Assessing the effect of schooling with cross-sectional data
Between grades differences addressed as a selection-bias problem

Hans Luyten; University of Twente, Enschede, the Netherlands
Bernard Veldkamp; University of Twente, Enschede, the Netherlands

This paper shows that it is possible to assess the effect of schooling accurately with cross-sectional data. The methodology applied requires that the data relate to two (or more) adjacent grades. The difference in learning outcomes between grades can be conceived as partly resulting from selection bias. Differences result primarily from the education received, but it should be taken into account that some of the factors - besides date of birth - that play a role in the assignment process also affect learning outcomes. Information on these factors may not be available, but this is not necessarily a problem. Heckman's two-stage procedure provides a solution. This is illustrated through analyses of TIMSS-95 data. These reveal substantial effects of education on mathematics and science achievement in fifteen different countries. With regard to attitudes, hardly any significant effects were found for science. The effects on attitudes toward mathematics are mostly negative.

Introduction

The primary aim of this paper is to illustrate how one can obtain an accurate assessment of the effect of one year schooling on student achievement and attitudes with data collected in cross-sectional surveys. The findings presented derive from a secondary analysis on data collected for the Third International Mathematics and Science Study, usually referred to as TIMSS-95 (Martin et al., 1997; Mullis et al., 1997). The analyses relate to primary school students (nine and ten-year-olds) in fifteen different countries. Assessments of the effects of one year schooling are reported for both achievement and attitudes with regard to mathematics and science. The methodology applied differs from the usual approach in educational effectiveness research, which focuses on variation in learning outcomes between schools. Since the publication of the famous Coleman report over four decades ago (Coleman et al., 1966), the term *school effect* and the percentage of variance in outcomes of learning situated at the school level have become nearly synonymous in the literature on educational effectiveness. It should be noted, though, that when measured in this way the "school effect" only expresses *relative* differences between schools. Numerous studies have confirmed the finding reported by Coleman et al. (1966) that these differences account for only a modest percentage (10-15%) of the variance in student achievement (Scheerens & Bosker, 1997).

The relatively small amount of variation between schools does not necessarily imply that the contribution of schooling to the development of students is a minor one. Lack of variation may be due to a high quality of instruction in most schools. Most researchers in the field of educational effectiveness would agree that a fair comparison between schools requires that differences in student backgrounds (like SES) are taken into account. In general, controlling for prior achievement is considered highly recommendable if not indispensable. Our approach is based on the difference in learning outcomes between grades rather than on variation among schools and does not require longitudinal data. It sets little demands on the dataset to be analyzed. It should comprise an age-cohort of students that are distributed over two (or more) adjacent grades. In addition, the learning outcomes from the various grades should relate to a common scale and information on the students' month of birth is required.

In contrast to the bulk of research on educational effectiveness, we also play explicit attention to the effect of schooling on non-cognitive outcomes. Although the officially stated goals of education in most industrialized countries include outcomes that relate to personal development, social skills and citizenship (Peschar, 1993), research on the effectiveness of education has thus far mainly focussed on cognitive outcomes. School examinations and public accountability systems also relate primarily to cognitive aspects of learning. As a result, both students and schools are stimulated to focus their efforts on the cognitive aspects of learning.

When trying to assess the effect of schooling one is faced with the problem that nearly everyone attends school. Heyns (1978) has pointed out that seasonality of learning may give a valuable indication, if one accepts the assumption the during the school year learning is affected by both school and non-school factors and that during the summer vacation only non-school factors have an impact. The difference between learning rates during the school year and summer learning rates may be taken as an estimate of the effect of schooling and has been termed *impact* by Downey, von Hippel and Hughes (2006). This approach requires longitudinal data at minimally three (but preferably more) specific points in time.

In a limited number of studies it has been shown that the effect of one year schooling can also be assessed with cross-sectional data (Cahan & Davis, 1987; Cahan & Cohen, 1989; Luyten, 2006; Luyten, Peschar & Coe, 2008; Tymms, Merrell & Henderson, 1997). The fact that in most countries students are assigned to grades primarily by date of birth is very helpful. If age is (virtually) the only factor that determines assignment to grades, the effect can be assessed simply by adjusting the differences between grades for the effect of age. Students born just a few days before the cut-off date usually end up in a higher grade than the ones that are just a little younger. In such cases application of the regression-discontinuity approach is appropriate (Shadish, Cook & Campbell, 2002; Trochim, 1984). Within each grade the effect of age on the outcome measure of interest is estimated and at the cut-off date a discontinuity in the age-achievement relationship is expected. Such a discontinuity between the oldest students in the lower grade and the youngest in the upper grade is considered to reflect the effect of one year schooling. This interpretation is based on the assumption that the impact of other factors besides age on assignment to grades is negligible.

For most countries, however, this is not a valid assumption. The cut-off date is usually applied with considerably flexibility. Especially grade repeating is a common phenomenon. In such cases a straightforward estimation of the statistical association between age and achievement is bound to produce a biased picture of the real relation. Grade repeaters are older than their classmates, but their test scores are typically below average. Gifted students, on the other hand, are sometimes allowed to skip a grade. Still, the estimated effect of age on achievement is an essential part of the analysis. Misspecification of the age effect will lead to a biased estimate of the schooling effect as well. Excluding students with non-standard school careers still yields biased estimates, as students with a birth date close to the cut-off point are (much) more likely to repeat a grade and most of the students with an accelerated career are born shortly after the cut-off date. See appendix 1 for an overview of the percentages of students with standard and non-standard careers in the fifteen countries on which this paper focuses.

In fact we are facing the following problem: We want to assess the effect of receiving schooling in the upper grade versus schooling in the lower grade, but it is very likely that assignment to grades is influenced by factors that also affect the outcome measures (i.e. student achievement). In other words: students in the lower grade perform less well, but some were assigned to that grade because they showed poor performance in the first place. This actually is a classic exmple of selection bias. Higher achievement scores in the upper grade result primarily from the education received, but it is necessary to take into account that other factors besides month of birth play a role in the assignment process. Information on the factors that determined the assignment may not be available, but this is not necessarily a problem. Heckman's two-step procedure provides a solution.

Controlling for selection bias

In the case of selection bias one is in fact faced with the problem that an explanatory variable is missing from the regression model (Heckman, 1979; Smits, 2003). The obvious solution is to add this variable to the model. In the two-step procedure proposed by Heckman (1976, 1979), the first stage involves the computation of a correction factor based on an analysis of the selection process. In the second stage this correction factor is included in the analysis as an additional explanatory variable. In order to determine the values on this correction factor one needs to involve the selection mechanism that causes the bias in the analysis. In the case of schooling the selection mechanism can be modelled as follows:

$$G = \alpha m + bc + e$$

This selection equation expresses that month of birth (m) and being born before or after the cut-off date (c) affect the probability that a student is assigned to the lower or upper grade (G). The effect of month of birth can be modelled as a polynomial function including linear, quadratic, cubic and higher order polynomial terms. In the above equation e represents a random residual that indicates to what extent the predicted value for G departs from the observed value. The selection model assumes an underlying continuous variable that determines a student's probability to be assigned to the upper grade, but this variable cannot be observed directly. In reality one can only assess whether or not a student has been assigned to the upper grade. The observed score for students in the lower grade is zero and one for students in the upper grade. In the present study the probabilities were estimated by means of binary logistic regression analysis. More frequently a probit analysis is conducted, but the probabilities obtained with logistic regression can easily be transformed into probit scores (Lee, 1983; Smits, 2003). The probabilities are not the main purpose of the analysis, but rather the residuals, as these contain the effects of all *unmeasured* variables that influence assignment to grades. As a result they provide the basis for constructing a correction factor (λ) that expresses the effects of these unmeasured variables, net of month of birth effect. In the second step of the procedure λ is added to the so-called substantial equation:

$$Y = \delta G + \zeta m + \theta\lambda + u$$

This regression equation includes three effects on the outcome variable (Y). The effect of being in the upper grade (G), the effect of month of birth (m) and the effect of the correction factor (λ). Deviations from the predicted score are represented by the random residual u.

An essential condition for application of the Heckman procedure is that the model needs to be sufficiently identified (Breen, 1996; Smits, 2003). The selection equation needs to include at least one variable that is not included in the substantial equation. Otherwise severe multicollinearity problems will arise, which will result in high standard errors and unreliable coefficients. In the analyses to be presented the effect of being born before or after the cut-off date was excluded from the substantial equation. This reflects the supposition that being born before or after an arbitrary cut-off point does affect assignment to grades, but is not causally related to the outcomes of learning. Any effect of the cut-off date is assumed to affect the outcome variable only indirectly trough the assignment to grades.

The correction factor λ is a monotone decreasing function of the probability that a student is assigned to the upper or lower grade (Heckman, 1979). The formula for computing λ differs for students in the upper and lower grade:

$\lambda = \varphi/\Phi$ if $G = 1$ (students in the upper grade)
$\lambda = -\varphi/(1- \Phi)$ if $G = 0$ (students in the lower grade)

Where Φ is the probability of being assigned to the upper grade and φ the corresponding standard normal density function. For students with a low probability of being assigned to the upper grade λ becomes large and positive if they actually are in the upper grade[1]. If the probability is .01, the corresponding density equals .0267. As a result λ equals 2.76. If the same student were in the lower grade, which is much more in line with expectations, λ would be close to zero (-.0267/.9900 = -.0269).

Research questions
The analyses aim to answer the following research questions:

1. What is the effect of one year of schooling on student achievement with regard to mathematics and science?
2. What is the effect of one year of schooling on student attitudes with regard to achievement in mathematics and science?

The effects will be reported as standardized regression coefficients and effect sizes.

Data
The data analysed derive from the cross-national TIMSS-95 survey. The aim of the TIMSS-surveys is to provide policy makers, educators, researchers and practitioners with internationally comparable information about educational achievement and learning contexts with regard to science and mathematics (Mullis et al.; 1997; Martin et al., 1997). Follow-up surveys have been conducted in 1999, 2003 and 2007. For the purpose of the present study the data from the more recent surveys cannot be used, as our approach requires data from students that are distributed over two adjacent grades.

Our analyses relate to students in primary education, often referred to as population 1. Twenty-six countries participated in the primary education component of TIMSS-95, which focused on the two grades with the largest proportion of nine-year-olds. In most countries these are the third and fourth grades. Data were collected in May and June of 1995, except for the countries on the southern hemisphere schedule, where tests were administered in late 1994. The analyses focus on fifteen countries and include only the students born in the three months after the cut-off date and the ones born in the nine months before. The aim was to obtain a nearly complete cohort of students born within a range of twelve months, of which a considerable percentage (about 25%) would be in the lower grade. In countries where 1

[1] Smits (2003) details how to compute the appropriate values for λ (LAMBDA) with the SPSS software. First of all one should transform the probabilities obtained with binary logistic regression analysis into probit scores. If PROB is the variable denoting the probabilities, individual probit scores (IPS) can be obtained with a straightforward compute command:
COMPUTE IPS = PROBIT (PROB).
Next these scores are used to compute the correction factor LAMBDA:
IF (grade =1) LAMBDA = ((1/sqrt(2*3.141592654))*(exp(-IPS*IPS*.05)))/cdfnorm(IPS).
IF (grade =0) LAMBDA = -((1/sqrt(2*3.141592654))*(exp(-IPS*IPS*.05)))/(1 - cdfnorm(IPS)).

January is the cut-off date, this means that students born in the months October through September are included.

As a result the analyses comprise only half the students of the original datasets, which include two intact grades. Besides, some students from the intended cohort could not be included anyway; first of all, the ones whose career was delayed or accelerated more than one year. These students were in a grade that was not included in the TIMSS-95 survey. Students born in the nine moths before the cut-off point are still included if their career was delayed by only one year. They are in the lower grade. The one year accelerated students born in the three months after the cut-off date are included as well, as they are in the upper grade. One-year delayed students born more than nine months before the cut-off date are not included, nor the accelerated ones born more than three months after. In nearly every country these students represent a relatively small percentage of the delayed and accelerated ones, as the probability of both delayed and accelerated school careers is highest for students whose birth date is close to the cut-off date. Inevitably some students were "lost", but by focusing on the ones born in the nine months before and the three months after the cut-off date the percentage of lost students was minimized. Countries are included in the analyses only if the percentage of lost students is below five percent. This applies to fifteen countries.

For two countries in TIMSS-95, Kuwait and Israel, application of our selection-bias approach is not possible, as tests were administered in only one grade. In other countries the cut-off date varies between regions within countries (e.g., Australia and the United States). In other countries schools are allowed to determine the cut-off date themselves (e.g., Ireland, New Zealand and Portugal). In such countries our approach could be applied, but only if additional information were available about the cut-off dates applied. Other countries were excluded because their percentages of lost students are above five percent. In most of these countries the number of students whose school career was delayed by more than one year is relatively large. This means that even if information on these students were available our approach would be faced with the complication that they are distributed over more than two grades. This would render the analysis of the selection process more complex. Table 1 lists the fifteen countries included in the analyses and provides some basic information (cut-off dates, ranges of birth dates of the cohorts included in the analysis, numbers of students and percentages of lost students).

Variables

The dependent variables in the analysis are student achievement and attitudes with regard to mathematics and science. Our analyses of student achievement are based on the 'international proficiency scores'. The user guide for the TIMSS international database (Gonzalez & Smith, 1997) recommends using these scores for both international and within-country comparisons. The mathematics tests given to the primary school students covered the following six content dimensions: 1- whole numbers; 2- fractions and proportionality; 3- measurement, estimation and number sense; 4- data representation, analysis and probability; 5- geometry; 6- patterns, relations and functions. The science test covered four content dimensions: 1- earth science; 2- life science; 3- physical science; 4- environmental issues and the nature of science. Table 2 shows the average math and science scores for each country by grade. These data relate to the findings in the entire datasets (not the constructed twelve month cohorts) and show the raw differences between grades.

Table 1: Basic information on countries and cohorts included in the analyses

	Cut-off date	Range of birth dates		Students included	Upper grade	Lower grade	Students lost
Austria	1 September	Dec. 1984	Nov. 1985	2,387	66.6%	33.4%	4.1%
Canada	1 January	April 1985	March 1986	7,413	65.9%	34.1%	4.9%
Cyprus	1 March	June 1985	May 1986	3,271	72.9%	27.2%	1.7%
Czech republic	1 September	Dec. 1984	Nov. 1985	3,241	60.7%	39.4%	2.3%
England	1 September	Dec. 1984	Nov. 1985	3,115	73.3%	26.7%	0.4%
Greece	1 April	July 1985	June 1986	2,997	72.4%	27.6%	1.5%
Hungary	1 June	Sept. 1984	Aug. 1985	2,727	68.7%	31.3%	4.7%
Iceland	1 January	April 1985	March 1986	1,743	77.5%	22.5%	0.4%
Japan	1 April	July 1984	June 1985	4,379	75.1%	24.9%	0.3%
Korea	1 March	June 1984	May 1985	2,669	70.2%	29.8%	4.1%
Netherlands	1 October	Jan. 1985	Dec. 1985	2,270	62.7%	37.3%	3.1%
Norway	1 January	April 1985	March 1986	2,226	76.0%	24.0%	0.2%
Scotland	1 March	June 1985	May 1986	3,137	72.9%	27.1%	0.9%
Singapore	1 January	April 1984	March1985	6,974	75.4%	24.6%	1.4%
Slovenia	1 January	April 1984	March 1985	2,492	77.8%	22.2%	1.9%

Table 2: Average scores and standard deviations for math and science achievement

	Mathematics Achievement				Science Achievement			
	Lower grade	Upper grade	Diff.	Pooled std. dev.	Lower grade	Upper grade	Diff.	Pooled std. dev.
Austria	487.02	559.25	72.23	85.26	504.56	564.75	60.20	83.70
Canada	469.46	532.13	62.67	80.93	490.43	549.26	58.83	87.03
Cyprus	430.42	502.42	72.00	81.88	414.72	475.44	60.72	74.41
Czech republic	497.18	567.09	69.90	84.55	493.67	556.51	62.84	82.93
England	456.50	512.70	56.20	89.30	499.23	551.46	52.24	98.29
Greece	428.05	491.89	63.84	87.57	445.87	497.18	51.31	82.74
Hungary	476.10	548.36	72.26	88.74	464.42	531.59	67.17	84.74
Iceland	410.11	473.77	63.67	69.80	435.38	504.74	69.36	83.52
Japan	537.91	596.83	58.92	78.25	521.78	573.61	51.83	72.70
Korea	560.90	610.70	49.80	72.03	552.92	596.90	43.98	69.79
Netherlands	492.88	576.66	83.78	67.72	498.84	556.67	57.83	64.59
Norway	421.30	501.87	80.57	72.86	450.28	530.27	79.99	88.04
Scotland	458.03	520.40	62.37	84.47	483.85	535.61	51.76	94.22
Singapore	552.08	624.95	72.87	102.06	487.74	546.69	58.94	98.00
Slovenia	487.64	552.41	64.77	79.91	486.94	545.68	58.74	77.05
Average	**477.71**	**544.76**	**67.06**	**81.69**	**482.04**	**541.09**	**59.05**	**82.78**

Student attitudes toward learning mathematics at school were assessed by means of their responses to the following three statements: 1- I like mathematics; 2- I enjoy learning mathematics; 3- Mathematics is boring (reversed scale). Students were asked to indicate their level of agreement with these three statements and the index of overall attitudes towards mathematics is based on a student's average response to these three items. With regard to science three similar questions were asked. Table 3 shows the average attitude scores with regard to math and science scores for each country by grade. When it comes to achievement the upper grade averages clearly exceed the lower grade averages in each and every country. With regard to attitudes towards mathematics the average score tends to be lower in the upper grade, while there appears to be no clear pattern for the science attitude scores.

Table 3: Average scores and standard deviations for math and science attitudes

	Mathematics Attitudes				Science Attitudes			
	Lower grade	Upper grade	Diff.	Pooled std. dev.	Lower grade	Upper grade	Diff.	Pooled std. dev.
Austria	3.27	3.08	-0.19	0.90	3.26	3.19	-0.07	0.88
Canada	3.30	3.29	-0.01	0.78	3.18	3.08	-0.10	0.81
Cyprus	3.53	3.55	0.02	0.63	3.23	3.34	0.11	0.74
Czech republic	3.23	3.06	-0.17	0.72	3.03	3.07	0.04	0.75
England	3.38	3.27	-0.11	0.86	3.15	3.10	-0.05	0.87
Greece	3.61	3.60	-0.01	0.63	3.55	3.54	-0.01	0.63
Hungary	3.28	3.04	-0.24	0.76	3.15	3.09	-0.06	0.79
Iceland	3.51	3.45	-0.06	0.76	3.11	3.16	0.05	0.89
Japan	3.03	2.89	-0.14	0.73	3.27	3.22	-0.05	0.68
Korea	3.18	2.97	-0.21	0.83	3.05	3.28	0.23	0.74
Netherlands	3.03	2.87	-0.16	0.90	3.00	2.92	-0.08	0.89
Norway	3.30	3.13	-0.17	0.85	3.03	3.11	0.08	0.86
Singapore	3.44	3.41	-0.03	0.66	3.42	3.31	-0.11	0.68
Slovenia	3.40	3.24	-0.16	0.72	3.28	3.24	-0.04	0.73
Average	**3.32**	**3.20**	**-0.12**	**0.76**	**3.19**	**3.19**	**0.00**	**0.78**

Students in Scotland were not asked to report their attitudes toward mathematics and science

Grade level and date of birth are independent variables of crucial importance in the present study. Grade level was re-coded to assign scores of zero to students in the lower grade and scores of one to students in the higher grade. The variable that denotes a student's date of birth is based on year and month of birth[2]. In each country the lowest score is assigned to the students born three months after the cut-off date and the highest to the ones born nine months before. A positive effect will indicate that older students get higher scores. The effect of the unmeasured variables that affect assignment to grades is captured by the correction factor λ.

Analysis

The first step in Heckman's two-step procedure entails an analysis of the assignment process. This results in a selection equation, which provides the basis for computing the correction factor λ. The probability for each student to end up in the upper grade was estimated with logistic regression analysis. Month of birth and a dummy variable indicating whether or not a student was born before the cut-off date served as the independent variables. In an exploratory manner polynomial terms were included to model the month of birth effect. For each country the analysis started with only the cut-off dummy and a linear term for month of birth. In subsequent stages higher order polynomial terms were included step by step, until the inclusion of a new term no longer presented a statistically significant improvement of the model fit (for a < .10 in a two-tailed test). The estimated probabilities were transformed into probit-scores, which were then used to compute the correction factor λ.

In the second step, the so-called substantial equations are estimated. These provide the answers to our main research questions. The effect of grade level on achievement and attitudes with regard to mathematics and science are estimated. Month of birth and the correction factor λ serve as control variables in these analyses. The cut-off dummy is not included in these analyses.

[2] In most countries exact information on dates of birth is available. Additional analyses (not reported here) based on the exact birth dates revealed only marginal differences with analyses based on month of birth.

Appropriate estimation of effects requires that the TIMSS-95 sample design be taken into account in the data analyses. This has been be achieved by using the appropriate sampling weights, which ensure the proper and proportional representation of the various subgroups constituting the sample in the computation of population estimates. The weight applied in the analyses is the one recommended in the user guide for the TIMSS-95 international database (Gonzalez & Smith, 1997) when performing significance tests. Its sum adds up to the sample size within each country.

Findings: selection equations (step 1)

The binary logistic regression analyses conducted to model the selection process yield a significant and positive effect ($a < .0005$, two-tailed) of the cut-off dummy in each of the fifteen countries. This indicates that, in line with expectations, the probability to be assigned to the upper grade is higher for students born before the cut-off date. In addition to the cut-off date, month of birth also affects a student's chance of being assigned to the upper grade. The only exception is Iceland. In eight countries a quadratic function suffices to model the relation between date of birth and assignment to grades. Regardless the exact shape of the function (quadratic or more complex), the patterns that emerge invariably indicate that for students born before the cut-off date the probability to end up in the upper grade is relatively low for the ones with a birth date close to the cut-off. Still, the strength of this effect varies considerably across countries (e.g., compare the Czech Republic with Japan or Norway). Among the students born after the cut-off date the chances of still being assigned to the upper grade increase, as their date birth approaches the cut-off point.

Figure 1 presents a graphical display of the relations between month of birth and assignment to grades per country. The horizontal axes denote a student's age (from young to old) and the vertical axes represent the probability of assignment to the upper grade. Each figure shows low probabilities for the younger students and high probabilities for the older ones. The effect of the cut-off date is depicted by the dashed line between the students born in the month before and after the cut-off point. Numerical details on the results of the logistic regression analyses are provided in appendix 2.

Findings: substantial equations (step 2)

In the second step of the Heckman procedure the substantial equations are estimated. These produce the findings we are really interested in. The effects of grade on achievement and attitudes are assessed controlling for the effects of age and the correction factor. The basic results with regard to achievement and attitudes are presented in tables 4 and 5 respectively. These tables report the standardized regression coefficients of grade, month of birth, λ and the percentage of variance explained. The regression models only include a linear term for the effect of age. Additional analyses (not reported) showed that including quadratic and cubic terms for month of birth do not produce an improved model fit.

The analyses reveal substantial effects of receiving education in the upper grade for both mathematics and science achievement, with the effect on mathematics being somewhat larger. Across countries the average standardized regression coefficient is 0.182 for mathematics and 0.142 for science. On average the effects of age and λ are similar for both subjects. The month of birth effect is 0.100 for both subjects. The effect λ of is 0.132 for mathematics and 0.123 for science. Note that for student achievement nearly all effects are positive, which is in line with expectations. Only one coefficient is slightly negative (effect of λ on science achievement in England), but not significant. The large majority of the effects are significant ($a < .05$; two-tailed).

Figure 1: Month of birth and assignment to grades

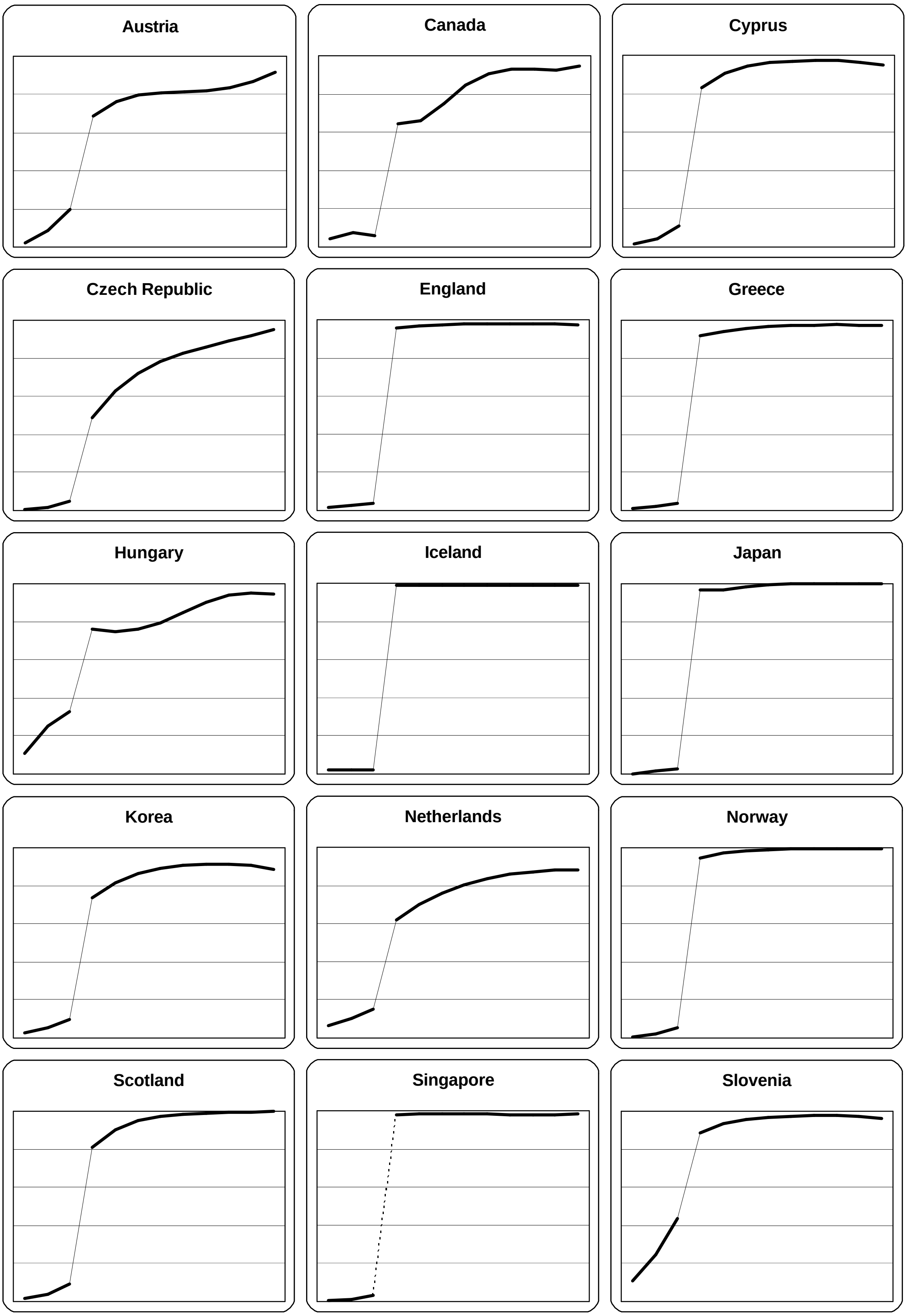

Table 4:
Effects of grade, month of birth and λ on achievement (standardized regression coefficients)

	Mathematics				Science			
	Grade	Month	λ	Var. expl.	Grade	Month	λ	Var. expl.
Austria	0.264	0.052	0.263	26.4%	0.132	0.080	0.309	19.4%
Canada	0.119	0.129	0.116	8.3%	0.044	0.149	0.215	9.1%
Cyprus	0.184	0.111	0.126	11.4%	0.138	0.137	0.155	11.0%
Czech republic	0.206	0.104	0.189	17.2%	0.183	0.099	0.157	13.2%
England	0.077	0.129	0.042	4.1%	0.068	0.116	-0.009	2.9%
Greece	0.174	0.078	0.090	7.7%	0.114	0.073	0.108	5.2%
Hungary	0.118	0.141	0.300	19.8%	0.197	0.104	0.181	16.1%
Iceland	0.212	0.089	0.054	8.8%	0.193	0.089	0.036	7.4%
Japan	0.176	0.086	0.056	6.9%	0.140	0.103	0.072	6.2%
Korea	0.167	0.089	0.073	7.5%	0.201	0.039	0.007	5.3%
Netherlands	0.355	0.052	0.223	32.1%	0.206	0.068	0.226	18.5%
Norway	0.297	0.100	0.048	15.3%	0.209	0.106	0.074	10.3%
Scotland	0.088	0.147	0.106	6.8%	0.039	0.134	0.063	3.3%
Singapore	0.171	0.058	0.023	4.9%	0.141	0.052	0.031	3.7%
Slovenia	0.122	0.138	0.270	15.4%	0.125	0.146	0.215	14.1%
Average	**0.182**	**0.100**	**0.132**	**12.8%**	**0.142**	**0.100**	**0.123**	**9.7%**

Figures printed in light grey denote statistically insignificant effects (a > .05; two-tailed)

Table 5:
Effects of grade, birth month and λ on attitudes (standardized regression coefficients)

	Mathematics				Science			
	Grade	Month	λ	Var. expl.	Grade	Month	λ	Var. expl.
Austria	-0.120	0.028	0.019	0.8%	0.050	-0.054	-0.042	0.1%
Canada	-0.067	0.007	0.052	0.1%	0.002	-0.064	-0.025	0.4%
Cyprus	-0.073	0.037	0.070	0.3%	0.062	0.010	0.020	0.6%
Czech republic	-0.184	0.026	0.071	1.5%	0.032	-0.004	-0.002	0.1%
England	-0.022	-0.032	-0.001	0.2%	-0.041	0.001	-0.028	0.2%
Greece	-0.041	0.018	0.018	0.1%	-0.035	0.011	0.033	0.1%
Hungary	-0.225	0.053	0.070	2.1%	-0.100	0.041	0.111	0.2%
Iceland	-0.095	0.030	-0.024	0.8%	0.043	-0.052	-0.049	0.3%
Japan	-0.099	-0.009	0.007	1.0%	-0.035	0.008	-0.013	0.1%
Korea	-0.100	-0.010	-0.005	1.2%	0.103	0.010	0.024	1.6%
Netherlands	-0.107	0.027	0.015	0.7%	-0.126	0.013	0.035	0.9%
Norway	-0.069	-0.017	-0.020	0.8%	0.055	-0.044	0.003	0.1%
Singapore	-0.102	0.051	0.020	0.5%	-0.095	0.009	-0.002	0.8%
Slovenia	-0.150	0.041	0.057	0.9%	0.046	-0.070	-0.028	0.0%
Average	**-0.104**	**0.018**	**0.025**	**0.8%**	**-0.003**	**-0.013**	**0.003**	**0.4%**

Figures printed in light grey denote statistically insignificant effects (a > .05; two-tailed)
Students in Scotland were not asked to report their attitudes toward mathematics and science

The analyses also reveal considerable differences between countries. The effect of schooling on mathematics achievement ranges from 0.077 (England) to 0.355 (the Netherlands). For science the effect ranges from 0.039 (Scotland) to 0.209 (Norway). The variation between countries is also substantial with regard to the effect of the correction factor: 0.023 – 0.300 for mathematics (Singapore – Hungary) and -0.009 – 0.309 for science (England – Austria). The variation of the age effect is more modest: 0.052 – 0.147 for mathematics (Austria – Scotland) and 0.039 – 0.149 for science (Korea – Canada). Table 4 also reports the percentages of variance explained by the three variables. On average these percentages or quite moderate, both for achievement in mathematics (12.8%) and science (9.7%).

With regard to attitudes more moderate, but in nine out of fourteen countries, significant effects of were found for mathematics. In each case the effect is negative. The standardized regression coefficients range from -0.022 (England) to -0.225 (Hungary). The average across countries equals -0.104. For science the average effect across countries hardly differs from zero (-0.003). Significant effects were found in only two countries, namely Korea (0.103) and Singapore (-0.095). On average, the effects of age and the correction factor are very small for both mathematics and science and insignificant in the large majority of cases. The percentages of variance explained are minimal (less than 1%), for both attitudes towards mathematics and science.

Tables 6 and 7 provide additional information on the findings. The non-standardized regression coefficients are reported, which allows for a straightforward comparison of the raw differences between grades (see tables 2 and 3) and the effects of schooling estimated with our approach. The non-standardized coefficients represent the differences between grades that can be attributed to the effect of schooling. For both mathematics and science achievement these are all smaller than the raw differences between grades, but it can be inferred that the effect of schooling accounts for a substantial portion of the differences between grades. In the case of Austria, e.g., the effect of schooling accounts for 69.0% of the raw difference in mathematics achievement between both grades (49.82/72.23). Across countries the schooling effect accounts on average for 51.0% of the raw differences in mathematics scores between grades. For science the percentage is 44.6%. The range in percentages is 28.6% – 72.3% for mathematics (England – the Netherlands) and 13.9% – 69.2% for science (Canada – Korea). With regard to attitudes a comparison between the non-standardized coefficients and the raw differences reveals that in some cases the estimated effect of schooling is even larger than the raw difference. In other cases the sign of the regression coefficient is negative, whereas the raw difference is positive or vice versa. Dividing the regression coefficient by the raw difference therefore yields percentages that are hardly interpretable. For that reason such percentages are not included in table 7.

The impact of schooling can also be expressed as and effect size. In that case one needs to divide the estimated effect (the non-standardized coefficient) by the pooled standard deviation in the control group and the experimental group (upper and lower grade). Expressing the effect of schooling in this way allows for a straightforward comparison with the efficacy of a broad array of psychological, educational, behavioural and medical treatments (Lipsey & Wilson, 1993). The pooled standard deviations are reported in tables 2 and 3. Thus one obtains an effect size of 0.58 (49.82/85.26) for mathematics achievement in Austria. This implies that due to schooling the students in the upper grade score 0.58 of a standard deviation higher than the ones in the lower grade. In other words: thanks to schooling 72% of the upper grade students get a score above the average in the lower grade (Cohen, 1988, pp. 21-23). This interpretation is based on the assumption of normal and equally variable distributions in both grades. Note that the actual percentage of Austrian upper grade students performing above the lower grade average is even higher because of the effects of age and the correction factor.

Table 6: Effects of grade on achievement

(non-standardized coefficients, standard errors, effect sizes and percentages of raw difference)

	Mathematics				Science			
	Coeff.	S.E.	Effect size	Perc.	Coeff.	S.E.	Effect size	Perc.
Austria	49.82	9.54	0.58	69.0%	24.81	9.91	0.30	41.2%
Canada	20.94	5.75	0.26	33.4%	8.81	6.07	0.09	13.9%
Cyprus	36.82	6.46	0.45	51.1%	24.21	5.70	0.33	39.9%
Czech republic	38.79	10.38	0.46	55.5%	32.58	10.02	0.39	51.9%
England	16.05	6.10	0.18	28.6%	15.04	6.53	0.15	28.8%
Greece	36.04	6.46	0.41	56.4%	21.59	6.01	0.26	42.1%
Hungary	24.03	12.37	0.27	33.3%	37.53	11.82	0.44	55.9%
Iceland	37.54	6.25	0.54	59.0%	40.23	7.41	0.48	58.0%
Japan	33.23	4.43	0.42	56.4%	23.85	4.00	0.33	46.0%
Korea	27.44	6.87	0.38	55.1%	30.43	6.40	0.44	69.2%
Netherlands	60.60	9.63	0.89	72.3%	30.63	9.20	0.47	53.0%
Norway	55.54	6.11	0.76	68.9%	44.55	7.15	0.51	55.7%
Scotland	17.99	6.96	0.21	28.8%	8.40	7.55	0.09	16.2%
Singapore	42.14	4.49	0.41	57.8%	32.66	4.23	0.33	55.4%
Slovenia	25.66	8.36	0.32	39.6%	24.34	7.81	0.32	41.4%
Average	**34.84**	**7.35**	**0.44**	**51.0%**	**26.60**	**7.32**	**0.33**	**44.6%**

Figures printed in light grey denote statistically insignificant effects (a > .05; two-tailed)

Across countries the average effect size is 0.44 for mathematics achievement and 0.33 for science achievement. With regard to attitudes the average effect size is -0.23 for mathematics and virtually zero (-0.01) for science. For mathematics achievement the smallest effect size occurs in the English sample (0.18) and the largest in the Dutch dataset (0.89). The effect sizes for science achievement are in between 0.09 (Canada) and 0.51 (Norway). For attitudes towards mathematics the effect sizes range from -0.05 (England) to -0.51 (Hungary). The extremes for attitudes toward science are -0.22 (Singapore) and 0.22 (Korea). The average effect size for mathematics achievement implies that due to schooling 67% of the upper grade students score above the lower grade average. For science the percentage is 63%. The negative effect size for attitudes toward mathematics indicates that because of schooling 59% of the students in the upper grade reveal an attitude that is more negative than the average in the lower grade. The largest effect size reported here (0.89; mathematics, the Netherlands) implies that thanks to schooling 81% of the Dutch students in the upper grade score above the lower grade average. A zero effect size implies complete overlap between the distributions in both grades. In that case 50% of the students in the upper grade get a score above the mean score in the lower grade.

The standard errors for the regression coefficients are reported in tables 6 and 7 as well. This shows that the standard errors are substantial in most cases. The average standard error across countries is almost identical for mathematics and science achievement (7.35 vs. 7.32). This implies that a 95% confidence interval would be nearly 29 points wide. For attitudes towards mathematics and science the interval would cover about 0.29 points. The standard errors reported have been corrected for heteroskedasticity. When conducting ordinary least squares regression with SPSS-software, the obtained standard errors are based on the assumption that the variance of the error terms is constant regardless the values on the independent variables. This assumption may not always be met. In the present case, however, the corrected standard errors hardly deviate from the ones obtained with ordinary least squares regression[3].

[3] Smits (2003) details how one can correct for heteroskedasticity using the SPSS-software.

Table 7: Effects of grade on attitudes
(non-standardized coefficients, standard errors and effect sizes)

	Mathematics			Science		
	Coefficient	S.E.	Effect size	Coefficient	S.E.	Effect size
Austria	-0.23	0.11	-0.26	0.10	0.11	0.11
Canada	-0.11	0.06	-0.14	0.00	0.06	0.00
Cyprus	-0.10	0.05	-0.16	0.10	0.06	0.14
Czech republic	-0.27	0.09	-0.38	0.05	0.09	0.07
England	-0.04	0.06	-0.05	-0.04	0.06	-0.05
Greece	-0.06	0.05	-0.09	-0.05	0.05	-0.08
Hungary	-0.39	0.12	-0.51	-0.17	0.12	-0.22
Iceland	-0.17	0.07	-0.22	0.09	0.08	0.08
Japan	-0.17	0.04	-0.23	-0.06	0.04	-0.08
Korea	-0.18	0.08	-0.22	0.16	0.07	0.22
Netherlands	-0.20	0.13	-0.22	-0.23	0.13	-0.26
Norway	-0.14	0.07	-0.17	0.11	0.07	0.13
Singapore	-0.16	0.03	-0.24	-0.15	0.03	-0.22
Slovenia	-0.26	0.08	-0.36	0.05	0.08	0.06
Average	**-0.18**	**0.07**	**-0.23**	**0.00**	**0.07**	**0.00**

Figures printed in light grey denote statistically insignificant effects (a > .05; two-tailed)
Students in Scotland were not asked to report their attitudes toward mathematics and science

Discussion

In this paper we have presented a methodology to assess the effect of schooling that, unlike the standard approach in research on educational effectiveness, is based on the difference in learning outcomes between students in adjacent grades. What makes the approach very convenient from a practical point of view is that it does not require longitudinal data. As cross-sectional data suffice, results with regard to the contribution of schooling to student development will be quickly available. In our view, the approach provides a great opportunity to increase the usefulness of large-scale national surveys and cross-national research projects like PISA, PIRLS and TIMSS. Like regression-discontinuity, it capitalizes on the fact that in most countries students are assigned to grades by date of birth, but it does not require a strict adherence to this criterion. In addition, the effect of the correction factor (λ) reveals the impact of unmeasured variables on learning outcomes. The fact that this correction factor yields consistently positive effects on achievement in mathematics and science provides further support for the validity of our approach.

In future research the approach can be extended in several ways. First of all, it can be applied within a multilevel framework. In that case the substantial equations should be estimated through multilevel analysis and grade should be modelled as a random effect at the school level. This means that not only the overall effect of schooling is estimated, as in the approach presented here, but that one also obtains an assessment of its variation across schools. Another extension of the basic model would be to include additional independent variables. This may be both school and student level variables. Including additional variables is not expected to affect the estimate of the grade effect, as the effects of unmeasured variables are assumed to be captured by the correction factor. However, a relevant question to be addressed is whether the effect of schooling is stronger for certain students (e.g., high SES vs. low SES) or certain schools (e.g. private vs. public). The main effects of additional independent variables on learning outcomes are not the main goals of the analyses, but rather their interaction effects with grade. These can tell us whether the effect works differently for certain schools or students.

Another promising extension would be to increase the number of grades. Instead of only two grades a wider range could be addressed. In that case the first step of the procedure would involve a multinomial logistic regression. For each student, multiple probabilities would be estimated (one less than the number of grades) and as a result, multiple correction factors would be computed (see Haas-Wilson & Savoca, 1990 for an example of this method). The great merit of such an approach would be that the effect of schooling on the development of children can be assessed across a wide age range. Moreover, the data collection will require less time and effort than a research design that draws on longitudinal data and the findings will be available much sooner.

Traditionally, research on educational effectiveness has mostly drawn on standardized tests that focus on basic cognitive skills. The limited number of effectiveness studies that have addressed alternative measures such as attitudes toward learning, social competencies and citizenship suggest that the school level variance with regard to these measures is (very) modest (Van der Wal, 2004; Van der Wal & Rijken, 2002). It should be noted, though, that lack of variation across schools does not preclude a strong impact of education in general. When it comes to learning outcomes like motivation and academic self-concept, many studies report declining scores as students grow older (Wigfield, Eccles & Rodriguez, 1998). Often this is interpreted as an effect of schooling (Eccles & Midgeley, 1989), although it may also reflect an autonomous development as students grow up (Baumert, cited in Wigfield et al., 1998). This study shows how our approach allows for a disentangling of the effects of age and schooling. For nine and ten-year-olds the effect of schooling on attitudes towards mathematics appears to be negative, whereas the impact of age is largely negligible.

The approach presented here could be applied to explore the effect of schooling on many more outcome measures. A few studies that are based on regression-discontinuity and seasonality of learning have addressed measures for which it is not that obvious that they may be affected by schooling. Von Hippel, Powell, Downey and Rowland (2007) address the question whether school or out-of-school environments contribute more to childhood overweight in the United States. They report that growth in body mass appears to be typically faster during the summer vacation than during the kindergarten and the first-grade school years. Cahan and Cohen (1989) have used the regression-discontinuity approach to assess the effect of formal education as opposed to chronological age on intelligence development. Their study relates to fifth and sixth grade students in Hebrew-language elementary schools in Jerusalem. The results unambiguously point to schooling as the major factor underlying the increase of intelligence test scores.

It is also important to point to some drawbacks of the approach presented here. First, it requires a sample of students within a certain age range that are distributed over (at least) two adjacent grades. Most educational studies focus on intact grades, which is more convenient from a data collection perspective. For the analyses presented we constructed an age cohort for each country. The result of this construction process was that we included only half of the students from the original surveys. Another point that needs to be mentioned relates to the size of the standard errors of the estimated effects. As the correlations among the independent variables (grade, age and λ) become large, so do the standard errors. This multicollinearity problem is strongest in countries where the cut-off date is applied most flexibly. Although our approach does not require a strict adherence to the cut-off date, it only works if there still is a substantial difference in the probability of assignment to the upper grade between students on either side of the cut-off point. Although the datasets analyzed in this study are quite large (over two thousand students in almost every country), the standard errors for the effects of schooling are still substantial. The most straightforward method for reducing standard errors is to increase sample size. Another option is to increase the range of the independent variables. In this case that would mean covering a wider range of grades.

The moderate percentages of variance explained for student achievement (see table 4) indicate that many other factors affect student achievement besides grade level, age and the unmeasured variables that determine assignment to grades. The main advantage of our approach, however, is not that it produces an impressively high R^2. Its merit is that we obtain unbiased estimates of the contribution schooling makes to the progress in student achievement by controlling for the relevant variables that affect the difference between grades as well. As they all derive from the same population, the upper and lower grade students constitute equivalent groups in most respects. One only needs to take into account the difference in age and the unmeasured variables that influence assignment to grades. Including additional variables in the regression models would hardly affect our estimates of the schooling effects, either because students from distinct grades do not differ on these variables (e.g., gender and SES) or because their effects are captured by the correction factor (e.g., ability and motivation). The analyses presented here were conducted with SPSS, which is the most commonly used statistical package by researchers in the field of social sciences. To our knowledge there is only one package available that comprises a component able to perform the analyses in a single routine, namely LIMDEP (for more information go to http://www.limdep.com/).

An attractive feature of the approach presented here is that the impact of schooling can readily be expressed as an effect size. Expressing it in this way allows for a comparison with the in numerous meta-analyses reported effects of a wide range of treatments. It is especially revealing to compare the effects of educational treatments to those of established medical interventions on mortality. The in meta-analyses reported effect of chemotherapy for breast cancer, e.g., ranges from 0.08 to 0.11 (Lipsey & Wilson, 1993). A careful review of over three hundred meta-analyses of psychological, educational and behavioural treatment research by Lipsey and Wilson (1993) shows an average treatment effect of 0.47. This is exactly equal to the effect on mortality of AZT for AIDS (as reported in 1991; see Lipsey & Wilson, 1993; p. 1199). It should be noted, however, that the review of educational treatments by Lipsey & Wilson is restricted to intervention packages with specific purposes and target populations. In the present paper we have addressed the effect of one year education for the entire student populations in fifteen different countries. The magnitude of the effect size is inversely proportional to the variability of the sample being studied. Given the wide scope of the "treatment", the cross-national effect sizes on student achievement (0.44 for mathematics and 0.33 for science) should still be considered substantial. The effect sizes of one year schooling may also serve as a benchmark in evaluations of specific interventions.

Our analyses also revealed a considerable amount of variation between countries with regard to the effect sizes. What factors may account for these cross-national differences would be an interesting topic for future research. Another important issue is the consistent finding that education has a negative impact on student attitudes toward mathematics.

References

Breen, R. (1996). *Regression Models, Censored, Sample Selected enor Truncated Data.* Thousand Oaks/London/New Dehli: SAGE Publications.

Cahan, S. & Cohen, N. (1989). Age versus Schooling Effects on Intelligence Development. *Child Development, 60,* 1239-1249.

Cahan, S. & Davis, 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.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences, Second Edition,* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Coleman, J.S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F. & York, R. (1966). *Equality of Education Opportunity.* Washington, D.C.: Government Printing Office.

Downey, D.B. & Von Hippel, P.T. & Hughes, M. (2006). *Are 'Failing' Schools Really Failing? Using Seasonal Comparisons to Evaluate School Effectiveness.* Paper presented to the International Congress for School Effectiveness and Improvement. January 2006, Fort Lauderdale, FL.

Eccles, J.S. & Midgeley, C. (1989). Stage/Development Fit: Developmentally Appropriate Classrooms for Early Adolescents. In R. Ames & C. Ames (Eds.), *Research on Motivation in Education* (Vol. 3, pp. 139-181). New York: Academic Press.

Gonzalez, E.J. & Smith, T.A. (1997). *User Guide for the TIMSS International Database. Primary and Middle School.* Chestnut Hill: TIMS International Study Center.

Heckman, J.J. (1976). The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement, 5,* 475-492.

Heckman, J.J. (1979). Sample Selection Bias as a Specification Error. *Econometrica, 47,* 153-161.

Heyns, B. (1978). *Summer Learning and the Effects of Schooling.* New York: Academic Press.

Haas-Wilson, D. & Savoca, E. (1990). Quality and Provider Choice: A Multinomial Logit-Least-Squares Model with Selectivity. *Health Services Research, 24*(6), 791-809.

Lee, L.F. (1983). Generalized Econometric Models with Selectivity. *Econometrica, 51,* 507-512.

Lipsey, M.W. & Wilson, D.B. (1993). The Efficacy of Psychological, Educational, and Behavioral Treatment, Confirmation from Meta-Analysis. *American Psychologist, 48*(12), 1181-1209.

Luyten, H. (2006) An Empirical Assessment of the Absolute Effect of Schooling: Regression-Discontinuity Applied to TIMSS-95. *Oxford Review of Education 32*(3), 397-429.

Luyten, H., Peschar, J. & Coe, R. (2008). Effects of Schooling on Reading Performance, Reading Engagement and Reading Activities of 15-year-olds in England. *American Educational Research Journal, 45*(2), (manuscript accepted).

Martin, M.O., Mullis, I.V.S., Martin, M.O., Beaton, A.E., Gonzales, E.J. Kelly, D. & Smith, T.A. (1997). *Science Achievement in the Primary School Years: IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill: TIMSS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., Beaton, A.E., Gonzales, E.J., Kelly, D. & Smith, T.A. (1997). *Mathematics Achievement in the Primary School Years: IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill: TIMSS international Study Center, Boston College.

Peschar, J. (1993). *Prepared for Real Life. Establishing Indicators for Non-Curriculum Bound Indicators (NOBS) in a Comparative Setting.* OECD Network A Meeting: Vilamoura, Portugal.

Scheerens, J. & Bosker, R.J. (1997). *The Foundations of Educational Effectivenes.* Oxford: Pergamon.

Shadish, W.R., Cook, T.D. & Campbell, D.T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston/New York: Houghton Mifflin Company.

Smits, J. (2003). *Estimating the Heckman Two-Step Procedure to Control for Selection Bias with SPSS.* Retrieved from the Internet at: http://home.planet.nl/~smits.jeroen/selbias/Heckman-SPSS.doc.

Trochim, W.M.K. (1984). *Research Design for Program Evaluation, the Regression-Discontinuity Approach*. London: SAGE Publications.

Tymms, P., Merrell, C. & Henderson, B.(1997). The First Year at School: A Quantitative Investigation of the Attainment and Progress of Pupils. *Educational Research and Evaluation,* 3, 101-118.

Van der Wal, M. (2004). *Competencies to Participate in Life.* Groningen, the Netherlands: University of Groningen
Van der Wal, M & Rijken, S. (2002). *Cross Curriculaire Competenties (CCC). De samenhang tussen cognitieve schoolprestaties, CCC en schoolkenmerken* (Cross Curricular Competencies: The correlation between cognitive performance in school, CCC and school characteristics). Paper presented to the annual conference of the Dutch Educational Research Association, Antwerp.
Von Hippel, P.T., Powell, B., Downey, D.B. & Rowland, N.J. (2007). The Effect of School on Overweight in Childhood: Gain in Body Mass Index during the School Year and during Summer Vacation. *American Journal of Public Health, 97,* 696-702.
Wigfield, A., Eccles, J. & Rodriguez, D. (1998). The Development of Children's Motivation in School Contexts. *Review of Research in Education, 23,* 73-118.

Appendix 1:
Standard and non-standard school careers

	Standard	Delayed	Strongly delayed	Accelerated
Austria	79.5%	16.0%	1.9%	2.6%
Canada	82.7%	12.9%	1.7%	2.6%
Cyprus	93.4%	4.2%	0.3%	2.1%
Czech republic	81.2%	16.9%	1.3%	0.6%
England	98.0%	1.1%	0.0%	0.9%
Greece	94.9%	3.5%	0.7%	0.9%
Hungary	79.5%	11.5%	2.4%	6.6%
Iceland	98.7%	0.7%	0.0%	0.5%
Japan	98.9%	0.5%	0.0%	0.6%
Korea	85.1%	11.0%	1.0%	2.9%
Netherlands	80.2%	16.0%	0.9%	2.9%
Norway	98.0%	1.0%	0.0%	0.9%
Scotland	94.9%	3.3%	0.1%	1.7%
Singapore	97.0%	1.7%	0.9%	0.4%
Slovenia	88.2%	4.2%	0.3%	7.4%
Average	**90.0%**	**7.0%**	**0.8%**	**2.2%**

Appendix 2:
Selection equations; effect of month of birth and cut-off on assignment to upper grade

	Cut-off	Month	Month**2	Month**3	Month**4	Month**5	Constant
Austria	1.515	17.660	-28.400	15.382	---	---	-3.755
Canada	3.601	16.112	-131.441	374.419	-417.353	160.625	-3.147
Cyprus	2.866	13.356	-9.032	---	---	---	-4.249
Czech republic	2.145	16.387	-20.554	10.304	---	---	-5.360
England	5.930	5.586	-3.766	---	---	---	-4.100
Greece	5.173	6.913	-4.154	---	---	---	-4.364
Hungary	1.806	15.772	-60.625	92.309	-44.356	---	-2.086
Iceland	8.430	---	---	---	---	---	-3.702
Japan	7.247	38.753	-189.401	349.484	-193.598	---	-6.291
Korea	2.738	8.465	-5.550	---	---	---	-3.589
Netherlands	1.759	5.837	-2.941	---	---	---	-0.901
Norway	4.784	12.235	-6.331	---	---	---	-4.821
Scotland	2.852	11.131	-3.907	---	---	---	-4.143
Singapore	6.499	19.437	-33.488	17.618	---	---	-5.863
Slovenia	1.573	11.441	-2.083	---	---	---	-2.083