The Effect of Relative Age in the First Grade of Primary School on Long-Term Scholastic Results: International Comparative Evidence using PISA 2003

Maresa Sprietsma

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

The Effect of Relative Age in the First Grade of Primary School on Long-Term Scholastic Results: International Comparative Evidence using PISA 2003

Maresa Sprietsma

Non technical summary

In most countries there exists a compulsory school starting age, which obliges cohorts of e.g. 6 year old pupils to start school each year. If the admission rule into the first grade of primary school is respected, then pupils that vary in age up to twelve months end up in the same class.

There are several reasons why being the youngest versus the eldest of their class may affect pupils performance. In the first grade of primary school, relative immaturity of the youngest pupils within their age cohort may disadvantage them, because they are less able to perform the same tasks. But they may also benefit from peer effects from their older classmates.

If such effects vanished over time, there would be little to worry about. However, relative age in the first grade of primary school may also have a more lasting effect on performance. Indeed, previous research argues that early childhood maturity can influence long-term cognitive performance through ability grouping. According to Allen and Barnsley (1993), it might be difficult for teachers to differentiate between ability and maturity for children at school starting age. As a result, if children have different levels of maturity at the start of primary school and they are grouped according to observed ability, the youngest ones may be more likely to end up in low-performing peer groups. Over time, they risk learning less and they may face a higher risk of repeating a grade in primary school (see e.g. Corman 2003).

In this paper, we provide causal evidence on the long-term effect of predicted relative age in the first grade of primary school on test scores at age 15. The predicted relative age is calculated using the pupil's birth month and the admittance rule into primary school. We also try to assess through which particular channels, for instance an increased risk of grade repetition, or selection into the vocational track, it occurs. This is done for 16 countries.

We find that predicted relative age does have a significant long-term impact on test scores, the eldest pupils having about twenty percent of a standard deviation higher test scores than the relatively youngest. Furthermore, according to our estimations, most of the effect of relative age comes from grade repetition and selection of pupils into different grades, as well as intra-class ability grouping.

The Effect of Relative Age in the First Grade of Primary School on Long-Term Scholastic Results: International Comparative Evidence using PISA 2003

Maresa Sprietsma[*]

Abstract

In this paper, we estimate the effect of pupil's relative age within the first grade of primary school on math and reading test scores at age 15. The main objective is to evaluate the long-term causal effect of relative age in the first grades of primary school on pupil's test in 16 different countries. We use the national rule for admission to primary school to construct the predicted relative age of each pupil. We find that relative age at the start of primary school has a significant positive effect on test scores in most countries. Moreover, we identify some of the channels through which the effect occurs.

Keywords: pupil performance, relative age, international comparison
JEL classification: I21, I29

1.   Introduction

In most countries there exists a compulsory school starting age, which obliges cohorts of e.g. 6 year old pupils to start school each year. If the admission rule into the first grade of primary school is respected, then pupils that vary in age up to twelve months end up in the same class.

There are several reasons why being the youngest versus the eldest of their class may affect pupils performance. In the first grade of primary school, relative immaturity of the youngest pupils within their age cohort may disadvantage them, because they are less able to perform the same tasks. But they may also benefit from peer effects from their older classmates. If such effects vanished over time, there would be little to worry about. However, relative age in the first grade of primary school may also have a more lasting effect on performance. Indeed, recent research argues that early childhood maturity can influence long-term cognitive performance through ability grouping. According to Allen and Barnsley (1993), it might be difficult for teachers to differentiate between ability and maturity for children at school starting age. As a result, if children have different levels of maturity at the start of primary school and they are grouped according to observed ability, the youngest ones may be more likely to end up in low-performing peer groups. Over time, they risk learning less[1] and they may face a higher risk of repeating a grade in primary school (see e.g. Corman 2003). Moreover, even if there is no ability grouping, we think that teachers may hold lower expectations from the youngest pupils which may lower their self esteem and effort in the long run. In the psychological literature there is qualitative evidence that teacher expectations and objectivity are important determinants of test scores[2].

In this paper, we provide causal evidence on the long-term effect of predicted relative age in the first grade of primary school on test scores at age 15. The predicted relative age is calculated using the pupil's birth month and the admittance rule into primary school. We also try to assess through which particular channels, for instance an increased risk of grade repetition, or selection into the vocational track, it occurs. This is done for 16 countries.

---

[1] See Bedard and Dhuey (2005) for a the model of human capital accumulation that corresponds to the intuition developed here.
[2] See e.g. Bulterman-Bos et al. (2002) for a recent paper

We construct our measure of relative age using arguably exogenous national rules for admission to primary school and birth month. Several recent papers use the exogenous variation provided by administrative cut-off dates for starting school to instrument school starting age and time in school in education production functions (e.g. Angrist and Krueger 1992, Leuven et al., 2004, Bedard and Dhuey 2005).

The majority of papers that investigate pupil's test scores as an outcome find a positive effect of being older when you start school. This seems to be true whether test scores are measured at kindergarten, the end of primary school or at the end of compulsory schooling. The 'age premium' seems to be decreasing over time, from ¾ of a standard deviation higher points for the oldest versus the youngest children in kindergarten (Datar 2004, Leuven et al. 2004), to about half a standard deviation at the end of primary school (Puhani and Weber 2005) and 20% of a standard deviation in secondary school (Frederiksson and Ökert 2005, Strom 2004). Fertig and Kluve (2005) on the other hand find no effect of age at school entry on grade retention or degree attainment in Germany. These studies respectively use data from the United States, the Netherlands, Germany, Sweden and Norway. Bedard and Dhuey (2006), using TIMSS[3] data from eleven OECD countries, find long lasting positive effects on student test scores in grade eight of being relatively older in grade one. These results seem to be more pronounced in countries with early tracking[4]. Furthermore, they find that relatively older children are more likely to attend college in New Zealand and the UK.

In our paper, all children have spent the same total time in school at the time of the test because they all belong to the same cohort (they were admitted in the same school year). However, we cannot distinguish relative from absolute age effects with our data because the relatively youngest kids in a given cohort will also be younger in absolute age at the time of the test. There is no agreement in the literature as to which of these effects dominates. Two of the mentioned papers comment on this issue, concluding to the dominance of differences in absolute age (maturity) based on the opinion of 25 primary school directors that answered a questionnaire (Puhani and Weber), or based on peer age composition in grade 9 (Frederiksson and Ökert 2005). However, we favour the relative age interpretation. First of all, differences in cognitive maturity in months seem unlikely to exist at age 15. Following Heckman(2000), "Cognitive ability differences are fairly well determined or stable by age 8-9". Moreover, if there were significant maturity

---

[3] Trends in International Mathematics and Science Study.
[4] Tracking refers to pupils being selected into different tracks at high school. Often there will be for instance a general versus a vocational track.

differences between children at age 15 we would expect the effect of RELAGE to be significant in all countries, which is not the case. Finally, Kawaguchi(2006) exploits a change in the relative age composition of a cohort of pupils born in 1966 due to a sudden drop in birth cohort size[5] in Japan to test the relative versus the absolute age effect. His evidence is consistent with the relative age effect hypothesis.

We contribute to existing research by providing causal evidence on the effect of predicted relative age in the first grade of primary school on test scores at age 15 in 16 different countries and by investigating the possible channels through which this effect may occur. The objective is to evaluate whether relative age in the first grade of primary school has a significant long-term effect on reading and maths test scores. Indeed, the causal literature on the effect of school starting age, summarised above, studies only a limited number of northern European countries, and does not proceed to inter-country comparisons (with the exception of Bedard and Dhuey 2005). Furthermore, we have seen the main theoretical explanation proposed for a long term effect of relative age is ability grouping i.e. grade repetition, skipping and non compliance to entrance rule related to relative age. We investigate whether this mechanism can indeed explain differences in performance of younger versus older pupils in the considered countries.

We find that predicted relative age does have a significant long-term impact on test scores, the eldest pupils having about twenty percent of a standard deviation higher test scores than the relatively youngest. Furthermore, according to our estimations, most of the effect of relative age comes from grade repetition and selection of pupils into different grades, as well as intra-class ability grouping. For instance, relatively younger pupils are more likely to repeat a grade, and grade repetition has a significant negative impact on test scores. Grade repetition is therefore considered a channel through which relative age affects test scores.

In the following section, we explain how our measure of relative age was constructed and why it is exogenous. In the third section, we present the dataset and clarify the choice of included countries. Our empirical strategy is the topic of section 4. In Section 5 we present the results of our estimations.

---

[5] The number of births decreased by 25 percent in 1966 due to a superstition about births in the 1966 'Fire-horse' year.

2.   Relative age

Our measure of relative age in the first grade of primary school is the predicted relative age of a pupil inside his/her age cohort based on the national rule for admission to primary school (RELAGE). In most countries with high scolarisation rates, there exists a legal cut-off date for admission to primary school. For instance, children may enter grade one if they have reached a certain age before the start of the school year. We construct the relative age of the pupils using these national cut-off dates and information on pupil's birth month. If e.g. the cut-off date is January 1st, the pupil should have reached the age of 6 (the compulsory school starting age) before the 1st of January in order to start school that calendar year. Therefore, pupils born in December are the youngest in class, because their birthday is just before the cut-off date. When the school year starts in e.g. September, they still have three months before they turn 6. Pupils born in January on the other hand have to wait 9 months after their birthday before they can start school: they are the eldest in class.

Note that all the children will have spent the same cumulated time in school at the time of the test. In effect, imagine the test takes place at the beginning of the calendar year following primary school admission. All children will then have had exactly four months of schooling (if the school year started in September).

The key advantage of using the national cut-off month for admission to primary school to construct RELAGE is that it provides us with an exogenous measure of a pupil's relative age inside his age cohort. Indeed, as explained above, it is the birth month of a pupil combined with the national cut-off month for starting school that determines its age position in class. We identify only two possible causes for endogeneity: birth month targeting by parents (in order to give their child a relative age advantage when starting school), or disease associated to birth months.

Bedard and Dhuey (2006) use the US Natality Files[6] and verify that parents do not target birth dates in a significant manner.
Mixed evidence exists in the psychological literature about seasonal variation in births of mentally ill children, the most important amount of research concerning autism and schizophrenia[7]. Such birth patterns would be related to different rates of occurrence of epidemics and differences in sunlight intensity and temperature between summer and winter (See e.g.

---

[6] These data contain the birth month and maternal socio-economic information for all US births.
[7] See e.g. Yeates-Frederikx et al.(2000), Landau et al.(1999) and Torrey et al.(1997) for reviews.

Yeates-Frederikx et al. 2000). Leaving aside a discussion on the quality and abundance of evidence as to the existence of such birth patterns in this mainly qualitative branch of research, we do not think this represents a major issue for the identification of the effect of relative age. First of all, we use birth months relatively to legal cut-off dates as a variable, not season of birth. Moreover, we found no evidence as to any statistically significant association with cognitive skills of these mental problems. And, for many diseases, there is one specific month that appears to jump out as concentrating the highest frequency of patient birth dates. For unclear reasons, this is not the same month for all diseases and we found no evidence of systematic differences of IQ between birth months. Finally, we should keep in mind the low occurrence rates of autism and schizophrenia (around 0.1 % of adults are affected by autism[8], 1% of adults by schizophrenia[9]).

3. The data and choice of included countries
3.1. The data

This paper uses cross-section data from the OECD survey conducted in 2003 as part of the Program for International Student Assessment (PISA). The PISA 2003 database contains comparable math and reading test scores of a sample of 15 years old pupils coming from 40 countries. All pupil and school related variables are measured in the same way in all countries. Most importantly, the mathematics and reading test scores, which we use as our dependent variables, are also comparable across countries. Indeed, all pupils in the sample have been submitted to the same standard tests at age 15, taking into account possible school program and cultural differences. The scores are not comparable across disciplines. The sampling procedure occurs in two steps. First schools are randomly selected, then pupils inside them. The sampled set of schools is representative of the country's school population. Please refer to the PISA manual (2003) for more details as to how the data were constructed.
We have standardised test scores to mean 500 and standard deviation 100 by country in order to ease interpretation and comparisons of estimates.

3.2. Choice of included countries

In our paper, we only use 16 countries of the 40 countries in the dataset because we need the PISA sampling period to correspond to the legal cut-

---

[8] Dougherty (2000)
[9] Gottesman (1991)

off date for entering primary school. In this section we explain how and why we proceeded to this restriction.

PISA is not sampled by grade. To the contrary, the sample of pupils in the dataset was created based on age. We have a representative sample of fifteen year-olds, this means that we observe all pupils that were born in a given year. However, these pupils will enter the first grade at the same time only if by chance the sampling was done from the legal cut-off date. Figure 1 illustrates why this is the case. If the cut-off date to enter primary school is for instance January and pupils where selected from January to January then we observe a complete grade. If on the other hand the cut-off date is September but sampling was done from January then we do not observe pupils born from September to December in grade x. These attend a grade lower for which we do not observe kids born from January to August. We can make similar reasoning for all cut-off birth month – sample birth month combinations that do not match.

The reason we want to observe all pupils entering school a given year is that the interpretation of our results will not be comparable. In effect, in those countries where we observe a whole cohort entering primary school, RELAGE varies from 1 to 12. This is different in the countries where only a fraction of the birth cohort enters a given grade: RELAGE will take values from 1 to 9 only for example. The other pupils have to wait one year to start school and therefore will not have spent the same time in school at the time of the test.

In order to find out what the PISA sampling period was in each country it suffices to look at the birth month of the eldest pupils in the sample. If sampling was done from January to December then the oldest pupils in the country sample should be from January, the youngest from December.

We also need to verify what the legal cut-off date for starting primary school is in each country. We use the frequency patterns of pupils born in certain months per grade to check the legal cut-off date in each country. In the previous example, we for instance expect a frequency pattern of pupils per grade such that pupils born between September 1st and January 1st are mostly in grade 9, whereas those born before September should be mostly in grade 10. We thus compare the frequency pattern of pupils from each birth month per grade to infer the national cut-off rule for admission to first grade. The cut-off dates deduced from the data are next verified to be

correct using the information we found on the websites of the countries' Ministries of Education and on that of Eurydice[10].

Finally, we compare the sampling birth month with the double-checked cut-off date. If they are the same we include the country in our dataset. This yields 16 useable countries. Most countries included have sampling cut-off dates in January. Exceptions are Japan, Korea and New Zealand where these are April, March and May respectively.

3.3. Spreading across grades

As mentioned before, the PISA sample contains only 15 year-olds. But pupils may start the first grade of primary school at a higher(lower) age than the national compulsory school starting age. Some of these are put in the grade they should have started in at the legal age but some are placed a grade lower(higher). Therefore, when they reach the age of 15 they are one grade 'behind' ('in advance') with respect to the other 15 year olds. These pupils may attend other grades than the one they should attend according to the admission rule to primary school and their birth date. Moreover, pupils may repeat or skip a grade. They then end up in a lower (higher) grade than other pupils that have the same age.

The considered countries present different degrees of spread in the distribution of pupils across grades. Figure 2 presents the proportions of pupils in each grade, by country. In some countries (Iceland, Japan, Norway, Korea, Poland, Sweden, and former Yugoslavia[11]) nearly all 15 year olds are in the same grade. There are less than 3 percent of the pupils that repeated a grade and most pupils respect the entry rule. In other countries (such as Belgium, France, Portugal and Spain), 20 up to 30% of 15 year olds are in a lower grade than the one they should be in according to their birth-date. In the remaining countries (Canada, Denmark, Italy, Latvia, and New Zealand) between 5 and 15 % of pupils are in a too low grade according the admission to primary school rule.

4. Estimation strategy

The aim of our estimation will be to assess first of all whether relative age in the first years of primary school has a significant effect or not on test

---

[10] The information network on Education in Europe, European Commission (2006) . We also check our dates are the same as those used by Bedard and Dhuey (2005) for countries present in both our studies.
[11] Actual Serbia and Montenegro. We refer to the country by former Yugoslavia throughout the paper, as this was its name at the time of the survey.

scores at 15 years old. We estimate a classical education production function where test scores are a function of relative age for all included countries separately. We do not include pupil and school control variables because RELAGE is exogenous and should therefore not be affected by the presence of the control variables[12].

We take into account the hierarchical structure of educational data: pupils in a same class are thought to have similar unobserved characteristics, as do pupils from a same school or country. As a result, we know that pupil's test scores may be correlated if they are in the same school. We therefore estimate our equation as a mixed model that allows correlation of test scores inside the same school. This appears as the inclusion of a school 'random' effect in the estimated equation.

This yields specification (1):

$$SCORE_{is} = C + \alpha.RELAGE_i + \gamma_s + \varepsilon_{is}$$

Where γ_s stands for the school 'random' effect.

We call the estimate from specification (1) the aggregate effect of relative age. Its coefficient will capture variation in test scores that may be due to intra-grade relative age effects but also to misplacement[13], grade retention and selection into the vocational track or intra-class ability grouping.

As mentioned in the introduction, the theoretical literature points to ability grouping, grade retention and selection into vocational tracks as possible channels though which relative age may have long-term effects on pupil performance. From our data, we know whether a pupil repeated a grade, attends a vocational school and we can deduce from the cut-off date, birth month and grade currently attended whether the pupil started on time and how long he or she attended kindergarten. TIMEKINDERGARTEN is a discrete variable indicating the time spent in kindergarten by the pupil. Possible values are 0 (did not attend kindergarten), 1 (one year of kindergarten) and 2 (more than one year of kindergarten). This variable is relevant as in some countries children may already acquire certain skills in kindergarten, before they start school. All these variables relate to spreading pupils across grades in a way that could be correlated with relative age.

We do not have information as to the degree of intra-class ability grouping that goes on at the start of primary school. However, we do know whether the school groups pupils by ability within classes at age 15.

---

[12] We tested this hypothesis.

[13] By misplacement we refer to children starting school too early or too late, that therefore attend a different grade than the one predicted by their birth month.

In order to assess whether the mentioned channels might play a role and to what extent, we then estimate the effect of relative age on the probability of grade retention (dummy RETENTION equal to one if the pupil was retained one or more times), and of ending up in a vocational track (dummy VOCED) for each country. Relative age is exogenous with respect to these dependent variables. We would also like to know whether the probability of starting primary school too early or too late is related to RELAGE. Because very few pupils are concerned in some countries, we merely compute descriptive statistics of the relation between relative age and late or early start (dummies LATESTART and EARLYSTART equal to one if the pupil started one year too late or too early) of primary school.

Secondly, we would like to know whether relative age has an effect on long term test scores only through these channels or whether there is a 'residual' effect that exists independently of 'spreading' the pupils across grades.
We thus estimate a second specification of the test scores production function (1) where we include RELAGE as well as potential 'channel' variables RETENTION, VOCED, LATESTART, EARLYSTART, TIMEKINDERGARTEN and GROUPBYABILITY as well as dummy variables for each grade (except the one most pupils attend). Pupils that attend different grades are indeed expected to obtain different test scores at the same age of 15 as they have not received the same amount of teaching[14].

We are interested in the significance of the effect of the 'channel' variables but foremost in the residual effect of RELAGE. Indeed, the 'channel' variables are most likely endogenous. The choice of entry into primary school as well as grade retention can be expected to depend on unobserved pupil and parent characteristics such as ability or parent motivations. However, RELAGE remains exogenous. Correlations between RELAGE and the channel variables can therefore be interpreted as an effect of RELAGE on them, not the reverse. As a result, we can interpret changes in the coefficient of RELAGE on test scores as meaning that the channel variables are the channels through which RELAGE affects test scores.
This yields specification (2):

---

[14] It is possible to include both grade dummies and RETENTION because RETENTION is a dummy variable equal to one if the pupil repeated one or more grades, it does not specify the number of grade repetitions.

$$SCORE_{is} = C + \alpha.RELAGE_i + \beta.X_{is} + \beta_r.RETENTION_i + \beta_v.VOCED_i + \beta_E.EARLYSTART_i +$$
$$\beta_L.LATESTART_i + \beta_T.TIMEKINDERGARTEN + \sum D_{GRADE} + \gamma_s + \varepsilon_{is}$$

For each country, we adjust this specification according to the existence of grade retention, a vocational track, and the number of grades pupils are spread over.

5. Results

5.1. The baseline model

Table 1 shows the estimates of the aggregate effect of relative age on test scores for each country (specification (1)). This effect is significantly positive in 10(9) out of 16 countries for Reading and Maths test scores respectively. Being one month older in your age-cohort in the first years of schooling increases test scores at age 15 by between 0.87 and 2.06 (Portuguese), and 0.98 and 2.77 (Mathematics) points (one point is one percent of a standard deviation). Though highly statistically significant, this effect may seem negligible in size. We should not forget however, that pupils differ in age up to 11 months in the same grade. The obtained coefficients imply that the oldest pupil in a cohort (RELAGE = 12) on average obtains between 9.6 and 22.6 (Reading), 10.8 and 30.47 (Mathematics) points higher test scores than the youngest cohort member (RELAGE = 1). Relative age hence has an important and significant positive effect on long-term test scores in most considered countries. The magnitude of the coefficient is in line with previous results in the literature. Indeed, the effect of being the oldest pupil in a cohort is generally estimated to yield around 0.2 (between 10 and 27%) standard deviations higher test scores than the youngest pupils in the cohort (Strom, 2004, Frederiksson and Ökert, 2005, Bedart and Dhuey, 2005). Our standard deviation is 100 test score points therefore these effects would correspond to about 20 points in our paper.

Our results are not significantly different when we include control variables for pupil, family or school characteristics. The average difference between the coefficient of relative age in a specification including such controls and the specification with no control variables is of 0.01[15]. This indicates that RELAGE is indeed exogenous with respect to the other explanatory variables.

---

[15] Detailed results of the specification with no controls are available upon request.

5.2. What channels could relative age work through?

The results of the estimations of the impact of relative age on the probability of ending up in a vocational track and of repeating a grade are reported in Tables 2 and 3, for the countries where grade retention and a vocational track exist. The relative age of a pupil when he enters primary school has a significant effect on the probability of grade retention in six out of eleven countries. In these countries, being the eldest rather than the youngest in your age cohort when starting school reduces the probability to repeat a grade by between 2 and 15 percent.

Other factors that seem to play an important role[16] in determining grade retention are gender, parents' material wealth (HOMEPOS) and the number of books at home (NBBOOKS). Boys are more likely to repeat a grade than girls. Noticeably, not living with both parents has a very strong effect on the probability of repeating a grade. We could interpret this as being (indirectly) related to the temporary problems at home related to e.g. parents divorce.

The probability of reaching a vocational track does not seem related to relative age at school entry in most countries. The coefficient is significant in only two out of seven countries: Italy and former Yugoslavia. Strong determinants of the probability of being tracked into vocational education are parent education (HISCED), material wealth and the number of books at home. Not living with both parents has no significant effect on the probability of following the vocational track. Entering a vocational track seems to be much more related to the 'structural' family background than to more punctual events such as parents divorce or relative age at school entry.

In Table 4[17], we observe the percentage of pupils that attend a too high grade that are relatively eldest of their cohort according to their birth date. In all countries where there are 15-year-old pupils in too high grades, these are mostly (more than 50%) the eldest pupils of their cohort (relative age between 9 and 12). This points to a relation between skipping a grade and relative age. Relatively older pupils seem to be more likely to skip a grade.

Table 5 presents the correlation between relative age and early or late start of primary school in each country. In 9 out of 15 countries more than 2%

---

[16] Presented results are conform to specification 3. Complete results are available from author upon request.
[17] Recall that due to the low number of observations of pupils in a higher grade we cannot perform a probit estimation here.

of pupils start school a year or more too early. In 11 out of 15 more than 2% start school a year or more too late. In most concerned countries, starting too early is significantly positively correlated with relative age. In these countries, it is more likely to start school too early when you are the eldest of your age cohort. This is in line with the intuition that school entry decisions are probably based on observed pupil maturity. Starting school too late is significantly correlated with relative age in 11 out of 15 countries, but the sign of the correlation is not clear.

We conclude from this section that there is a strong relation between relative age and the variables we expect to be channels through which age affects long-term test scores. An exception is selection into the vocational track which does not seem related to relative age but rather to more 'structural' family characteristics (based on the restricted set of control variables included). We shall now look at potential residual effects of relative age (effects that would not go through these channels) and at the significance of the different channels in each country.

5.3. What channels does relative age work through?

Table 1 presents the estimates of specification 2 in the right column (residual effect). In the majority of countries (Belgium, Canada, Denmark, Latvia, New Zealand, Norway, Spain) the positive effect of relative age on test scores seems to happen mainly through the proposed channels. In effect, in 7(6) out of ten (nine) countries where RELAGE has a significant aggregate effect, the coefficient becomes non significant or is importantly reduced when including the 'channel' control variables.

Which channels are active in each country is reported in Table 6 (both disciplines). It is based on Tables 7 to 8 bis, which give the estimates for the channel variables of specification 2, and on the significance of the relation between relative age and these channels. Overall, the most important channels for RELAGE appear to be LATESTART, RETENTION and attending a grade higher than one should have according to the entry rule (GRADE+1). Starting school one year too late, repeating a grade or skipping a grade are strongly correlated with RELAGE, have a significant effect on test scores, and reduce the coefficient of RELAGE. This leads us to conclude that in those countries where RELAGE has a lasting effect on student performance, this mainly occurs through the pupils being spread across grades according to their relative age. Selection into the vocational track on the other hand is not related to RELAGE, except in Italy.

In France, Latvia and Portugal, the aggregate effect of RELAGE was not significantly different from zero. Including the channel variables, we obtain a significantly negative coefficient for the residual effect of RELAGE and the channel variables have a significant effect on test scores as well. This implies that spreading across grades also occurs in these countries, in a way that is unfavourable to younger pupils. However, given the spreading, i.e. inside a given grade, the remaining younger pupils perform better. Adding up this negative residual effect of RELAGE with the effect through the channels, the aggregate effect of RELAGE is not significantly different from zero in those countries.

In three countries, relative age has no effect on test scores in either discipline or specification. These are Iceland, Poland, and former Yugoslavia. In Iceland all pupils are in the same grade, there is no tracking into vocational education or grade retention, and pupils even all start at the right age. In Poland and former Yugoslavia, as well, the vast majority of pupils (96 and 97 percent of 15 year olds respectively) attend the same grade. It is therefore not surprising to us that relative age in the first years of primary school has no effect on test scores in these countries as none of the proposed channels through which we assumed it may work exist.

Denmark, Sweden, Japan, and Korea do not fit into this frame for at least one discipline. In these countries, relative age has a significant positive effect on test scores in both specifications. In Denmark and Sweden, the coefficient of RELAGE is significantly reduced by including GRADE+1 in Reading, but remains significantly positive and of the same magnitude in Maths. In Korea, relative age in the first years of primary school has a significant residual positive effect on Reading test scores only.
This implies that there is an effect of relative age at the start of primary schools on test scores in these countries, but it does not work through the proposed channels. In effect, in all four countries, most 15 year olds are in the same grade and there is very limited grade retention so few of the proposed channels apply.

The significantly positive aggregate effect of relative age in Japan, Korea (Reading) and Sweden, in the absence of the proposed channels, as well as the absence of an aggregate effect in countries where the channels seem active (France, Latvia, Portugal) point to the existence of other explanations for the effect of relative age. As yet unobserved characteristics of the schooling systems could allow certain countries to avoid the age premium, despite age related spreading across grades,

whereas others face an age advantage without any apparent ability grouping.

Moreover, we only control imperfectly for ability grouping. As mentioned before, we only have information on intra-class ability grouping at age 15 therefore we cannot test the relevance of intra-class ability grouping at the start of primary school as a channel for RELAGE. Nevertheless, Figure 3 shows that there is a considerable amount of intra-class ability grouping at age 15 and that many pupils attend different levels of maths classes in countries with a comprehensive schooling system such as Iceland, Norway and Sweden. There exists a positive correlation between the estimated residual effect of RELAGE and the degree of intra-class ability grouping at age 15 as can be seen in Table 9.

This does not hold for Japan and for Korea (Reading test scores). However, these countries have another particularity in that the schooling system is extremely competitive from the very start[18]. Pupils have to go through admittance examinations from an early age. This may imply that the eldest pupils, being more mature at time of first entry exam, are selected into the better schools. Again, we cannot test these channels due to the absence of information on the pupils at age 6. There is a positive correlation between attending a school where academic record is a prerequisite and RELAGE in Japan but further research using primary school data on these countries is required to check these hypotheses.

6. Conclusion

Relative age when entering primary school has a positive effect on long-term test scores in 10 out of the 16 studied countries. The size of the effect is of between 0.2 to 0.3 standard deviations in test scores, in line with previous findings in the literature. According to our results, in most countries the effect of relative age on long term test scores works through the spreading of pupils across grades. In particular, the effect seems to occur through increased grade retention and/or late start among the youngest pupils in an age cohort. Exceptions to this rule are Denmark, Japan, Korea and Sweden, where intra-class ability grouping and very early pupil competition could play a role. Moreover, there are also some countries where relative age has no effect on test scores in spite of the existence of ability grouping. We think these differences may be due to as yet unobserved characteristics of schooling systems. Further research as to

---

[18] Table 10 shows the percentage of pupils that attend schools where academic record is a prerequisite for admittance in each country. Noticeably, Korea and Japan are very selective schooling systems as compared to the rest of the countries at age 15.

how schools manage the maturity differences at the start of primary schools is necessary to disentangle the effects of teacher expectations, ability grouping, and peer effects as vectors of a lasting effect of relative age. In the meantime, more flexible admission rules into primary school such as allowing new pupils each semester instead of each year would make the relative age differences smaller.

Bibliography

Allen J., and Barnsley R. (1993), 'Streams and tiers: the interaction of ability, maturity, and training in systems with age-dependent recursive selection', *The Journal of Human Resources* 28(3): 643-659.

Angrist J. (2004), American education research changes tack', *Oxford Review of Economic Policy* 20:198-212.

Angrist J. and Krueger A. (1992), 'Does compulsory school attendance affect schooling and earnings?', *Quarterly Journal of Economics* 106: 979-1014

Bedard K., and E. Dhuey (2006), "The persistence of early childhood maturity: International evidence of long-run age effects", *Quarterly Journal of Economics*, forthcoming.

Bound J. and Jaeger D. (1996), 'On the validity of season of birth as an instrument in wage equations: a comment on Angrist and Krueger's "Does compulsory school attendance affect schooling and earnings?", *NBER Working Paper* No.5835.

Bulterman-Bos J., Terwel J., Verloop N., and Wardekker W. (2002), 'Observation in teaching: towards a practive of objectivity', *Teacher College Record* 104(6):1069-1100.

Corman H. (2003), 'The effects of state policies, individual characteristics, family characteristics and neighbourhood characteristics on grade repetition in the United States', *Economics of Education Review* 22: pp 409-420.

Datar A.(2005), 'Does delaying kindergarten entrance give children a head start?', *Economics of Education Review*, forthcoming.

Dougherty M.J. (2000), 'The genetics of Autism', Action Bio Science (peer reviewed online journal), available online: www.actionbioscience.org.

European Commission (2006), DG Education and Culture, Eurydice, 'The information network on education in Europe'. Available online : http://www.eurydice.org.

Fertig M., and J. Kluve (2005), "The effect of age at school entry on educational attainment in Germany", *IZA Discussion Paper* Nr. 1507.

Frederiksson P. and B. Öckert (2005), 'Is early learning really more productive? The effect of school starting age on school and labour market performance', *IZA Discussion Paper* Nr. 1659.

Heckman J.(2000), 'Policies to foster human capital', *Research in Economics* 54(1).

Kawaguchi D.(2006), 'The effect of age at school entry on education and income', *ESRI Discussion Paper Series* No.162, Tokyo, Japan.

Landau E.C., D.V. Cicchetti, A. Klin and F.R. Volkmar (1999), 'Season of birth in autism: A fiction revisited', *Journal of Autism and Developmental Disorders* 29.

Leuven E., M. Lindahl, H. Oosterbeek, D. Webbink, "New evidence on the effect of time in school on early achievement", mimeo, Universiteit van Amsterdam.

Musch J. and Grondin S. (1994), 'Unequal competition as an impediment to personal development : a review of the relative age effect in sport', *Developmental Review* 21 : 146-167.

Puhani P.A. and A.M. Weber (2005), 'Does the early bird catch the worm? Instrumental variable estimates of educational effects of age of school entry in Germany', *IZA Discussion Paper* Nr. 1827.

Strom B. (2004), "Student achievement and birthday effects", mimeo, Department of Economics, Norwegian University of Science and Technology.

Stipek D. (2002), "At what age should children enter kindergarten? A question for policy makers and parents', *Social Policy Report* 16: pp 3-16.

Torrey E.F., J. Miller, R. Rawlings and R.H. Yolken (1997), 'Seasonality of births in schizophrenia and bipolar disorder: a review of the literature', *Schizophrenia Research* Vol. 28.

Yeates-Frederikx M.H.M, H. Nijman, E. Logher and H.L. Merckelbach (2000), "Birth patterns in mentally retarded autistic patients", *Journal of Autism and DevelopmentalDisorders* 30(3).

Figure 1: Matching PISA sampling dates and national cut-off dates.

Birth months

    Jul.   Sep.   Nov. Jan. Mar. May. Jul. Sep. Nov. Jan. Mar. May
--|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|$\rightarrow$
Jun. Aug. Oct. Dec. Fev. Apr. Jun. Aug. Oct. Dec. Fev. Apr. Jun.

PISA sampling rule (example)
-------------------------|--|--------------------
               Jan.   Jan.
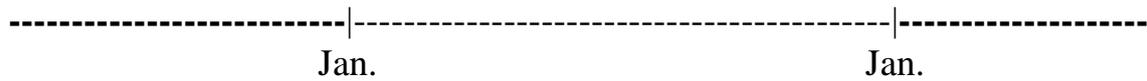
Cut-off date for admission into primary school (example)

-------------|--|---------------------------------$\rightarrow$
     Sep.   **grade 10**   Sep.   **grade 9**
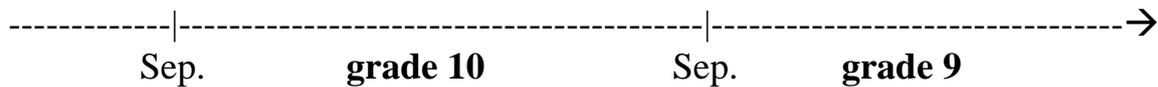
Missing observations if the PISA sampling rule is different from the admission rule.

-------------|--------------|-------------------------|--------------|------------------$\rightarrow$
     Sep.   Jan.   Sep.   Jan.

            Grade 10   **Grade 9**
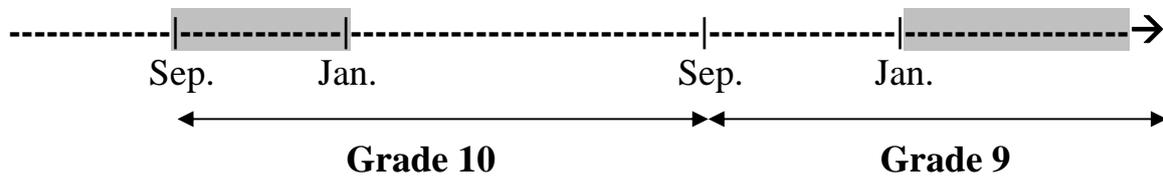
Figure 2 : Spreading of pupils over grades, by country.

Portugal
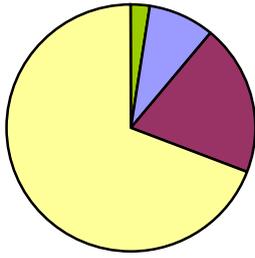
Spain
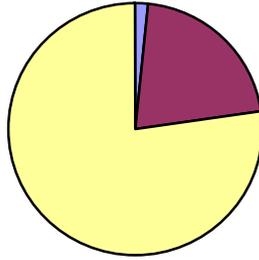
Sweden

Yugoslavia

Legend:

One grade too high

Two grades too low

One grade too low

Main grade
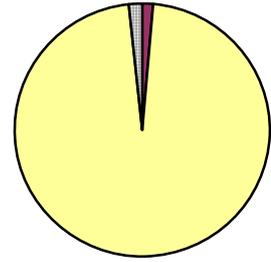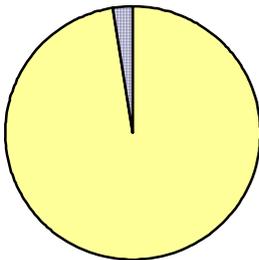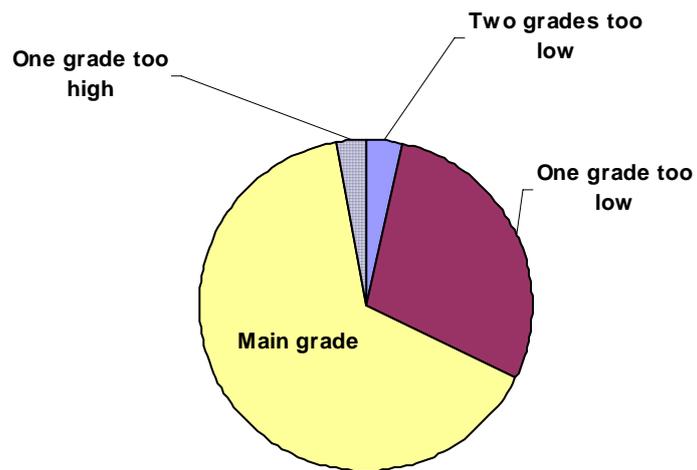
Figure 3:

Percentage of schools that practice ability grouping

Legend: ☐ Intra class grouping (maths)   ■ Different levels in maths
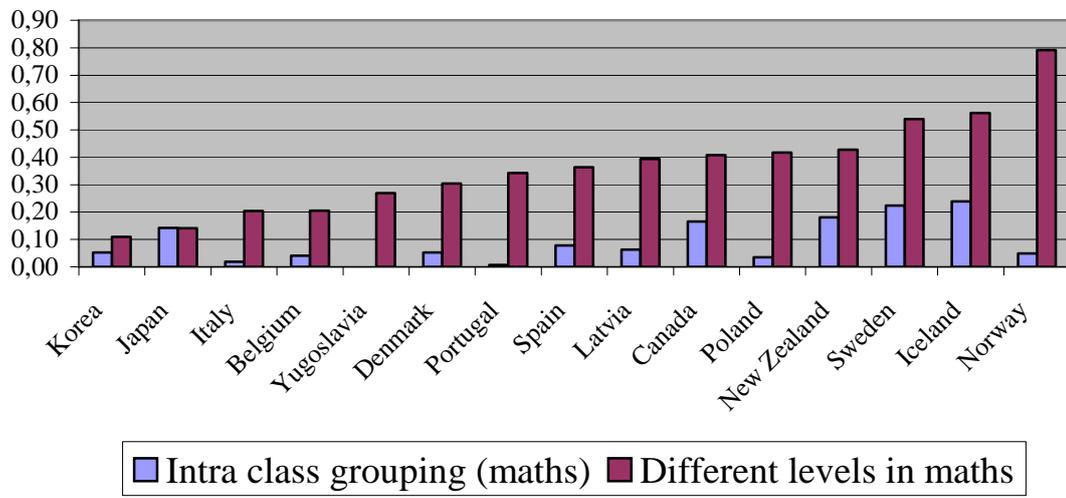
Table 1: Aggregate and residual effect of being one month older at the start of primary school
Dependant variables: reading and maths test scores at age 15

	READ						MATHS					
	Aggregate effect			Residual effect			Aggregate effect			Residual effect		
	Coeff.		StdErr	Coeff.		StdErr	Coeff.		StdErr	Coeff.		StdErr
Belgium	1,36	***	0,34	0,38		0,32	0,98	***	0,33	-0,16		0,31
Canada	0,93	***	0,29	0,23		0,28	1,40	***	0,28	0,51	*	0,28
Denmark	2,06	***	0,58	0,65		0,61	2,06	***	0,58	1,68	***	0,61
France	-0,50		0,48	-1,14	**	0,47	-0,59		0,48	-1,31	***	0,47
Iceland	0,18		0,58	0,18		0,58	0,90		0,58	0,90		0,58
Italy	0,94	***	0,25	0,68	***	0,25	0,36		0,24	0,08		0,24
Japan	1,11	***	0,38	1,11	***	0,38	0,99	***	0,34	0,99	***	0,34
Korea	1,13	***	0,36	1,20	***	0,37	0,26		0,34	0,30		0,35
Latvia	0,99		0,60	-1,16	*	0,64	1,57	**	0,60	-0,81		0,63
Norway	1,02	*	0,58	0,69		0,59	0,42		0,58	0,20		0,59
NZ	2,06	***	0,56	0,00		0,61	1,98	***	0,56	-0,03		0,60
Poland	0,19		0,44	-0,04		0,42	0,28		0,45	0,03		0,43
Portugal	-0,42		0,47	-1,19	***	0,44	0,36		0,48	-0,66		0,45
Spain	0,87	***	0,31	0,08		0,29	1,28	***	0,32	0,42		0,29
Sweden	1,28	**	0,52	0,78		0,52	1,93	***	0,52	1,53	***	0,52
Yugoslavia	-0,09		0,48	-0,29		0,49	0,62		0,48	0,28		0,49

*, ** and *** respectively stand for significance at a 10, 5 and 1% level of confidence

Aggregate effect :

Coefficient of RELAGE : specification 1.

Residual effect:

Coefficient of RELAGE when we include grade retention, VOCED, late and early entry and grade attended by pupil: specification 2.

Table 2: Probit estimation, coefficients of relative age (effect of being one year older). Dependant variable: Being in a vocational track.

Country	Coeff.	
Belgium	-0.03	
France	0.01	
Italy	-0.11	***
Japan	0.01	
Korea	0.01	
Portugal	0.02	
Yugoslavia	-0.04	*

*, ** and *** respectively stand for significance at a 10, 5 and 1% level of confidence

Only countries where there exists a vocational track are included.

Table 3: Probit estimation, coefficients of relative age (effect of being one year older). Dependant variable: Pupil has repeated a grade at least once.

Country	Coeff.	
Belgium	-0.09	***
Canada	-0.03	***
Denmark	0.00	
France	-0.12	***
Italy	-0.02	***
Latvia	-0.01	
New Zealand	0.00	
Poland	-0.01	
Portugal	-0.15	***
Spain	-0.07	***
Sweden	0.00	

*, ** and *** respectively stand for significance at a 10, 5 and 1% level of confidence

Only countries where there exists significant grade retention are included.

Table 4: Percentage of pupils attending a grade higher than their age cohort, that are the eldest of their age cohort (relative age between 9 and 12).

	Percentage	Nb. of pupils
Belgium	77%	51
Canada	70%	178
Denmark	90%	56
France	64%	65
Iceland	NA[19]	0
Italy	70%	196
Japan	NA	0
Korea	100%	4
Latvia	66%	156
Norway	NA	0
New Zealand	93%	81
Poland	58%	19
Portugal	80%	21
Spain	NA	0
Sweden	63%	54
Yugoslavia	90%	67

---

[19] Not applicable

Table 5: Correlation coefficient between RELAGE and early and late start of primary school.

Country	Corr. Coeff. Early start		Corr. Coeff. Late start		Percentage Early start	Percentage Late start	Nb. of Obs
Belgium	0.02	*	-0.03	**	2.0	10.6	5093
Canada	0.04	***	0.09	***	0.3	24.7	10465
Denmark	0.02		-0.16	***	4.5	7.0	2599
France	0.04	*	0.06	***	6.6	10.0	2380
Iceland	No obs.		0.00		0.0	1.7	2473
Italy	0.12	***	0.05	***	1.9	3.3	8717
Japan	No obs.		No Obs.		No obs.	No obs.	3709
Korea	-0.11	***	-0.01		1.2	7.0	4352
Latvia	0.19	***	-0.26	***	7.2	10.6	2160
New Zealand	0.03		-0.18	***	4.4	5.8	2429
Norway	-0.14	***	-0.06	***	5.2	0.4	2572
Poland	0.04	***	-0.01		0.6	5.5	3998
Portugal	0.04	**	0.04	**	0.8	11.4	2997
Spain	0.19	***	0.02	**	9.5	2.5	7449
Sweden	-0.02		0.035	**	3.6	0.7	3244
Yugoslavia	0.17	***	0		2.4	1.5	2577

*, ** and *** respectively stand for significance at a 10, 5 and 1% level of confidence.

<u>Early start</u> is a dummy equal to one if the pupil started school one or more years too soon compared to the compulsory school starting age.

<u>Late start</u> is a dummy equal to one if the pupil started school one or more years too late compared to the compulsory school starting age.

Table 6: Summary of effect of RELAGE and channels through which it occurs (both disciplines)

Country	Total effect of RELAGE	Voced	Earlystart	Latestart	Retention	Grade+1	Kindergarten
		Channels through which it occurs					
Belgium	Positive		x	x	x	x	
Canada	Positive				x	x	x
Denmark	Positive			x		x	x
France	None		x			x	x
Iceland	None						
Italy	Positive	x		x	x	x	
Japan	Positive						
Korea	Positive						x
Latvia	Positive		x			x	
New Zealand	Positive			x		x	
Norway	Positive		x				
Poland	None		x				
Portugal	None			x	x	x	x
Spain	Positive			x	x		
Sweden	Positive			x		**x**	
Yugoslavia	None					x	

	Kindergarten			Groupbyability			Repeat grade			Start late			Start early		
	Estimate		StdErr	Estimate		StdErr	Estimate		StdErr	Estimate		StdErr	Estimate		StdErr
Belgium	19,70	***	3,92	-9,85		6,64	-15,83	**	7,23	-13,37	***	3,74	-16,81	**	8,25
Canada	16,02	***	1,62	4,28		3,42	-75,21	***	5,14	0,32		2,29	11,01		21,11
Denmark	21,95	***	3,86	-18,47	**	8,13	-46,98	***	12,94	-10,51		9,52	5,36		11,56
France	12,35	*	6,64				-8,07		18,44	2,24		5,36	-20,09	***	6,81
Iceland	-3,38		3,71	-10,3	*	5,68				-25,47		15,43			
Italy	0,17		1,79	-7,17		5,99	-27,16	***	9,11	-11,04	**	4,92	-15,91		11,78
Japan	18,54	***	5,01	13,33		11,84									
Korea	-3,51		2,77	5,30		7,84	-72,57	**	28,44	-16,68	***	6,05	-10,79		12,14
Latvia	-6,19	***	2,22	1,20		9,20	-55,90	***	11,54	-4,46		9,65	36,40	**	16,52
Norway	14,40	***	3,39	-8,73		5,88				20,08		48,37	-25,57	***	8,88
New Zealand	9,42	***	3,31	-9,56		8,15	-57,73	***	10,53	-27,99	***	10,30	-7,91		9,18
Poland	9,72	***	2,63	10,00		6,52	-150,74	***	14,39	-34,92	***	6,47	-24,49		27,99
Portugal	-1,23		1,68	5,77		4,92	-102,31	***	4,41	-9,34	*	5,18	-29,84		21,43
Spain	11,99	***	2,03	-15,31	***	4,49	-25,03	**	11,78	-15,63	**	6,34	2,41		3,44
Sweden	14,48	***	2,56	1,40		4,50	5,24		14,71	-71,81	***	20,72	-4,25		11,37
Yugoslavia	-2,07		2,2	4,45		10,02	-67,18	**	25,37	-16,66		13,71			

Table 7: Coefficients of the control variables of specification 2, Reading test scores.

	Grade -3			Grade -2			Grade -1			Grade +1			Vocational track		
	Estimate		StdErr	Estimate		StdErr	Estimate		StdErr	Estimate		StdErr	Estimate		StdErr
Belgium	-232,62	***	48,95	-122,07	***	12,57	-40,56	***	7,23	41,28	***	11,61	-71,57	***	3,53
Canada				-100,34	***	22,68	-36,69	***	5,35	29,29	***	11,18			
Denmark							-44,01	***	11,09	62,41	***	17,28			
France				-170,35	***	20,93	-105,80	***	19,03	62,33	***	10,48	-80,34	***	8,23
Iceland															
Italy	-302,18	***	87,78	-145,09	***	22,97	-20,46	**	8,87	36,02	***	10,91	-62,16	***	5,32
Japan													-32,22	**	13,48
Korea							-22,84		19,63	92,47	**	41,49	-93,95	***	7,57
Latvia				-81,78	***	24,30	-37,58	***	10,16	6,94		16,78			
Norway							-69,54		46,00						
New Zealand				-131,48		91,69	-41,21	***	11,91	65,63	***	10,68			
Poland							-1,29		15,43	100,50	***	37,12			
Portugal				-116,60	***	9,88	-58,47	***	5,99	93,93	***	21,96	-35,93	***	5,16
Spain				-117,70	***	13,87	-68,98	***	11,78						
Sweden							-143,65	***	20,18	76,16	***	19,26	-34,36		30,58
Yugoslavia										17,38		10,82	-90,1	***	10,02

Table 7 bis : coefficients of the control variables of specification 2, Reading test scores, continued.

	Kindergarten			Groupbyability			Repeat grade			Start late			Start Early		
	Estimate		StdErr	Estimate		StdErr	Estimate		StdErr	Estimate		StdErr	Estimate		StdErr
Belgium	22,17	***	3,75	-6,72		6,82	-25,69	***	6,91	-10,73	***	3,58	-20,65	***	7,88
Canada	15,29	***	1,61	-1,90		3,58	-74,27	***	5,09	6,60	***	2,26	36,94	*	20,87
Denmark	22,51	***	3,88	-17,63	**	7,45	-72,45	***	13,01	-21,77	**	9,57	-4,53		11,63
France	18,83	***	6,58				-8,38		18,24	-2,76		5,30	-18,73	***	6,74
Iceland	4,86		3,74	-3,88		5,87				-14,61		15,51			
Italy	3,71	**	1,75	-14,04	**	6,85	-18,33	**	8,89	-22,18	***	4,80	-0,17		11,48
Japan	13,52	***	4,58	11,60		12,84									
Korea	-5,25	**	2,63	6,99		8,78	-19,71		27,12	-18,23	***	5,80	-6,74		11,55
Latvia	-2,65		2,18	10,08		9,19	-68,32	***	11,33	-6,12		9,47	11,80		16,22
Norway	14,67	***	3,40	-5,45		5,54				13,94		48,59	-21,63	**	8,91
New Zealand	10,79	***	3,29	-3,68		8,30	-60,29	***	10,48	-33,02	***	10,25	-6,36		9,13
Poland	16,19	***	2,68	10,50	*	6,01	-142,69	***	14,68	-38,96	***	6,59	-48,61	*	28,51
Portugal	5,29	***	1,70	10,39	**	4,46	-110,64	***	4,31	-7,51		5,24	-25,37		21,72
Spain	10,58	***	2,03	-11,38	**	4,40	-22,85	**	11,75	-19,36	***	6,32	5,76	*	3,43
Sweden	9,47	***	2,55	3,04		4,84	-0,43		14,79	-44,08	**	20,66	2,56		11,35
Yugoslavia	2,49		2,18	6,62		9,83	-38,51		25,20	-22,18	***	4,80			

Table 8: coefficients of the control variables of specification 2, Maths test scores.

	Grade-3			Grade-2			Grade-1			Grade+1			Vocational track		
	Estimate		StdErr	Estimate		StdErr	Estimate		StdErr	Estimate		StdErr	Estimate		StdErr
Belgium	-207,51	***	47,39	-122,58	***	12,06	-34,07	***	6,91	59,07	***	11,09	-83,21	***	3,44
Canada				-99,40	***	22,50	-44,12	***	5,31	43,17	***	11,06			
Denmark							-26,01	**	11,14	50,00	***	17,19			
France				-171,66	***	20,60	-109,71	***	18,71	74,84	***	10,37	-82,65	***	7,82
Iceland										31,96	*	16,48			
Italy	-341,67	***	88,29	-129,83	***	24,23	-33,55	***	8,65	38,25	***	10,64	-43,86	***	5,77
Japan													-40,19	***	14,63
Korea							-30,90		20,84	51,75		39,40	-103,37	***	8,04
Latvia				-112,64	***	23,86	-39,34	***	9,97	31,96	*	16,48			
Norway							-24,24		46,21						
New Zealand				-91,80		91,20	-39,18	***	11,85	57,68	***	10,63			
Poland							8,05		15,74	144,84	***	36,72			
Portugal				-106,31	***	9,96	-44,09	***	6,06	118,02	***	22,26	-38,87	***	5,19
Spain				-135,24	***	13,83	-73,22	***	11,75						
Sweden							-106,72	***	20,18	72,62	***	19,63	-19,77		30,67
Yugoslavia										40,53	***	10,75	-89,24	***	9,85

Table 8 bis: coefficients of the control variables of specification 2, continued.

Table 9: Correlation coefficient of the estimated effect of RELAGE (in maths) and the percentage of schools that group pupils by ability in each country.

	Aggregate effect	Residual effect
Intra-class ability grouping at age 15	0.35***	0.43***
Different level courses in maths	0.09**	0.08***

*** stands for significance at a 1% level of confidence.

Table 10: Percentage of pupils that attend a school where academic record is a prerequisite for admittance, by country.

	Academic record prerequisite
Average other countries	0.02
Belgium	0.15
Korea	0.41
Japan	0.67
Yugoslavia	0.78