

Estimated regression model[13]:

PVread = 495.8 + 1.30 * (*age*) + 26.3 * (*school year*) (age effect per month)

PVread = 495.8 + 15.6 * (*age*) + 26.3 * (*school year*) (age effect per year)

Both the age effect and school year effect are statistically significant with t-values of 5.8 and 7.2.14

Table 1 shows that the total effect, which represents the average difference in results for the two school years (normal-age) is 41.7 points (see Table 1). The results from the regression estimates above show that this difference is made up of an age effect of 15.6 points and a school year effect of 26.3 points.[15] This means that the net effect of being one whole year older i.e. when the number of years in the school is held constant, is approximately 16 points on average. The average net effect of an extra year in school, i.e. after holding student age constant, is approximately 26 points. The school year effect is significantly greater than the age effect, and it accounts for approximately 63 percent of the total effect.

The first coefficient in the estimated regression equation, b_0, estimated as 495.8 can be interpreted as the expected value for a student who has a value 0 in age and 0 as the value of the dichotomous variable. In this specific case, it refers to the expected value for a student born in January 1992, i.e. one month later than the youngest students in the data material and attending school year 3. But in practice it has no interpretive value apart from being included in the equation for determining the expected value for a given student.

These estimates of the effect of age in relation to school year are identical to the estimates obtained by Gustafsson (see Djurfeldt & Barmark 2009). On the other hand, the t-values differ marginally, which is probably due to the use of different statistical software: M-plus and IDB Analyzer.[16]

To further extend the analysis, separate regression equations were estimated for girls and boys. Compared to girls, the age effect for boys is somewhat larger, whilst the school year effect is somewhat lower. However, the differences are relatively small and not statistically significant.[17]

---

12. Two separate simple regression equations were estimated, one for each school year. The effect of age (expressed per month) is estimated as 1.24 for school year 3 and 1.34 for school year 4. The standard errors for different effects were 0.34 and 0.28 respectively. The standard error for the difference 0.1 is approximated to 0.44. The difference is thus far from significant
13. The software used is IDB Analyzer, which partly takes account of cluster effects, or expressed differently that students from the same school are not entirely independent of each other, and that performance is expressed in plausible values. The model uses "houseweight".
14. As a rule of thumb t-values of 2 or greater indicate that the coefficient is statistically significant.
15. The discrepancy of 0.2 points ($15.6 + 26.3 = 41.9 \neq 41.7$) is probably because the statistical programme (IDB Analyzer) uses different procedures for estimating regressions and mean values. But this is no more than a guess.
16. Gustafsson's t-value for the age effect is 5.6, comparable to 5.8 in this replication. For the school year effect, Gustafsson obtains a t-value of 8.2, comparable to 7.2 in the replication.
17. For girls, the age effect amounted to 15.2 and for boys 18.2. The effect of school year was measured as 27.4 for girls and 23.5 for boys.

PIRLS 2006 Iceland and Norway, school years 4 and 5

Table 2 presents some descriptive data for students in school years 4 and 5 who took part in PIRLS 2006 in Norway and Iceland.

Table 2 Descriptive picture of data for Iceland and Norway in PIRLS 2006.

(Grade)	Student type	Number students	Total houwgt	Proportion in school year	Mean PVread	Standard-deviation
Iceland						
4	Normal-age	3 658	3 657	99.6 %	510.5	68.0
5	Normal-age	1 375	1 371	99.6 %	549.1	63.5
Norway						
4	Normal-age	3 812	3 815	99.3 %	498.1	66.6
5	Normal-age	1 796	1 793	99.3 %	540.4	64.6

Estimated regression models[18]:

Iceland:

PVread = 495.3 + 28.0 * (*age*) + 10.3 * (*school year*) (age effect in years)

t-values for estimated coefficients: 8.4 and 1.8

Norway:

PVread = 481.2 + 30.1 * (*age*) + 12.1 * (*school year*) (age effect in years)

t-values for estimated coefficients: 4.6 and 1.96

The results presented by Martin et al. are identical.

Table 3. Comparison of results for Sweden, Iceland and Norway

Study	Country	Subject and school year	Age effect	School year effect	Coverage of birth cohort
PIRLS 2001	Sweden	Read 3–4	15.6 (2.6)	26.3 (3.6)	96 %
PIRLS 2006	Iceland	Read 4–5	28.0 (3.5)	10.3 (5.3)	99.6 %
PIRLS 2006	Norway	Read 4–5	30.2 (6.1)	12.1 (7.2)	99.3 %

Standard error in parenthesis

Table 3 summarises the results for Sweden, Iceland and Norway. The dominant pattern is that the proportions between the age and school year effect are more or less reversed. Both Iceland and Norway exhibit significantly greater age effects and at the same time significantly lower school year effects compared to Sweden. The total effect (total of age and school year effect), however, is approximately just as large for each country.

18. The software used is IDB Analyzer, which partly takes account of cluster effects or expressed differently that students from the same school are not entirely independent of each other, and that performance is measured by plausible values. In the model the weight "totwgt" has been used.

Unfortunately there is more than one variable which differs between Sweden on the one hand, and Iceland and Norway on the other. For Sweden, the results apply to school years 3–4, whilst for Iceland and Norway, they apply to school year 4–5. Possibly the age effect could have been expected to be largest in the early years, and then become progressively less important? Or does the transition from "low" to "intermediate" school imply a more substantial increase in level? Another factor that differs is the time period when the studies were done. The results for Sweden apply to 2001, whilst for Iceland and Norway they apply to 2006. It is difficult to assess how this may have affected the estimates.

A final factor that differs and which could explain some of the differences shown, is the degree of coverage of the cohort, namely how large a proportion of the birth cohort are of normal age, and thus included in the sample. Sweden is at a level of 96 percent, whilst Iceland and Norway are very close to 100 percent.19 In Chapter 2 the argument was put forward that the more students who were excluded (lower coverage), the more the age effect can be expected to have been underestimated, even though under-age and over-age students were excluded. It is reasonable to assume that the students excluded have some impact on the results, but it seems unlikely that this would change the results in any significant way.

19. In reality these figures are merely an approximation of the coverage as the proportions represent the proportion of normal-age students in the school year. This is not the same thing as the proportion of the birth cohort, but can be regarded as an approximation. The possible underestimate of the age effect is not due to the exclusion of over-age and under-age students from the school year (they were born a different year), but because there is a lack of over- and under-age students from the birth cohort attending the second school year.

TIMSS 1995, Sweden, school years 6, 7 and 8

Sweden took part in TIMSS 1995 with students from three school years, 6, 7 and 8. Below, the results of the estimates are presented, using exactly the same principle as earlier in the document and thus represent an attempt to replicate Cliffordson. Since the results from TIMSS 1995 have been subsequently rescaled, the effects have been estimated for data based on both the original scale and the new scale.[20]

Table 4a TIMSS 1995, Sweden, original scale

Study	Subject and school year	Age effect	School year effect	Comments
Mathematics				
TIMSS 1995	Maths Year 6–7	13.1 (3.8)	28.4 (5.8)	Original scale
TIMSS 1995	Maths Year 7–8	16.8 (4.4)	17.5 (6.2)	Original scale
Natural sciences				
TIMSS 1995	Sci Year 6–7	11.5 (3.8)	34.4 (5.9)	Original scale
TIMSS 1995	Sci Year 7–8	21.1 (5.6)	13.6 (6.8)	Original scale

Standard error in parenthesis

Table 4b. TIMSS 1995, Sweden, new scale

Study	Subject and school year	Age effect	School year effect	Comments
Mathematics				
TIMSS 1995	Maths Year 6–7	13.7 (3.2)	24.0 (5.4)	New scale
TIMSS 1995	Maths Year 7–8	18.0 (3.8)	8.4 (5.9)	New scale
Natural sciences				
TIMSS 1995	Sci Year 6–7	13.0 (4.1)	26.7 (6.5)	New scale
TIMSS 1995	Sci Year 7–8	20.9 (4.4)	8.9 (5.8)	New scale

Standard error in parenthesis

An interesting pattern is that for school year 6–7, the school year effect is greater than the age effect, but for school year 7–8 the age effect is greater. A possible explanation for this result could well be that there is a larger academic leap when transferring from school year 6 to 7 (i.e. from primary to lower secondary school) compared to school year 7 to 8. At the same time the reasoning earlier was that the age effect could be expected to decline in later years, which does not appear to be the case in these estimates. It should also be pointed out that the differences between the estimates lie within the confidence interval and are not significant.

Another interesting result is that the estimates are not entirely independent of how the results are scaled. What this is related to is not known and lies outside the scope of this document, but is obviously something that should be further studied, and is an area for further research.

Compared with Cliffordson, the general picture of the results is similar, but not identical. See the appendix for a detailed review of the replication of Cliffordson.

20. Cliffordson has only used the original scale in her analyses. The difference in scaling does not mean that the nominal values are at entirely different levels, but rather that the statistical method used to estimate student performance on the basis of responses to test questions differs. The original scale is based on a 1-parameter model whilst the new scale is based on a 3-parameter model. See IEA 2004.

TIMSS Advanced 2008, Sweden, school year 11–12

In TIMSS Advanced 2008 Sweden added a supplementary sample of students from school year 11 (second year in upper secondary school), in addition to the conventional school year 12 sample (year 3 in upper secondary school). It should be noted that there are some obvious problems in attempting to estimate the age- and school year effect in this material as students have not necessarily studied the same courses, even though they attend the same school year. Yet it could still be interesting to see how the results of these estimates come out using this method.

A quick glance at the descriptive results shows that they contain some unusual features. (See Table 5).

Table 5. Average results in mathematics for students in school years 11 and 12 who participated in TIMSS Advanced 2008.

School year	Number of students in data	PVma	Standard error
Year 11	1 058	411.4	6.2
Year 12	2 043	414.0	5.3

The difference between results in mathematics for school year 11 and 12 students is marginal. Since the total effect is virtually non-existent, neither the effect of age nor school year will probably be other than marginal and hardly significant.

$$PVma = 421 - \underset{(10.3)}{3.8} * (age) + \underset{(13.5)}{6.4} * (school\ year)$$

(age effects expressed per year)
(standard error in parenthesis)

Neither the age effect, nor that of school year, is significant. The reason for the absence of any effect is probably due to the fact that a fairly large proportion (about 40 percent) of school year 12 students did not study mathematics in the last school year and have probably "lost" knowledge rather than gained new knowledge.[21] This in its turn is partly due to the fact that Mathematics E is not a compulsory course. Unfortunately data is not specifically available to identify which school year 11 students intend to study Mathematics E in year 12. With this data, the analysis could have been performed only on students in school years 11 and 12 who either studied Mathematics E or intended to study it in school year 12. This would have made it possible to estimate a "forgotten" effect by comparing results for only those students who studied up to Mathematics D. With the data available, the analysis regrettably does not contribute to our knowledge of the real relationship between the age effect and school year effect in higher school years.

21. See National Agency for Education 2009

ICCS 2009, Sweden, year 8–9

In the international study ICCS 2009, which measures student knowledge in democracy and civics, a supplementary sample was used so that students from both school years 8 and 9 took part in the study. The descriptive results are shown in Table 6.

Table 6. Average results in the Civic and Citizenship Education Study for students in school years 8 and 9 who participated in ICCS 2009, Sweden.

School year	Number of students in data (normal-age)	Excluded (under- and over-age)	PVciv (results in points)	Standard error
Year 8	3 183	6.4 %	541	3.0
Year 9	3 250	6.0 %	578	3.6

The average difference in results between school years 8 and 9 is 37 points. Note that the proportion excluded, which represents a measure of non-response, amounts to approximately 6 percent which is somewhat higher than the rule of thumb figure of 5 percent.

Estimated regression:

$$PVciv = \underset{}{311} + \underset{(5.6)}{15.6} * (age) + \underset{(6.7)}{21.3} * (school\ year)$$

(age effect expressed per year[22])
(standard error in parenthesis)

The total effect of 37 points can be divided into an age effect of approximately 16 points and a school year effect of approximately 21 points. These results are thus consistent with other results from Sweden concerning other knowledge domains, namely that the school year effect is somewhat greater than the age effect.

22. The variable for age has not been standardised with the youngest students assigned age 0, but their real age is coded in years. The constant of 311 points is thus the model's theoretical estimate of what a newly born person would achieve, but obviously in this context has no sensible meaning.

Problems with the model and its extension

Weighting

One problem observed in connection with replicating these studies is that the researchers have generally used "houseweight", which is one of the weights available in the data set from the IEA Studies, TIMSS and PIRLS.[23] When using "houseweight" (houwgt), sample size is kept constant, which has the advantage of not exaggerating the standard error in SPSS. If "totwgt" were used instead, sample size would be increased in relation to population size, and then SPSS would regard the sample as much larger than it is, and thus underestimate the standard error.

However, SPSS should not be used directly in any case since the sample design for these studies is so complicated that using SPSS would provide incorrect estimates for standard errors. Neither should the estimate of a population's mean value be affected by whether "houwgt" or "totwgt" is used. The problem occurs, however, when we make use of students from two school years (as we do in these contexts) and the samples of the different school years are of different sizes. Since the number of students in two consecutive school years can be expected to be approximately equally large, they should also be given essentially the same weight when regression parameters are estimated. This is also the case where "totwgt" is used, even though the sample is not as large. With "totwgt" the sample is inflated to the size of the population, that is the size of the population which the sample is intended to represent. But with "houwgt" the sample size is retained and if the sample is relatively smaller for one school year, it will be given relatively less weight in the regression.

Because of this, "totwgt" should be used. However, in this report the weight used by the researchers has also been used, usually "houwgt" as the aim has been to try to replicate results from the researchers who initially carried out the analyses. For Iceland and Norway in PIRLS 2006, the regressions on the other hand in this study have used "totwgt" since the researchers who originally carried out these analysis also used "totwgt", which provides slightly different results compared to using "houwgt", but these differences are marginal.[24] The sample sizes in Norway were 3 812 for school year 4, and 1 796 for school year 5. For Iceland the corresponding sample sizes were 3 658 and 1 371, thus significant asymmetry existed in the samples in relation to population. The fact that the results are nevertheless stable can be interpreted as signifying that the age effect is approximately of the same size for both school years (i.e. the regression for age in different school years), so that it is not important if one of the school years is given greater weight. In purely theoretical terms, however, the author's view is that it would be more correct to use "totwgt".

23. The exception is the study by Mullis using PIRLS 2006 data from Iceland and Norway where "totwgt" was used.
24. In Iceland the age effect decreased by about 1 point and the school year effect also by about 1 point. In Norway, however, the age effect increased by about 2 points, whilst the school year effect decreased by about 2 points.

Scaling

For TIMSS 1995, it is evident that the scaling of the performance variable (the plausible values) is of some importance for the estimates. The causes of this could be studied in greater detail and it is an area for further research. One possible approach would be to repeat the TIMSS 1995 analyses for a number of countries but using a different scaling, possibly taking Luyten's (2006) study as a starting point to determine if the deviations are general.

Multi-collinearity

Multi-collinearity is where two explanatory variables are closely correlated with each other. This can lead to difficulties in separating the effects of different variables. One indication that multi-collinearity is a potential problem is when the estimated model has a relatively good fit (high R^2) whilst the individual regression coefficients are not significant. It is fairly evident that the two independent variables, age and school year, cannot be regarded as independent. For PIRLS 2001, Sweden, the correlation amounts to 0.87. But in nearly all cases the estimates in this report are significant with good margins. Multi-collinearity can be a major problem where two independent variables provide a measure for similar variables. However, since this is not the case here, multi-collinearity is not considered to be a problem in estimates from this model.[25]

Standardised coefficients

Based on the empirical observation that the students who are not "normal-age" tend to be born either early or late in the year, the age distribution for normal-age students in a school year will not be perfectly evenly distributed. This could potentially have an impact on the estimates, and one means of compensating for this would be to standardise the coefficient for age.

Multi-level approach

Luyten (2006) (see research review in the introduction) uses a multi-level model which means that he is able to estimate variation in the school year effect between schools in different countries. He finds that the variation between schools is substantial, but that the variation is of varying magnitudes in different countries. He also finds that in some countries the school year effect is greater in schools where students have a more advantageous socio-economic background. It would, of course, be interesting to apply a similar analysis to Swedish data

Luyten has also together with a number of other researchers (Luyten et. al., 2008) carried out a similar multi-level analysis on data from PISA 2000 for England concerning reading comprehension, reading involvement and reading attitudes in school years 10 and 11. They find just one significant school year effect for reading comprehension, and in addition a relatively weak effect but nevertheless still greater than the age effect. They also find in contrast to Luyten (2006), that the school year effect is weaker in schools where students have a more advanta-

25. See pages 276–277 in Gustafsson (2009) for a discussion on this.

geous socio-economic background. What makes this study particularly interesting is that it is the only one, as far as the author is aware, that has estimated the effects of age and school year at higher age levels in the school. Since the analysis is specific to reading comprehension, it is perhaps not entirely surprising that both the effects of school year and age are relatively small compared to earlier school years. Yet it is not clear that this is also the case for e.g. mathematics.

Concluding discussion

The above replications and reporting shows how it is possible to separate the effects of age and school year in data covering students from two consecutive school years. It also requires that information be available on students' date of birth, and that differences regarding the age at which they start school are not excessive.

In general, the results show that the school year effect is often greater than the age effect, but that there are exceptions. In addition, estimates of the effects vary greatly between different studies and between different countries.[26] This makes it relatively risky to argue that there is a general measure for school year- and age effects that can be used to adjust the results from International studies in order to make comparisons between results more accurate, which could otherwise be regarded as the main purpose of these analyses. The analyses show, however, that due to the large amount of variation in the results, this cannot be easily done.

In an appendix to one of the Government commissions carried out by the National Agency for Education, Gustafsson (see Gustafsson 2009) used for mathematics a standardised measure for age and school year, of 0.17 and 0.24 respectively, where the school year effect is greater than that for age. He uses 0.17 and 0.28 in the natural sciences for the effect of age and school year respectively. In reading comprehension, the effects of age and school year for younger students (years 3–4) were set to 0.20 and 0.33, and for older students (year 8–9) 0 and 0.20. Based on these effects, results were then subsequently corrected for the different countries. For Sweden the consequence is that the results are somewhat "better" in the PISA studies which were based on an age sample, whilst the results in TIMSS and PIRLS were somewhat "worse" as these were based on a school year sample. The qualitative shift (i.e. the direction of the correction) is fully consistent with expectations. However, the quantitative size of the correction depends entirely on the assumptions made by Gustafsson regarding the size of the age and school year effects.

It is important to point out that these corrections have no impact on trends in e.g. the PISA studies. This is because the correction factor is just as large for all the PISA studies. On the other hand interpretation of overall knowledge trends in Sweden is affected during the period from 1960 and the 1970s compared with today: This is because the first IEA studies at that time used an age-based sample before going over to a school year sample. The consequence was that Sweden's results in these early studies were better than would have been the case if the results had not been corrected, at least from a relative time perspective.[27] This correction of the results may thus be justified in this specific case, as comparisons are made between Sweden's relative results from the early 1960s and up to the present time, since the basis on which comparisons were made changed during later decades.

26. See in particular Luyten (2006) who analyses 8 different countries using TIMSS 1995 data for school year 3–4.
27. The correction factor has different signs (i.e. goes in opposite directions) for the early IEA studies compared to the later IEA studies.

Based on the wide variation in results from different studies, the author's view is that carrying out such corrections is not recommended in public contexts, such as e.g. routinely reporting corrections when publishing international results. The problem is also that even if age and school year effects for Sweden can be estimated with an acceptable level of accuracy, the same effect must be applied to all countries if results are to be comparable. It can also be argued that the effects are not necessarily of the same magnitude in different countries, partly because of how students are assigned to different school years when starting school, or later during their schooling. But since it is not possible to estimate these effects for each individual country, there is a risk that these corrections could distort the accuracy of the results. However, it is evident that further research in the area is required before these results can be used in connection with international studies.

References

Cahan, S., Cohen, N. 1989. *Age versus schooling effects on intelligence development*. Child Development, 60, 1239–1249.

Cahan, S. & Davies, D. 1987. *A Between-Grades-Level Approach to the Investigation of the Absolute Effects of Schooling on Achievement*, American Educational Research Journal, 24(1), 1–12.

Cliffordson C, 2010. *Methodological issues in investigations of the relative effects of schooling and age on school performance: the between-grade regression discontinuity design applied to Swedish TIMSS 95 data*. Educational research and Evaluation, Vol. 16, No. 1, February 2010, pp. 39–52.

Cliffordson C, Gustafsson J-E, 2008. *Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling.* Intelligence 36 (2008), pp. 143–152.

Gustafsson J-E, 2009. *Kunskaper och färdigheter I grundskolan under 40 år: En kritisk granskning av resultat från internationella jämförande studier* (Knowledge and skills in compulsory school over 40 years: A critical examination of results from international comparative studies). Bilaga till Skolverkets svar på regeringsuppdrag. Skolverkets bild av utvecklingen av kunskapsresultaten i grundskolan och av elevers studiemiljö – redovisning av uppdrag att utarbeta ett sammanfattande underlag avseende utvecklingen av kunskapsresultaten i grundskolan (2009) Dnr 2008:3010. (The Agency's picture of the development of knowledge results in compulsory school and pupils' study environment – Report of commission on drawing up summary material concerning the development of knowledge results in the compulsory school (2009) Dnr 2008:3010.)

Gustafsson J-E, 2009. *Strukturella ekavtionsmodeller.* kapitel 10 i Statistisk verktygslåda 2 – multivariat analys (Structural equation models. Chapter 10 in Statistical Toolbox 2 – multivariate analysis (Ed:) Djurfeldt G, Barmark M). Studentlitteratur, 2009.

IEA 2004. TIMSS 2003 Technical Report.

Luyten H, 2006. *An empirical assessment of the absolute effect of schooling: regression-discontinuity applied to TIMSS-95*. Oxford Review of Education, Vol. 32, No. 3, July 2006, pp. 397–429.

Luyten H., Peschar J. & Coe R., 2008. *Effects on Schooling on Reading Performance, Reading Engagement and Reading Activities of 15-Year-Olds in England*. American Educational Research Journal, Vol. 45, No. 2, pp. 319–342.

Luyten H, Veldkamp B, 2008 *Assessing the effect of schooling with cross-sectional data: Between grades differences addressed as a selection-bias problem*. Conference paper, presented at the IEA Conference in Taipei September 2008.

Martin M, Mullis I, Foy P, 2008. *Interrelationships among Reading Achievement, Grade level and Age in PIRLS 2006*. Conference paper, presented at the IEA Conference in Taipei September 2008.

Skolverket, 2004. *Internationella studier under 40 år – svenska resultat och erfarenheter* (International studies over 40 years – Swedish results and experiences). Skolverkets aktuella analyser, Beställningsnr: (Current Analyses of the National Agency for Education, Order number:) 04:843. National Agency for Education

Appendix 1

Effect of school year and age for PIRLS 2001, Sweden: documented data processing.

Processing

1. Based on the two data sets **[ASGSWER1.sav]**, containing school year 4 students and **[ASGSE3R1.sav]**, containing school year 3 students, a single combined data set was created and named as: **[ASGSWER1_arbetskopia med ak3–4.sav].**

   The data set contains 11 315 students with 5 271 students from school year 3 and 6 044 from school year 4.

2. Compute **[student type]** = **itbirthy**

3 a. select if **idgrade** =3, transform student type to: Accelerated = 3 (born 92)
Normal =2 (91)
Delayed =1 (89+90)

b. select if **idgrade** = 4, transform student type to: Accelerated = 3 (born 91)
Normal =2 (90)
Delayed =1 (89)

4. Create variable **[age]** = age in months, where 1 = Dec 92, …..48 = Jan89
By : compute **[age]** = **[itbirthm]** * (–1) + 13 if **[itbirthy]** = 92
+ 25 if " = 91
+ 37 if " = 90
+ 49 if " = 89

5. Create variable **[grade_dummy_34]** = 1 if **idgrade** = 4, och =0 if **idgrade** = 3

Results

A. descriptive

			Replication PIRLS 2001			
Grade	**Student type**	**Number students**	**Total houwgt**	**Proportion %**	**Mean PVma**	**S.D.**
3	Accelerated	181	139	2.6	527.3	71.3
	Normal	4 987	5 040	95.6	520.2	76.9
	Delayed	103	92	1.7	482.0	88.0
4	Accelerated	170	117	1,9	576.0	63.7
	Normal	5 778	5 830	96.5	561.9	64.9
	Delayed	170	94	1.6	491.6	79.8

B. Regressions

Carried out using IDB Analyzer with dependent variable: **[ASREA01-05]** and explanatory variables **[age]** and **[grade_dummy_34]**

Model: PVread = **495.8 + 1.30 * (age_{months}) + 26.31 * (grade)**
(t=5.8) (t=7.2)

which corresponds to: **15.6 * (age_{year})**

According to this estimate, the effect of school year (26.3) is significantly greater than that of age (15.6). The results are identical to those of Gustafsson (see Djurfeldt and Barmark, 2009).

Appendix 2

Replication of Cliffordson "Effects of schooling and age on performance in mathematics and science: A between-grade regression discontinuity design applied to Swedish TIMSS 1995 data" (working paper presented at the IEA conference in Taipei 2008).

In this part of the appendix, a study by Cliffordson is replicated using a "between grade w discontinuity design".

Data

TIMSS 1995 data for Sweden

Original data set: **[bsgswem1.sav]** (student file with background variables and plausible values in maths and science). The file contains 8 855 students from school year 6 (2 831 students), 7 (4 075 students) and 8 (1 949 students), which corresponds to Cliffordson.

An initial descriptive picture

STUDENT'S DATE OF BIRTH\YEAR

Table 1. Proportion students by year of birth

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	78	1	0.0	0.0	0.0
	79	60	0.7	0.7	0.7
	80	2 001	22.6	22.7	23.4
	81	3 995	45.1	45.3	68.7
	82	2 753	31.1	31.2	99.9
	83	11	0.1	0.1	100.0
	85	2	0.0	0.0	100.0
	Total	8 823	99.6	100.0	
Missing	System	32	0.4		
Total		8 855	100.0		

Variable: itbirthy, unweighted

As can be seen in Table 1, data is not available for 32 students. Cliffordson states that the oldest student was born in Jan 1979 and the youngest in Dec 1983. In this data set there were an additional 3 students (born 1978 and 1985). These three students were excluded from the analyses in order to achieve comparability with Cliffordson. 8 820 students thus remain.

In the next stage, the 8 820 students were assigned a value between 1 and 60 which identifies their age (variable **age**) in months where 1 = Dec 83 and 60 = Jan 79 (in the same way as in Cliff). The variables **itbirthy** (year of birth) and **itbirthm** (month of birth) were used for this purpose.

Then based on the variable **idgrade** (school year) in combination with **itbirthy** (year of birth) students were classified as “accelerated”, “normal” and “delayed” depending on whether they are part of the normal cohort for the school year or if they started later or earlier. The variable created is **student type** (where 1=accelerated, 2=normal, 3=delayed). All the computations were carried out using SPSS 15.0.0.

Average result in maths

Table 2 shows average results by school year and student type. The table should be compared with Table 1 in Cliffordson.

Table 2. Descriptive picture of student material with average results in mathematics by school year and student type

		Replication					Cliffordson	
Grade	**Student type**	**Number students**	**Total houwgt**	**Proportion %**	**Mean PVma**	**S.D.**	**Mean PVma**	**S.D.**
6	Accelerated	10	10	0.3	510.85	64.54	507.52	59.04
	Normal	2 724	2 783	96.7	479.56	75.97	479.12	75.82
	Delayed	82	85	3.0	421.01	76.86	422.38	80.97
7	Accelerated	30	21	0.7	560.30	70.06	553.87	66.50
	Normal	3 901	2 816	96.1	520.90	85.03	520.84	84.72
	Delayed	130	94	3.2	453.32	77.10	446.79	75.94
8	Accelerated	13	20	0.7	643.74	67.54	647.24	69.39
	Normal	1 872	2 898	96.4	555.52	89.99	555.86	89.25
	Delayed	58	89	3.0	479.33	93.85	475.10	103.44

Software: IDB analyzer, weight: houseweight (houwgt), Plausible values from the original scale for 1995.

Since there are so few students belonging to the group “accelerated” or “delayed”, there is no reason to be particularly concerned over the deviations with Cliffordson for these groups, as individuals can substantially affect the results and also because the number of students (expressed by total “houwgt”) does not really correspond for these groups. On the other hand, it appears that the number of normal students in each school year exactly matches Cliffordson (cannot be seen in this table, must be compared with Table 1 in Cliff), which should be regarded as evidence that we are using exactly the same students in the data set. The average result in mathematics for the normal groups in different school years also corresponds more or less exactly with Cliffordson even though differences exist at the decimal level.

Replication of regressions in mathematics

A regression model has been estimated based on normal students in school years 6 and 7 and a model based on normal students from school years 7 and 8.[28]

Regression model: **PVma** = **b0** + **b1** * **(AGE)** + **b2** * **(GRADE)** + e

Where **AGE** is age in months and **GRADE** is school year in years.

Table 3: Estimate of effects of age and school year for school years 6–7 and 7–8 in mathematics. All the regressions are based on “normal students”.

				b0	Coefficient B1			Coefficient B2			
	Grade	Birth month	N*	Constant	AGE (month)	s.e.	t-value	AGE (year)**	GRADE (year)	s.e.	t-value
Rep	6–7	Jan–Dec	6 625	458.96	1.09	0.32	3.4	13.08	28.35	5.81	4.9
Cliff	6–7	Jan–Dec	5 599	No data	1.12	0.32	3.6	13.44	28.43	5.28	5.4
Rep	7–8	Jan–Dec	5 773	477.94	1.40	0.37	3.8	16.8	17.51	6.19	2.8
Cliff	7–8	Jan–Dec	5 714	No data	1.34	0.39	3.5	16.07	18.67	6.99	2.7

*Cliffordson gives N as the weighted number of students, whilst the replication is based on the real number of students.
**This coefficient is only AGE (month) multiplied by 12 for comparability with GRADE.

Average results in natural sciences

Table 4. Estimate of effects of age and school year for school years 6–7 and 7–8 in natural sciences. All the regressions are based on “normal students”. Note. There is no table for comparison with Cliffordson who only gives the pattern of results in text format for the natural sciences.

		Replication					Cliffordson	
Grade	Student type	Number students	Total houwgt	Proportion %	Mean PVsci	S.D.	Mean PVsci	S.D.
6	Accelerated	10	10	0.3	489.1	49.6		
	Normal	2 724	2 783	96.7	490.2	84.8		
	Delayed	82	85	3.0	427.4	91.5		
7	Accelerated	30	21	0.7	574.0	69.8		
	Normal	3 901	2 816	96.1	537.0	88.9		
	Delayed	130	94	3.2	476.4	95.6		
8	Accelerated	13	20	0.7	602.4	81.4		
	Normal	1 872	2 898	96.4	572.1	92.9		
	Delayed	58	89	3.0	489.3	96.1		

28. The regressions were carried out in IDB Analyzer after the variables for grouping normal_67 and normal_78 were created with the value 1 if the student is not only a normal student but also belongs to school year 6 or 7 (and 7 or 8), otherwise missing.

Replication of regressions in natural science

Regression model: **PVsci** = **b0** + **b1 * (AGE)** + **b2 * (GRADE)** + e

Table 5. Estimate of effects of age and school year for school years 6–7 and 7–8 in natural sciences. All the regressions are based on “normal students”.

					Coefficient				Coefficient		
	Grade	Birth month	N*	Constant	AGE (month)	s.e.	t-value	AGE (year)**	GRADE (year)	s.e.	t-value
Rep	6–7	Jan–Dec	6 625	473.1	0.96	0.32	3.0	11,5	34.4	5.91	5.8
Cliff	6–7	Jan–Dec	5 599	No data							
Rep	7–8	Jan–Dec	5 773	483.2	1.76	0.47	3.8	21.1	13.6	6.79	2.0
Cliff	7–8	Jan–Dec	5 714	No data							

There are no exact figures for Cliffordson, but she states that the school year effect was greater in science for school year 6–7 regressions, but less than maths for the 7–8 regression, results which are verified in Table 5. It is interesting to note that the effect of school year is three times larger than the age effect for years 6–7, whilst the school year effect is less than the age effect for school year 7–8. For maths, the school year effect was twice as large as the age effect for school years 6–7, and for school year 7–8 the corresponding effects were approximately the same.

A possible explanation for the pattern where the effect of school years is greater between school year 6 and 7 is that a more significant academic leap takes place when students transfer from school year 6 to 7, particularly as they often change schools at the same time.

Corresponding analyses with rescaled data

When the results from TIMSS 1995 were released, a two parameter model was used to scale results in mathematics and natural sciences. Later international studies have used a three parameter model. To achieve comparability, results from TIMSS 1995 have been rescaled using a three parameter model.

In the tables below, the same analysis and rations have been carried out as in the earlier section, but this time using the new rescaled results in mathematics and natural sciences. Note that everything is the same apart from the plausible values in mathematics and natural sciences.

Table 6. Descriptive picture of student material with average results in mathematics by school year and student type. N.B. Rescaled plausible values (not comparable with Cliffordson).

		Replication					Cliffordson	
Grade	**Student type**	**Number students**	**Total houwgt**	**Proportion %**	**Mean PVma**	**S.D.**	**Mean PVma**	**S.D.**
6	Accelerated	10	10	0.3	502.8	54.8		
	Normal	2 724	2 783	96.7	477.7	69.4		
	Delayed	82	85	3.0	417.0	75.6		
7	Accelerated	30	21	0.7	552.4	56.8		
	Normal	3 901	2 816	96.1	515.2	69.0		
	Delayed	130	94	3.2	451.7	68.5		
8	Accelerated	13	20	0.7	605.4	49.1		
	Normal	1 872	2 898	96.4	541.9	74.9		
	Delayed	58	89	3.0	468.1	88.9		

Software: IDB analyzer, weight: houseweight (houwgt), Plausible values rescaled for 1995.

Table 7: Estimate of age- and school year effects for school years 6–7 and 7–8 in mathematics. All the regressions are based on "normal students". Note. Rescaled plausible values. Not comparable with Cliffordson.

					Coefficient			Coefficient			
	Grade	**Birth month**	**N***	**Constant**	**AGE (month)**	**s.e.**	**t-value**	**AGE (year)****	**GRADE (year)**	**s.e.**	**t-value**
Rep	6–7	Jan–Dec	6 625	456.2	1.14	0.27	4.3	13.7	24.0	5.45	4.4
Rep	7–8	Jan–Dec	5 773	469.3	1.50	0.32	4.7	18,0	8.4	5.88	1.43

Table 8. Descriptive picture of student material with average results in natural sciences by school year and student type. N.B. Rescaled plausible values (not comparable with Cliffordson).

		Replication					Cliffordson	
Grade	**Student type**	**Number students**	**Total houwgt**	**Proportion %**	**Mean PVma**	**S.D.**	**Mean PVma**	**S.D.**
6	Accelerated	10	10	0.3	487.0	59.0		
	Normal	2 724	2 783	96.7	485.4	74.6		
	Delayed	82	85	3.0	413.0	83.9		
7	Accelerated	30	21	0.7	563.1	58.4		
	Normal	3 901	2 816	96.1	524.9	74.6		
	Delayed	130	94	3.2	459.1	84.4		
8	Accelerated	13	20	0.7	594.0	63.3		
	Normal	1 872	2 898	96.4	555.0	80.3		
	Delayed	58	89	3.0	473.8	89.1		

Software: IDB analyzer, weight: houseweight (houwgt), Plausible values rescaled for 1995.

Table 9: Estimate of age- and school year effects for school years 6–7 and 7–8 in natural sciences. All the regressions are based on "normal students". Note. Rescaled plausible values. Not comparable with Cliffordson.

					Coefficient			Coefficient			
	Grade	**Birth month**	**N***	**Constant**	**AGE (month)**	**s.e.**	**t-value**	**AGE (year)****	**GRADE (year)**	**s.e.**	**t-value**
Rep	6–7	Jan–Dec	6 625	465.2	1.08	0.34	3.2	13.0	26.7	6.48	4.1
Rep	7–8	Jan–Dec	5 773	471.6	1.74	0.37	4.7	20.9	8,9	5.78	1.5

Table 10. Comparison of results with original and new scales

School years			**Age effect (AGE years)**	**t-value**	**School year effect (Grade year)**	**t-value**
Year 6–7						
	PVma					
		Original	13.1	3.4	28.4	4.9
		New	13.7	4.3	24.0	4.4
	PVsci					
		Original	11,5	3.0	34.4	5.8
		New	13.0	3.2	26.7	4.1
Year 7–8						
	PVma					
		Original	16.8	3.8	17.5	2.8
		New	18,0	4.7	8.4	1.4
	PVsci					
		Original	21.1	3.8	13.6	2.0
		New	20.9	3.2	8,9	1.5

Comments: For school year 6–7 there are no major differences between the original scale and the new scale even though the school year effect is somewhat weaker with the rescaled data. On the other hand, the difference in the school year effect for 7–8 is important, both for maths and science. The school year effect is only half as large with new rescaled data and is not significant.

Skolverket

www.skolverket.se