CPB Netherlands Bureau for Economic Policy Analysis

CPB Discussion Paper | 255

How much do children learn in school?

International evidence from school entry rules

Dinand Webbink
Sander Gerritsen

How much do children learn in school?

International evidence from school entry rules

Dinand Webbink* and Sander Gerritsen**

Abstract

This study provides the first estimates of the causal effect of time in school on cognitive skills for many countries around the world, for multiple age groups and for multiple subjects. These estimates enable a comparison of the performance of education systems based on gain scores instead of level scores. We use data from international cognitive tests and exploit variation induced by school entry rules within a regression discontinuity framework. The effect of time in school on cognitive skills strongly differs between countries. Remarkably, we find no association between the level of test scores and the estimated gains in cognitive skills. As such, a country's ranking in international cognitive tests might misguide its educational policy. Across countries we find that a year of school time increases performance in cognitive tests with 0.2 to 0.3 standard deviations for 9-year-olds and with 0.1 to 0.2 standard deviations for 13-year-olds. Estimation of gains in cognitive skills also yields new opportunities for investigating the determinants of international differences in educational achievements.

JEL Codes: I2, J24

Keywords: cognitive skills, educational policy, international comparison

* Erasmus School of Economics, Rotterdam, Tinbergen Institute, IZA Bonn;

** Erasmus School of Economics, Rotterdam, CPB, Netherlands Bureau for Economic Policy Analysis, The Hague. Corresponding author: webbink@ese.eur.nl;

We thank Roel van Elk, Bas ter Weel, Eric Hanushek, Coen Teulings, Thomas Buser, Hessel Oosterbeek, Erik Plug, Sacha Kapoor and seminar participants at the Tinbergen Institute, University of Amsterdam, University of Rotterdam, CPB Netherlands, ESPE 2013 conference in Aarhus for helpful comments.

1. Introduction

Many studies have found a strong association between the economic outcomes of nations and their cognitive skills (e.g. Hanushek & Woessman 2008). It is therefore important to study international differences in the production of cognitive skills, and to examine how much children learn in school and whether this differs between countries. International tests, such as PISA, TIMSS or PIRLS, measure differences in cognitive skills of students between countries. The outcomes of these tests are increasingly used for the benchmarking of education systems and for designing educational policies.[1] However, it is difficult to investigate how much children learn in school because of the complex nature of the production of human capital. In the economic literature that investigates the so-called educational production function, student achievement at any point in time is typically seen as a cumulative result of the entire history of all inputs, for instance from family, peers, teachers and school, and the individual's ability (Hanushek & Rivkin 2006). The multitude of observed and unobserved factors that might be important pose challenges for identifying the effect of time in school on cognitive skills and for assessing the performance of a country's education system. Previous studies in economics have addressed these challenges by applying quasi-experimental designs for estimating the effect of completed schooling (Cascio & Lewis 2006; Hansen et al. 2004), pre-primary education (Berlinski et al. 2009; Gormley and Gayer 2005; Leuven et al. 2010) or grade retention (Jacob & Lefgren 2009) on cognitive skills for specific countries and specific age groups. To our knowledge, however, previous studies in the economic literature have not attempted to identify the effect of spending one additional year in school on cognitive skills across countries, age groups and subjects enabling comparisons between countries. Moreover, the recent literature that investigates the determinants of international differences in educational achievement has mainly focused on identifying cross-country associations (Hanushek & Woessmann 2011).[2]

[1] For instance, Germany, Denmark and Japan have experienced a 'PISA-shock' that resulted in a range of educational reforms. Lower-than-expected results triggered intense public and political debate on educational performance (Breakspear 2012). TIMSS and PIRLS results have been used to inform policy considerations in for example Hong Kong, Norway, New Zealand, The Russian Federation and The Republic of South Africa. Participating countries use TIMSS and PIRLS for establishing achievement goals and standards for educational improvement, stimulating curriculum reform, and improving teaching (IEA, 2011).

[2] Some recent studies apply a quasi-experimental approach for investigating specific factors such as the effects of class size (Woessmann & West 2006), central exams (Jürges et al. 2005), relative age (Bedard & Dhuey 2006) or private school competition (West & Woessmann 2010) using data from international cognitive tests.

This study provides the first estimates of the effect of time in school on cognitive skills for many countries around the world, multiple age groups and multiple subjects which enable a comparison of the performance of education systems based on gain scores instead of level scores. We use data from international cognitive tests and exploit variation in time in school induced by school entry rules.[3] Students born in adjacent months are assigned to different grades due to these school entry rules. As a result, students that are almost the same age differ in their time spent in school. This provides the opportunity to isolate the effect of time spent in school from the effect of time spent outside of school.[4] We apply this framework for estimating the effect of spending one year in school for samples of countries that participated in international cognitive tests. Within this framework we also address issues, such as sampling bias and violations of the exclusion restriction, that have been neglected in previous studies that exploit variation induced by school entry rules (see Section 2).

This framework enables us to perform four types of empirical analyses. First, we estimate the average effect of one year of school time on student performance in math and science. This yields estimates across countries, for two age groups and for two subjects. Second, we are able to estimate the gains in cognitive skills for each separate country. These estimates capture the gain in student achievement from the last year in school before the test was taken which can be interpreted as a measure of the performance of the education system. Third, we rank countries based on this measure of performance and compare this ranking with the ranking based on the level of the scores in international cognitive tests that is currently used for benchmarking of education systems. Fourth, this framework provides new opportunities for investigating the determinants of international differences in student achievement. We illustrate this by examining the effect of external exit-exams on student achievement using a specification that yields estimates of gains in achievement, and compare these results with previous studies that used a control strategy for estimating cross-country associations.

For applying this framework data are needed that include students in adjacent grades that took the same test in the same period. The data collected in the 1995 TIMSS study offer the

[3] School entry cut-off dates have also been used for investigating the effects of relative age (Bedard and Dhuey, 2006), or the effects of school starting age on student performance (Black et al. 2011; Fredriksson and Öckert, forthcoming) and the effects of education on earnings (e.g. Angrist and Krueger, 1991).

[4] This approach was introduced by development psychologists for separating schooling and age effects on test scores in a regression discontinuity framework (e.g. Baltes and Reinert 1969; Cahan and Davis 1987; Cahan and Cohen 1989) and was recently applied in the economic literature for estimating the effect of completed schooling (Cascio and Lewis 2006) or the effect of early childhood education (Gormley and Gayer 2005) on cognitive skills.

opportunity to apply this framework.[5] In the TIMSS study 9-year-olds and 13-year-olds were tested in math and science. The achievement tests were based on a curriculum framework developed through an international consensus-building process by all participating countries. For the analysis we only use data from countries that apply clear nationwide school entry rules; 21 countries for the 9-year-olds and 34 countries for the 13-year-olds.

Our empirical results can be summarized in three main findings. First, we find large differences in the effect of time in school on student learning between countries for both subjects and age groups. Some countries produce high gains in cognitive skills whereas in other countries additional time in school does not increase cognitive skills. Countries that achieve higher gains in cognitive skills for math also achieve higher gains in science. Moreover, countries with higher gains for 9-year-olds also have higher gains for 13-years-olds. Across countries we find that time in school on average matters for student performance in international cognitive tests. A year of school time increases performance in cognitive tests with 0.2 to 0.3 standard deviations for 9-year olds and with 0.1 to 0.2 standard deviations for 13-year olds. Hence, the effect of time in schools seems to reduce with age. This might indicate that later grades add less to the knowledge base or that the tests do a poorer job at measuring the full range of skill differences.

Second, and most remarkable, we find no association between the estimated gains in achievement and the level of test scores of countries. At all levels of test scores we observe countries with high achievement gains and countries with low gains in achievement. The lack of association has been found for both tests (math, science) and for both age groups (9-year olds and 13-year olds). This implies that assessments of the performance of education systems based on the estimated gains in achievement often are inconsistent with performance assessments based on level scores, and raises concerns about the current use of the outcomes of international cognitive tests in educational policy. A mere focus on test score levels is likely to yield misleading information about the performance of the education system. Low levels of test scores, or declining trends in test scores, could not be the result of low performing education systems. High levels of test scores could mask low performing education systems. Our estimated gains in achievement tell a different story about the performance of education systems. Using these gain scores as an additional instrument for the assessment of the performance of education systems is likely to reduce the risk of providing misleading policy information. The estimation of gain

[5] More recent TIMSS studies only sample students in one grade.

scores, which becomes possible when the current collection of international data is extended towards samples of students in adjacent grades, is likely to improve decisions on educational policies.

Third, the estimation of gain scores can be important for investigating international differences in educational achievement. For instance, studies that use control strategies have consistently found that students perform better in countries with external exit exams (Hanushek & Woessmann 2011; Bishop 1997; Woessmann 2001, 2003). However, we do not find higher gains in achievement in countries with an external exit exam for 9-year-olds and 13-year-olds compared to countries that do not have external exit exams.

Our paper makes several contributions to the current economic literature. First, we contribute to the literature on the educational production function by applying a method for measuring gains in achievements across countries. To our knowledge no previous study has estimated causal effects of time in school for different countries using a quasi-experimental approach. This method produces estimates of gains in achievement by different education systems, which enable a comparison of the performance of education systems based on gain scores instead of level scores. We compare the assessment of the performance of education systems based on the estimated gains with the performance assessments based on level scores. This comparison reveals that educational policies solely based on test score levels are potentially misguided because they ignore the gains that have been achieved.

Second, we contribute to the literature that investigates the effect of time in school on student performance (e.g. Gormley and Gayer (2005), Hansen et al. (2004), Cascio and Lewis (2006), Berlinski et al. (2009), Leuven et al. (2010)). We add to this literature by investigating the effect of time in school across countries, age groups and subjects. Third, previous studies that exploited variation induced by school entry rules have neglected various issues that might bias the estimates, such as sampling bias, relative age effects or violations of the exclusion restriction. In this paper we explicitly address these problems. In particular, we interpret our main estimates as lower bounds and also generate estimates that are adjusted for sampling bias. Fourth, we contribute to the literature that uses international cognitive tests for investigating the determinants of international differences in educational achievements. The typical features of human capital production pose major challenges for the identification of the effect of characteristics of education system and the cross-sectional structure of the international tests hinders value-added or panel estimations. Therefore, it has been argued that 'further exploration of quasi-experimental settings in the international data should be high on the agenda' (Hanushek & Woessman 2011). This is exactly what this paper does. We apply a quasi-experimental approach using international data and this approach might yield new opportunities for identifying the effects of characteristics of education systems. We illustrate this by comparing estimates of the effect of external exit exams using level scores with the estimates based on gain scores.

This study is organized as follows. Section 2 explains the empirical strategy used for estimating the effect of one year of school time on test performance. The data used in the analyses are described in Section 3. Section 4 shows the estimates of the effect of one year of time in school for pooled samples of countries. In Section 5 differences between countries are investigated. Section 6 compares country rankings of level scores with country rankings of gains scores. Section 7 illustrates the opportunities of the quasi-experimental approach for investigating international differences in student achievements and Section 8 concludes.

2. Previous studies and empirical strategy

The basic framework in the economic literature that studies the effects of educational inputs models student achievement as a function of family, peer, community, teacher and school inputs and student ability (Hanushek & Rivkin, 2006). Student achievement at any point in time is seen as a cumulative result of the entire history of all inputs and the individual's initial endowment (e.g. innate ability). A common approach for modeling this so-called educational production function is to assume that the cumulative achievement function is additively separable and linear (e.g. Boardman and Murnane 1979; Todd and Wolpin 2003). Estimating the effect of input factors, such as time in school, is complicated because in any actual application we will generally not be able to control for all relevant school, family or student characteristics. If some omitted variables are correlated with time in school, then the estimated parameters will be biased. Hence, the cumulative character of the production of human capital poses challenges for identifying the effect of time in school.

Previous studies in economics have addressed these challenges by applying quasi-experimental designs for estimating the effect of schooling on cognitive skills.[6] The effect of schooling has been analyzed from different perspectives. A first strand of the literature focuses on the effect of completed schooling on cognitive skills. Several studies have used quarter of birth as an instrument for completed schooling (Neal and Johnson 1996; Hansen et al. 2004) as in the seminal paper by Angrist & Krueger (1991). These studies find that one additional year of completed schooling increase cognitive skills with approximately 0.2 standard deviations. A recent study investigates the effect of an increase of compulsory schooling by one or two years on cognition (Meghir et al. 2013). They find that the reform increased cognitive skills on average, with 7 to 10 percent of a standard deviation. Cascio and Lewis (2005) exploit variation induced by school entry rules for estimating the effect of completed schooling on cognitive skills measured by the Armed Forces Qualification Test (AFQT). They find that an additional year of high school raised scores of minorities with 0.3 standard deviations.

[6] For surveys of the development psychology literature on the estimation of schooling effects, see Ceci (1991) and Stipek (2002).

A second strand of the literature focuses on variation in schooling from pre-primary education.[7] For instance, Gormley and Gayer (2005) and Gormley et al. (2005) estimate the impact of Oklahoma's pre-K program for 4-year-olds in Tulsa on cognitive/knowledge test scores, motor skills and language scores by exploiting cutoff requirements for enrolment in pre-K. Attendance increases test scores by approximately 0.4 standard deviations.[8] A third strand of the literature focuses on grade retention. For instance, Jacob & Lefgren (2009) estimate the effect of grade retention on high school completion by exploiting a nonlinear relationship between current achievement and the probability of being retained. They find that retention among sixth-grade students does not affect the likelihood of high school completion, but retention of eight-grade students increases high school dropout.[9] Our study is also related to a fourth strand of the literature which uses so-called value-added models for estimating gains in student achievement or the rate of learning over specific time periods. These models include measures of prior achievement to eliminate confounding by past unobserved parental and school inputs, for instance for estimating teacher fixed effects which can be linked to teacher characteristics (e.g. Rivkin et al. 2005; Hanushek et al. 2005). Dynamic sorting of teachers and students might bias the estimated effects in these models (Rothstein 2010). Our approach also focuses on the estimation of gains in cognitive skills but uses a quasi-experimental approach instead of controlling for prior achievements.

Empirical strategy

In this paper we focus on estimating the effect of time in school on cognitive skills. For identifying the effect of time in school we use a quasi-experimental design that was first applied by development psychologists (e.g. Balter and Reinert 1969; Cahan and Davis 1987; Cahan and Cohen 1989) and recently also applied in economic studies (Cascio and Lewis 2006; Gormley and Gayer 2005). The key idea for identification is that school entry rules create variation in time in school for children born close to the cut-off date. Students that are almost the same age differ in their time spent in school. A comparison of the test scores of students around this cut-off date yields estimates of the effect of one school year. In this paper we apply this approach to samples of countries that participated in international cognitive tests. Figure 1 illustrates the approach

[7] Early childhood interventions like Head Start or the Perry Preschool Project have been studied intensively. For surveys, see Currie (2001) and Almond & Currie (2010).
[8] For other recent studies, see Berlinsky et al. (2009) and Leuven et al. (2010).
[9] For other recent studies, see Manacorda (2012) or Schwerdt & West (2012).

using scores from the math and science tests of the 1995 TIMSS study for 9-year-olds in two adjacent grades.[10] The top panel shows results for Singapore, the bottom panel shows results for England. The left figure shows the assignment of students to grades on both sides of the cut-off date of the school entry rule; the middle (right) figures show the scores on the math (science) test. The horizontal axis shows the age of the student relative to the cut-off date. Each dot represents a monthly average of the grade-level or the test score.

The left figures show that both countries quite strictly apply the school entry rule for assigning students to grades. Nearly all students to the left of the cut-off date are in the higher of the two adjacent grades and nearly all students to the right of the cut-off date are in the lower of the two grades.[11] In both countries we also observe that scores in math and science decline with age which confirms previous findings about the importance of age at entry for test performance (Bedard & Dhuey 2006). The cut-off date divides students of very similar age into groups that differ in the number of years they have spent in school. Students on the left hand side of the cut-off date have spent one more year in school than students on the right hand side. For students from Singapore we observe a discontinuity in the math and science scores around the cut-off date. This discontinuity can be interpreted as the effect of one year spent in school in Singapore. For students in England we do not observe a discontinuity in test scores. This suggests that one year spent in school in England does not add more to the performance of students in math and science measured in the 1995 TIMSS study than one year spent out of school.

Estimating the effect of time in school by exploiting school entry rules

In a situation of full compliance with the school entry rules, the effect of one year in school on student performance can be estimated by using a regression discontinuity model that exploits the discontinuities created by the entry rule. The basic assumption in this model is that students on both sides of the discontinuity are very similar and that the relationship between date of birth and student performance is smooth around the discontinuity.[12] For each country, the effect of one year in school can be estimated using the following specification:

[10] It should be noted that these grades also include 8-year-olds and 10-year-olds.

[11] The first stage estimates (Equation (2)) for Singapore and England respectively are 0.96 and 0.93. Section 5 presents the first stage estimates of all countries used in the estimations.

[12] Cascio & Lewis (2006) exploit variation in school-entry dates across states in the USA and use individuals in other states as controls. With this approach they don't need to assume that relationship between date of birth and student performance is smooth.

(1) $Y_i = \alpha_0 + \alpha_1 G_i + f(birthmonth - C) + \alpha_2 X_i + \varepsilon_i$

where Y_i is the student performance of student i, G_i is a dummy variable for being in the higher grade, $birthmonth_i$ is the month of birth of the student, C is the cut-off date of the country, X_i is a vector of control variables and ε_i are unobserved factors. In this specification $f(.)$ is a smooth function of age which is allowed to be different at either side of the cut-off (f_l and f_r), as suggested by Lee and Lemieux (2010):

$$f(birthmonth_i - C) = f_l(birthmonth_i - C) + S_i[f_r(birthmonth_i - C) - f_l(birthmonth_i - C)]$$

The main parameter to be estimated is α_1 which can be interpreted as the effect of one year of school time on the test performance. Identification of α_1 is based on the non-linear relationship between age and time in school around the cut-off date.

A concern with this approach is non-compliance with the school entry rules. The grade level of a student (G) might differ from the time spent in school because of retention or acceleration, or because of schools that do not comply with the country's school entry rule. In that case, Equation (1) would probably yield biased estimates of the effect of time in school because it is likely that students that deviate from the regular path are not a random draw from the population. This problem has been recognized in development psychology and in economics. Studies in development psychology have dealt with this problem by excluding non-compliers (e.g. Cahen & Cohen 1989). This creates, however, a non-random sample that might induce biased estimates. Studies in the economic literature on schooling or starting age have often dealt with non-compliance by using an instrumental variable approach in which the school entry rule is used as an instrumental variable for the grade level (e.g. Cascio and Lewis 2006; Bedard and Dhuey 2006). In this approach the variation in time in grade that is induced by the school entry rule is used for estimating the causal effect of time in a specific grade on cognitive skills. The first stage and second stage equations can then be estimated using Two Stage Least Squares (2SLS):

(2) $$G_i = \beta_0 + \beta_1 S_i + f(birthmonth_i - C) + \beta_2 X_i + \eta_i$$

$$S_i = 1[birthmonth_i < C]$$

(3) $$Y_i = \gamma_0 + \gamma_1 \hat{G}_i + f(birthmonth_i - C) + \gamma_2 X_i + \vartheta_i$$

where S_i is a dummy variable for being born on the left side of the cutoff date, which is equivalent to being eligible for one extra year of time in school. Estimation of γ_1 will yield the causal effect of time in grade if the usual IV-assumptions hold (see below). This estimate can then be interpreted as the effect of time in grade for those students who move to the next grade if their expected time in school, due to the school entry rule increases by one year. Hence, for students who follow the regular path through education without deviations such as retention or acceleration. The estimate of the effect of time in school on grade level (β_1) in the first stage equation indicates the proportion of the students of a specific country that stays on the regular track of the education system of that country. For applying this IV-approach three assumptions should hold. First, the school entry rule should have an effect on the grade level of students. Hence, there should be no weak instrument problem. The empirical analysis in the next sections shows that in all selected countries the school entry rule is an important determinant of the observed grade level. The second assumption is that the cut-off date should not be correlated with unobserved determinants of cognitive skills. We will address this assumption below. The third assumption, which is neglected in previous studies on schooling or starting age, is the exclusion restriction; the instrument should only have an effect on cognitive skills through the endogenous variable. In our application this means that the difference in cognitive skills between students born on either side of the cutoff date should only be the result of the time spent in the highest grade by students that are on track. However, all students on the left side of the cutoff have been treated with an additional year in school; a year in a higher grade or a year of being retained. Hence, it is assumed that grade retention or acceleration of students has no effect on cognitive skills. Given the recent studies on grade retention (see above) it seems not likely that this assumption holds. We, therefore, focus our analysis on estimating the reduced form of this IV-approach:

(4) $$Y_i = \delta_0 + \delta_1 S_i + f(birthmonth_i - C) + \delta_2 X_i + \varepsilon_i$$
$$S_i = 1[birthmonth_i < C]$$

For the identification of the effect of time in school on cognitive skills in Equation (4) several further issues are important. First, school entry rules not only induce a difference in the time spent in school for students of nearly similar age but also induce a difference in relative age in class (school starting age). Students on the left of the cut-off not only receive an additional year of education but are also assigned to be the youngest in their grade. Students on the right of the cut-off are assigned to be the oldest in their grade. Differences in relative age have been shown to be important for short-term and long-term cognitive outcomes (Bedard & Dhuey 2006). We address this identification issue by using a model specification that allows the effect of the assignment variable age to be different at either side of the cutoff. Age and relative age are perfectly correlated because both are measured from the cut-off date. This means that in our specification the age effect on both sides of the cut-off not only controls for maturity but also for relative age in grade.[13]

Second, some countries apply a clear school entry rule but also use a rolling admission of students. For these countries the school entry rule does not create a one-year difference in time in school, but leads to a different timing of grade promotion. Hence, for these countries students on the right side of the cut-off date have spent more time in lower grades than students on the left side of the cut-off date. We address this issue in the sensitivity analysis in which we exclude countries with rolling admission from the estimation sample.

Third, Equation (4) yields the causal effect of one year of school time on cognitive skills if the conditional independence assumption holds. Hence, the critical assumption is that students near the cut-off date are very similar on observed and unobserved characteristics. This assumption seems plausible since parents are unlikely to plan the exact date of birth of their child. However, there is evidence that parents in the U.S. schedule births in order to avoid taxes (Dickert-Conlin and Chandra 1999). Several recent studies have investigated whether birth

[13] Bedard and Dhuey (2006) estimated the effects of age relative to the cut-off data and frame the estimates in terms of relative age. These estimates are the combined effect of maturity and age at entry. Black et al. (2011) isolate the effect of these two variables.

around the school entry cut-off dates is random.[14] For the US (Dickert-Conlin and Elder 2010; McCrary and Royer 2010), Chile (McEwan and Shapiro 2008) and Argentina (Berlinski et al. 2011) no evidence has been found for the non-randomness of births around cut-off dates. However, the timing of births in Japan seems to be related with school entry cut-off dates (Shitgeoka, 2013). The number of births sharply increases in the first days after the cut-off date. Hence, some Japanese parents seem to have a preference for their children to belong to the oldest in class. This might induce a bias for the estimated effect because it is not clear which parents try to postpone the birth of their child. To address this issue we exploit the fact that our data contains information about the exact date of birth. We will perform sensitivity tests by using estimation samples in which we exclude students born on the first days around the cut-off date (see Section 4).

Fourth, a further and related issue, which is not addressed in previous studies that use school entry rules, is sampling bias. Our sample consists of students in two adjacent grades that contained the largest proportion of students from a specific age group; 9-year-olds or 13-year-olds (see next section). The disadvantage of this sampling strategy is that that we do not observe students from these age groups that are not in these grades. If we imagine a country in which the 9-year-olds are evenly distributed over the two adjacent grades, then the higher grade will contain the oldest 9-year-olds (group B) together with the youngest 10-year-olds (group A), and the lower grade contains the youngest 9-year-olds (group C) together with the oldest 8-year-olds (group D). Groups A and B are on the left side of the cutoff in Figure 1 and, groups C and D are on the right side of the cutoff. For our main estimation sample we use students from group B and group C, and we compare the difference in performance of these two groups at the cut-off. However, in group B we do not observe students that have been accelerated, and in group C we do not observe students that have been retained. It might be expected that this will induce a downward bias for the estimated effects because students that have been accelerated will probably have a relatively high ability, and students that have been retained will probably have a relatively low ability.[15] This implies that the estimated effects should be interpreted as lower bound estimates. It should be noted that the 'missing students' in our estimation sample are the

[14] Several studies have raised concerns about the randomness of season of birth (Bound and Jaeger, 2000; Cascio and Lewis 2006; Dobkin and Ferreira 2010; and Buckles and Hungerman 2012).

[15] In Table A.4 it can be observed that retained students born at the left side of the cut-off on average score lower than students that are on track and, that accelerated students born at the right side of the cut-off on average score higher than students that are on track.

students that, because of their relative age in grade, are the least likely to be accelerated or retained. To further address this issue we will perform two types of sensitivity analysis. First, we will estimate the main models for samples of countries in which most students are on track; countries with a first stage estimate (Equation (2)) of at least 0.75.[16] In these countries only a very small proportion of students will not be observed. Second, the advantage of the sampling strategy is that we also have data of students in groups A and D which we can exploit to approximate the sampling bias of the 9-year-olds in groups B and C. In group A, which contains the youngest 10-year-olds, we can observe students that have been retained. We use these students to adjust for sampling bias in group C, which contains the youngest students of the 9-year-olds. Moreover, in group D, the oldest 8-year-olds, we can observe students that have been accelerated. We use these students to adjust for sampling bias in group B, which contains the oldest students of the 9-year-olds. By assuming that the proportion of retained and accelerated students, and the relative score of these retained or accelerated students compared to students that are on track does not change between grades, we can adjust their scores and include them in the main estimation sample. Hence, for our approximation of the sampling bias we adjust the scores of some students from groups A and D, and include them in the main estimation sample consisting of students in groups B and C. We use these samples to obtain estimates that are adjusted for sampling bias (appendix A.4 provides further details about this procedure). For the main models we will show the lower bound estimates and the estimates that are adjusted for sampling bias.

3. Data

The data used in this study come from the 1995 TIMSS study.[17] The 1995 TIMSS study collected mathematics and science achievement results from third and fourth graders in 26 countries and from seventh and eighth graders in 41 countries.[18] These achievement tests are based on a curriculum framework developed through an international consensus-building process by all participating countries. International experts in mathematics, science, and measurement contributed to the development of the achievement tests and the tests were endorsed by all

[16] This first stage estimate should not be directly interpreted as the proportion of missing students. The missing students can be observed in groups A and D, the first stage estimate is based on group B and C.
[17] See, http://timss.bc.edu/timss1995i/Database.html for TIMSS data.
[18] These data were also used in a quasi-experimental study on international differences in class size effects (Woessmann & West, 2006).

participating countries. The sampling focused on the two adjacent grades that contain the largest proportion of 9-year-olds – third and fourth graders in most countries – or the largest proportion of 13-year-olds – seventh and eighth graders in most countries. These samples also include students that are one year younger or older than the age groups that were targeted. This sampling strategy enables us to apply the regression discontinuity framework that we discussed in Section 2. After the 1995 TIMSS study the sampling strategy was changed and focuses only on one grade, which makes it impossible to apply our estimation framework.

From the 1995 TIMSS study we include all countries in the analysis that apply clear nationwide school entry rules. For the nine-year-olds we included 21 out of 26 participating countries. Australia, the USA and Ireland have been excluded because the rules regarding the school cutoff date vary across regions or are at the discretion of educators or parents. Kuwait and Israel have been excluded because in those countries only one grade was sampled or no information on test scores was available. For the thirteen-year-olds we included 34 out of 41 participating countries. Again Australia, USA, Ireland, Kuwait and Israel have been excluded. Columbia has been excluded because there is no clear cut-off date in average grade. The Republic of South Africa has not been included because the teacher and school data were not deemed internationally comparable. Bedard and Dhuey (2006) excluded more countries from their estimation sample because of concerns about the strict application of the school starting age rules or measurement error in the date of birth. They additionally excluded Germany, the Netherlands, Hungary, Switzerland and Korea. In our analysis, which focuses on differences between grades, it seems that these countries can be included because we observe sharp discontinuities in average grade around the cutoff date. We test the robustness of our findings by replicating our main estimations for the sample of countries used by Bedard and Dhuey (2006).

As dependent variables we use the TIMSS test scores in math and science. These scores have been standardized with a mean of 500 points and a standard deviation of 100 points which can be easily translated into the usual effect sizes from a standard normal distribution. TIMSS uses an incomplete or rotated-booklet design for testing children on the major outcome variables. For each student and each test TIMSS selects five plausible values. In the estimation we use all five plausible values and adjust standard errors as recommended when using the plausible values methodology (Von Davier et al. 2009). Our main control variable is the date of birth of the student measured by month. For many countries we also have the exact date of birth, which we will use in the sensitivity analysis. Other control variables that we use are gender, born in country of test, living with mother/father, language of test spoken at home, number of books at home, and mother's and father's educational level.

School entry rules are crucial in our analysis. We use information from Bedard & Dhuey (2006) and several online sources, and empirically checked this information in our data. For some countries we could not obtain information about the cutoff dates. In those cases we used the cutoff date from the data (see Table A.1 in the Appendix).

4. The effect of one year of school time on cognitive skills across countries

This section presents the first part of our empirical analysis. We estimate the average effect of one year of school time for students in different age groups, on the performance in cognitive tests for the pooled samples of countries. This estimate can be interpreted as the effect of spending one year in school across countries, and might be considered as an international benchmark for gains in cognitive skills for specific age groups and subjects.

To obtain estimates of the average effect of one year of school time in all the selected countries from TIMSS we have pooled the data for each test (TIMSS 9, TIMSS 13) and estimated Equation (4). In this model we have also included country dummies and interactions of these dummies with a linear function of age and we have allowed the functional form of age to be different at either side of the cut-off for each country. The critical assumption in applying this model is that students near the cut-off date are very similar on observed and unobserved characteristics. To investigate this assumption we compared the covariates of students born in the months around the cut-off date (see Table A.2 in the appendix). In addition, we performed balancing tests, in which observed characteristics are regressed on a dummy for being born at the left side of the cut-off and a function of age (Table A.3 in the appendix). For 9-year-olds we find that students on both sides of the cut-off are very similar. For 13-year-olds, however, we find a difference with respect to the educational level of the mother. This difference might be the result of sampling bias which we will address below.

Table 1 shows the estimation results based on Equation (4) for both subjects and age groups. In panel A and B we show the estimation results using the TIMSS achievement tests in math and science respectively for 9-year-olds and 13-year-olds. Columns (1) to (4) show the reduced form estimates from Equation (4). The odd columns only control for age, the even columns also control for gender, born in country of test, lives with mother/father, language of test spoken at home and number of books at home. These columns also report the first stage estimate from Equation (2) which can be interpreted as the proportion of students that is on track. We use two discontinuity samples around the cut-off date: ± 3 months and ± 6 months. Columns (5) to (12) show the result from various sensitivity analyses. All sensitivity analyses use the sample of students born six month before or after the cutoff date, and include all controls like in column (4). Columns (5) and (6) respectively include a quadratic or cubic function of month of birth. Columns (7) to (9) focus on the sample of countries for which the exact date of birth is available; 17 countries for 9-year-olds, 30 countries for 13-year-olds.[19] Column (7) uses a linear specification of month of birth like in column (4), column (8) uses the exact date of birth as assignment variable in a linear specification, and column (9) excludes students born three days before or after the cutoff date. Column (10) uses the same sample of countries as used by Bedard and Dhuey (2006). The last two columns address the issue of sampling bias. In column (11) the estimation sample only includes countries in which most students are on track; countries with a first stage estimate (Equation (2)) of at least 0.75. For the 9-year-olds the sample includes 9 countries; for the 13-year-olds 12 countries are included. Column (12) shows estimates that have been adjusted for sampling bias. The sampling bias is approximated by using data of students that are one year younger or one year older in the two adjacent grades. The standard errors of the estimates are adjusted for using plausible values (Von Davier et al. 2009), which causes a slight increase.

The estimates in column (1) to (4) of Panel A of Table 1 show that one year of time in school increases performance of 9-year olds between 25 and 27 points in math and between 18 and 20 points in science, which is between 0.2 and 0.3 standard deviations of test scores (a standard deviation of test scores is 100 points). The estimates are precise, and robust to the discontinuity sample. The inclusion of controls only slightly reduces the estimated effects, which confirms that students born around the cut-off date are quite similar in observed characteristics. Columns (5) to (12) show the results from various sensitivity analyses. Columns (5) and (6) show that including a quadratic or cubic function of age, measured by year and month of birth, slightly increases the estimated effects. The estimates in columns (7) to (9) test the sensitivity of the results for using the exact date of birth as assignment variable and for a potential non-randomness of births around the cutoff. The estimates are very similar if we include the exact date of birth (column (8)) or exclude students born three days before or after the cutoff date (column (9)). Column (10) shows the estimation results for the sample of countries used by Bedard and Dhuey (2006). This sample is more restrictive and also excludes countries with a rolling admission. The estimated effect for this sample remains similar to the results in the other columns. Columns (11) and (12) aim to assess the sensitivity of the results to sampling bias. The estimated effects are 0.03 to 0.04 standard deviations higher than those in column (4), which suggests that for 9-year-olds

[19] The number of observations for Greece has reduced substantially due to missing values of the day of birth.

sampling bias is small. For the sample of countries used for the estimations in column (11) the difference between the lower bound estimates and the unbiased estimates is 0.01 standard deviations (not shown in Table 1). Hence, the lower bound estimates are probably not very different from the unbiased effects of one year of time in school.

The estimates in column (1) to (12) should be interpreted as the effect of one additional year in school at the age of 9. As in previous studies, we can also attempt to estimate the effect of time in grade on cognitive skills by using an IV-approach. The IV-estimates can be obtained as the ratio of the reduced form estimates and the first stage estimates shown in column (1) to (4). If the IV-assumptions hold we would find that one year of time in grade increases the scores in math and science by approximately 40 and 30 points respectively. As mentioned above, this approach assumes that grade retention or acceleration has no effect on cognitive skills.

Panel B shows the effects of time in school for 13-year-olds. For the lower bound estimates in columns (1) to (10) we find that one year of time in school increases performance by 6 to 8 points in math and by 11 to 12 points in science. The estimates of the lower bound effect are robust to the discontinuity sample and to various sensitivity tests. The estimated effects in columns (11) and (12), which are approximations of sampling bias, are larger than the lower bound estimates. The proportion of 13-year-olds that are not on track is larger than the proportion of 9-year-olds, which explains the increase in the difference between the lower bound estimates and the approximations of the unbiased effects. For the sample of countries used for the estimation in column (11) the difference between the lower bound estimates and the unbiased estimate is 0.02 standard deviations (not shown in Table 1). Hence, for these 12 countries sampling bias is likely to be quite small.

The cross-country estimates of time in school yield three important findings. First, a year of school time matters for the performance of all age groups in math and science. Across countries a year of school time increases performance in cognitive tests with 0.2 to 0.3 standard deviations for 9-year olds and with 0.1 to 0.2 standard deviations for 13-year-olds. These effects are consistent with the results of previous studies based on credible research designs (Gormley and Gayer (2005), Gormley et al. (2005), Hansen et al. (2004), Cascio and Lewis (2006) and Berlinski et al. (2009)). Second, an additional year of time in school matters more at the age of 9 than at the age of 13. Hence, the effect of time in school seems to reduce with age. This might indicate that later grades add less to the knowledge base or that the tests do a poorer job at measuring the full range of skill differences. Third, the difference between the lower bound estimates and the unbiased estimates seems to increase with the proportions of student that are not on track. For 9-year-olds the lower bounds estimates are likely to be quite close to the unbiased effects. For 13-year-olds this also holds for the sample of 12 countries with high proportions of students that are on track. The difference between the lower bound estimates and the approximations of the unbiased effects are larger for countries in which substantial proportions of students are not on track.

5. International differences in gains in cognitive skills

The second step in the empirical analysis is to investigate differences between countries. Which countries produce the largest effects of one year of school time on performance in cognitive tests?

Differences in achievements of 9-year-olds between countries

We start by analyzing the achievement of 9-year-olds. Column (1) of Table 2 shows the reduced form estimates (RF) of the gain in math skills caused by one year of additional school time. This estimate can be interpreted as the effect of spending one year in school in a specific country. The countries are ranked with respect to this estimate. We observe that the education systems of Norway and Singapore have produced the highest gain in achievement in math for 9-year-olds; the lowest gain in achievement in math has been produced by the education systems of New Zealand and Thailand. Column (2) shows estimates of the effect of one year of school time which are adjusted for sampling bias. Column (3) shows the mean level score of the highest of the adjacent grades for each country. Singapore and South-Korea have the highest scores, whereas Iceland and Iran have the lowest level scores in math in the upper grades. We call these average scores the country level scores and in the next section we will compare them with the estimates of the effect of one year of school time. Columns (4), (5) and (6) show the reduced form estimate, the estimate that is adjusted for sampling bias, and the mean level score of the upper grade for the science test. Column (7) shows the first stage estimates (FS) of the effect of being born on the left side of the cutoff date on the grade level. This estimate indicates to which extent a country keeps students on track. For instance, in Singapore, Iceland, Japan and England most students move through the education system in line with the prediction based on the school entry rule. All models control for age (in months) separately specified for both sides of the cut-off and use the sample of students born in the period between six month before and six month after the cutoff date. The standard errors of the estimates are adjusted for using plausible values.

The estimates of the lower bound effect of one year of school time in columns (1) and (4) show that there are large differences in the gains in cognitive skills between countries. The estimated effects differ between 0 and 0.4 standard deviations of the test scores. High gains in achievement for both tests are found for Norway, Singapore and Iceland. On the other hand, we also find very low gains in cognitive skills; in five countries the estimated effects do not significantly differ from zero. Hence, one year of time in school does not yield a gain in cognitive skills in these countries. The country specific estimates remain quite similar when we use the exact birth date as assignment variable and exclude children born very close to the cutoff date (Table A.5a in the appendix). In general, countries with high gains in achievement in math also achieve high gains in science: the correlation between the estimates in columns (1) and (4) is 0.9, which is significant at the 1%-level. For most countries the lower bound estimates are very close to the estimates that are adjusted for sampling bias (columns (2) and (5)). However, for several individual countries, in particular Iran, Austria, Latvia and Thailand, the adjusted score is substantially larger than the lower bound estimates. In these countries the proportion of students that is not on track is relatively high, as indicated by the first stage equation in column (5). The ranking of countries based on the adjusted estimates is quite similar to the ranking based on the lower bound estimates (the correlation between these estimates for both math and science is 0.94, which is significant at the 1% level).

Differences in achievements of 13-year-olds between countries

The sample of countries that can be used for estimating the effect for 13-years-olds consists of 34 countries. Table 3 shows the estimation results. Again we observe large differences between countries. The lower bound estimates of the gains in cognitive skills (columns (1) and (4)) are substantially smaller for 13-years-olds than for 9-year-olds, which is in line with the results from the previous section. Singapore achieves the highest gains in cognitive skills in both subjects; the results for science are remarkably far ahead of all other countries. Another remarkable finding for the 13-year-olds is the large number of countries for which the lower bound estimate of the effect of one year of school time is statistically insignificant. This suggests that in these countries one additional year of school time does not matter for the performance on the TIMSS math or science tests. More countries generate a statistically significant effect in science than in math. The estimates are quite similar when we use the exact birth date as assignment variable and exclude children born very close to the cutoff date (Table A.5b in the appendix). Again we observe that countries with high gains in achievement in math also achieve high gains in science: the correlation between the estimates in columns (1) and (4) is 0.73, which is significant at the 1%-level. For countries with high proportions of students that are on track, indicated by high first stage estimates, we observe that the lower bound estimates are quite similar to the estimates that are adjusted for sampling bias (columns (2) and (5)). However, for countries with a relatively low first stage estimate the adjusted scores can be substantially higher than the lower bound estimates.

We have also investigated whether countries that have high achievement gains for 9-year-olds also have high achievement gains for 13-year-olds, and whether countries with low gains for 9-year-olds also have low gains for 13-year-olds. We find a correlation of 0.51 for the reduced form estimates for math and a correlation of 0.34 for the reduced form estimates for science. The correlations for the math tests are statistically significant. This implies that education systems that are more effective in producing cognitive skills for 9-year-olds are also more effective in producing cognitive skills for 13-year-olds.

6. Do gains scores and level scores yield a consistent assessment of education systems?

This section shows the results of the third part of our empirical analysis. We compare the estimates of the gains in cognitive skills with the levels of the cognitive skills as currently used for the benchmarking of education systems. Gains scores and level scores can both be considered as measures of the performance of an education system. An interesting question is whether the rankings of the estimates of gains in cognitive skills in Tables 2 and 3 are consistent with the ranking based on the level of the test scores. On the one hand we would expect a positive correlation between gain and level scores because the level scores are the sum of all gains in cognitive skills caused by time in and out of school. On the other hand gain scores and level scores might differ because both measures have limitations. Level scores do not isolate the contribution of time in school from the contribution of time out of school. High level scores could mask a low performing education system if the conditions outside schools are favorable for learning, which means a high contribution of time out of school to student performance. Low level scores might also be misleading about the performance of the education systems if the conditions outside school are unfavorable for learning.[20] Gain scores isolate time in school from time out of school but only measure the effect of one year in school. It follows that low gain scores could be the result of a low quality of education but also the result of the timing of the curriculum. The latter, however, seems to contrast with the way the TIMSS tests have been developed (see Section 3).

For investigating whether the two measures show a consistent ranking of countries we compare the reduced form estimates of the gain scores with the mean upper grade scores.[21] We use the mean of the upper grades scores, instead of the mean of the scores from both adjacent grades, because the gain scores measure the effect of time in school between the lowest and the highest grade and, therefore, are included in the mean upper grade scores. In Figures 2, 3 and 4 we have plotted the mean upper grade scores for the different age groups and subjects on the vertical axis against the estimates of the gains in achievement on the horizontal axis. Figure 2 shows the results for the 9-years-olds in math and science. Figure 3 shows the results for the 13-years-olds. In addition, Figure 4 plots the level score at age 13 against the gains in cognitive

[20] A similar concern arises when the performance of schools is compared. School with low level scores might actually have high 'value added' (Figlio & Loeb 2011).

[21] See for the mean upper grade scores TIMMS 9 http://timssandpirls.bc.edu/timss1995i/TIMSSPDF/P1HiLite.pdf and for TIMMS 13 http://timssandpirls.bc.edu/timss1995i/TIMSSPDF/P2HiLite.pdf.

skills of 9-year-olds. We have included axes at the median level of gains scores and level scores in all figures which generate four quadrants of the performance of education systems: low level – low gain; low level – high gain; high level –low gain; high level – high gain.

In Figure 2 we observe no association between the mean upper grade scores and the gains in cognitive skills for both subjects. For math we observe a large variation in gains in cognitive skills for countries below the median level scores. Hence, countries with a low level score are not only observed in the low gain quadrant but also in the high gain quadrant. For instance, Norway has the highest gain of all countries but also a level score below the median level. The gain scores for countries above the median level scores are less dispersed and more concentrated around the median gain scores. For science we observe a more even distribution of countries across the four quadrants of performance. A similar pattern is found for the 13-year-olds (Figure 3).[22] We observe no association between mean level scores and gains in cognitive skills. Countries with high level scores are not consistently found in the top of the ranking based on the gain scores. Similarly, countries with low level test scores do not systematically have low gains in achievement. Singapore can be considered as a (positive) outlier for both subjects in Figure 3. Figure 4 shows the association between the gain score at the age of 9 and the level score at the age of 13. We observe that the country rankings for the 13-year olds in TIMSS are not related with the gains in achievement for 9-year olds.

The results from Figures 2 and 3 have been summarized in Table 4. This table shows correlations between the estimates of the gains in achievement and the country level scores by test and age group. The table also includes the results of sensitivity analyses with respect to the size of the discontinuity samples (Panel A), the non-random timing of birth (Panel B), and the inclusion of countries not used in the sample of Bedard & Dhuey (2006) and sampling bias (Panel C). The main finding of Table 4 is that the correlation between the gain scores and the country level scores is close to zero and statistically insignificant for all subjects and age groups. This result is found for the main results as shown in Figure 2 and 3 (middle panel of Panel A) and is robust to a series of sensitivity analyses as shown in the other panels of Table 4. In panel C we find higher correlations for 13-year-olds in the sample of countries with a first stage estimates above 0.75. However, this correlation is completely driven by the large gains of

[22] Figure 3 also distinguishes countries with and without an external exit exam which is relevant for the analysis in the next section.

Singapore. The estimates that are adjusted for sampling bias suggest a negative correlation between gain scores and level scores.

The low correlations imply that country level scores and country gain scores often tell different stories about the performance of education systems. Hence, countries that are top ranked in the test are not necessarily characterized by high gains in achievement, and low ranked countries are not necessarily characterized by low gain scores. The current use of the outcomes of international cognitive tests in educational policy focuses on the ranking along the vertical axis. The figures in this section show that these rankings hide large variation in gains in cognitive skills between countries illustrated by the variation along the horizontal axis. As such, gain scores add a second dimension for assessing the performance of education systems. For educational policy it seems useful to focus not only on the ranking along the vertical axis but also take the horizontal axis into account, for instance by looking at the four quadrants of performance. For countries in the low level – low gain quadrant or in the high level – high gain quadrant the assessment of the performance seems clear. But for countries in the other two quadrants, the assessment of the performance of education system is less clear. For example, the below median level score of Norway can be interpreted as a signal of low quality education. However, the high gain scores tell a different story and suggest that other factors are likely to explain the low level scores.[23] For countries in these two quadrants a mere focus on the ranking along the vertical axis might yield misleading information for educational policy.

[23] It might be speculated that the relatively late school starting age in Norway lowers the level scores.

7. Investigating the determinants of international differences in cognitive skills

The estimated effects of the effect of time in school are also of interest to the literature that uses international cognitive tests for investigating the determinants of international differences in educational achievement (Hanushek & Woessmann 2011). This literature investigates whether differences in school inputs or institutions, such as school accountability and autonomy, central exams, competition between schools or tracking, can explain the large differences in achievements of students between countries. Most studies in this recent literature have focused on identifying cross-country associations.[24] However, due to the complex nature of the production of human capital it remains unclear whether these associations can be interpreted as causal effects. Therefore, it has been argued that 'further exploration of quasi-experimental settings in the international data should be high on the agenda' (Hanushek & Woessman, 2011). This paper provides such a quasi-experimental setting and the previous section shows that assessments of the performance of education systems based on level scores might be different from assessments based on gain scores. The approach that is applied in this paper might also yield new opportunities for investigating the determinants of international differences in education achievements between countries. Whereas the current literature tries to relate differences in student outcomes to differences in input factors or institutions, it is also possible to relate differences in gains in achievement to differences in input factor or institutions. The advantage of our approach is that it isolates the effects of time in school from the effects of time out of school.

To illustrate these opportunities, we re-examine the impact of curriculum-based external exit exam systems (CBEEE). Previous studies have investigated the effects of external exit-exams and provide a consistent picture about the beneficial effect of external exit-exams. The effects might be even larger than a whole-grade level equivalent, between 0.2 to 0.4 standard deviations of the respective tests (Hanushek & Woessmann 2011). These results have also been found for the 1995 TIMSS math and science achievements of 13-year-olds in a study that uses country level data (Bishop 1997) and in a study that uses micro-level data (Woessman 2003). We conduct a similar analysis using gains in achievements. Table 5 shows the estimations results; the left panel shows the results using country level-data, the right panel shows the results using micro-level data.

[24] Several studies have used a quasi-experimental design, see Section 1.

The first column of panel A of Table 5 replicates the estimates from Bishop (1997). The mean upper grade scores of 34 countries are regressed on a dummy for having a curriculum-based external exit exam. The estimates show that countries that have an external-exit exam score 29 points higher on math and 33 points higher on science tests. Bishop (1997) reports similar results (23 points for math and 34 points for science) in models with more controls. In columns (2) and (3) the estimated gains in achievement instead of the mean upper grade scores are used as dependent variable. Column (2) uses the lower bound estimate of the gains in cognitive skills; column (3) uses estimates that are adjusted for sampling bias. The estimates show that countries that have an external exit-exam do not produce higher gains in achievement than countries that do not have an external exit exam; the point estimates are negative and in column (3) we even find statistically significant negative effects. The right panel of Table 7 uses micro-level data; column (4) uses a specification as in Woessman (2003). The models in columns (5) and (6) are based on Equation (4), and include dummies for having an external exit exam and an interaction of time in school (grade) with the external exit exam (CBEEE) like in Equation (5):

$$(5) \qquad Y_i = \lambda_0 + \lambda_1 S_i + \lambda_2 CBEEE + \lambda_3 S_i * CBEEE + f(birthmonth_i - C) + \lambda_4 X_i + \varepsilon_i$$

The estimates show that students in countries with external exit exams score 25 (30) points higher on math (science) and the estimated effects are statistically significant. Woessman (2003) also finds statistically significant positive effects, but these effects are smaller after including an extensive set of family background and school-input controls (11 (16) points for math (science)). However, we do not find a positive effect of external exit-exams on the gains in achievement in the last year before the test. This result can also be observed in Figure 3 which distinguishes countries with and without an external exit exam. We do not observe that countries with external exams have higher gains in achievement than countries without an external exit exam.

A limitation of the gain score approach is that it only refers to the achievements in the last year before the test. Hence, it is possible that we fail to find an effect of external exit-exams because they only affect the results in earlier years in school. The TIMSS test scores of 9-year-olds provide an opportunity to observe what happened in one of the earlier years. The estimation results, based on the same models, are shown in panel B of Table 5. Again we observe substantial positive effects of external exit exams in columns (1) and (4). The estimates indicate that central exit-exams increase test scores in math by 51 points and in science by 36 points. However, the estimated effects become statistically insignificant when we focus on gains in achievement. This means that we do not find higher gains in achievement during two school years for students in countries that have an external exit exam compared to students in countries that do not have external exit exams. Although this finding does only relate to two school years, which is one quarter of the total amount of time in school, it raises concerns about unobserved differences between countries that have external exit exams and countries that do not have external exit exams in the studies that previously used the 1995 TIMSS data for estimating the effect of external exit-exams.

In sum, our re-examination of previous results on the effect of external exit exams using the 1995 TIMSS data shows that results based on level scores might differ from the results based on gain scores. This illustrates that an approach that focuses on gains in achievement might offer new opportunities and insights for investigating the determinants of international differences in educational achievement.

8. Conclusions

This study applies a quasi-experimental approach for estimating the effect of one year of school time on the performance in international cognitive tests by exploiting the assignment of students to different grades based on school entry rules. This method produces estimates of gains in cognitive skills for students in different age groups in the year before the test for worldwide samples of countries and for individual countries. This method also enables a comparison of the performance of education systems based on gain scores instead of level scores.

We find that time in school on average matters for student performance in international cognitive tests. For the pooled sample of countries we find that a year of school time increases performance in cognitive tests with 0.2 to 0.3 standard deviations for 9-year olds and with 0.1 to 0.2 standard deviations for 13-year-olds. These effects are consistent with the results of previous studies based on credible research designs (Gormley and Gayer (2005), Gormley et al. (2005), Hansen et al. (2004), Cascio and Lewis (2006) and Berlinski et al. (2009)). We also find large differences in the estimated gains in achievement between countries for both subjects and age groups. Countries that achieve higher gains in cognitive skills for math also achieve higher gains in science. Moreover, countries with higher gains for 9-year-olds also have higher gains for 13-years-olds.

The sampling strategy of the TIMSS-project, which focused on two adjacent grades, might induce a downward bias for our estimates. Therefore, the main estimates should be interpreted as lower bound estimates. For 9-year-olds the lower bound estimates are probably quite close to the unbiased effect. However, for 13-year-olds the lower bound estimates will probably underestimate the gains in cognitive tests for countries in which a large proportion of students is not on track.

Remarkably, we find no association between the estimated gains in achievement and the level scores of countries. At all levels of test scores we observe countries with high achievement gains and countries with low gains in achievement. The lack of association has been found for both tests (math, science) and for both age groups. Hence, assessments of the performance of education systems based on the estimated gains in achievement often are inconsistent with performance assessments based on level scores. This inconsistency might be explained by limitations of both measures. Level scores do not distinguish between the contribution of time in school and the contribution of time out of school. The gain scores only refer to the gain in achievement in the year before the test. The inconsistency of the two measures implies that the benchmarking of education systems based on level scores might yield misleading information about the performance of education systems. Low levels of test scores, or declining trends in test scores, might not be the result of low performing education systems. High levels of test scores could mask low performing education systems. Using gain scores as an additional instrument for the assessment of the performance of education systems is likely to reduce the risk of providing misleading policy information.

The quasi-experimental approach for estimating gains scores used in this paper can also be important for investigating international differences in educational achievements. For instance, studies that use control strategies have consistently found that students perform better in countries with external exit exams. However, we do not find higher gains in achievements in countries with a central exit exam for 9-year-olds and 13-year-olds compared to countries that do not have central exit exams.

This study shows that time in school is important for acquiring cognitive skills and that there are large differences in the effects between countries. Estimates of the gains in achievement for separate countries provide a different assessment of the performance of education systems, and of the effect of specific elements of education systems, than level scores. The estimation of gain scores, which becomes possible when the current collection of international data is extended towards samples of students in adjacent grades, is likely to improve decisions on educational policies and could offer new opportunities for investigating the determinants of international differences in student achievement.

References

Angrist, Joshua. D., and Alan B. Krueger. (1991). "Does compulsory school attendance affect schooling and earnings?" *Quarterly Journal of Economics*, 106(4): 979–1014.

Almond, D and J. Currie, (2010), Human capital development before age 5, in: Eric A. Hanushek, Stephen Machin and Ludger Woessmann, *Handbook of Labor Economics,* Volume 4b.

Baltes, P.B. and G. Reinert, (1969), Cohort Effects in cognitive development as revealed by cross-sectional sequences, Development Psychology, 1(2): 169-77.

Bedard, Kelly, and Elizabeth Dhuey. (2006) "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects," *Quarterly Journal of Economics*, 121(4): 1437–1472.

Berlinski, S., Galiani, S., & Gertler, P. (2009). "The effect of pre-primary education on primary school performance," *Journal of Public Economics*, 93, 219–234.

Berlinski, Samuel, Sebastian Galiani, and Patrick J. McEwan. (2011) "Preschool and maternal labor supply: Evidence from a regression discontinuity design," *Economic Development and Cultural Change*, 59(2): 313-344.

Bishop, J. H, (1997), The effect of national standards and curriculum-based examinations on achievement, *American Economic Review*, 87 (2), 260-264.

Black, Sandra, Paul J. Devereux, and Kjell G. Salvanes. (2011) "Too Young to Leave the Nest? The Effects of School Starting Age," *Review of Economics and Statistics*, 93(2): 455–467.

Boardman, A.E., Murnane R. (1979) "Using panel data to improve estimates of the determinants of educational achievement," *Sociology of Education*, 52, 113-121.

Bound, J. and D.A. Jaeger, (2000), Do compulsory school attendance laws alone explain the association between quarter of birth and earnings? *Research in Labor Economics*, 19, 83-108.

Breakspear, S. (2012), "The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance", *OECD Education Working Papers*, No. 71, OECD Publishing. http://dx.doi.org/10.1787/5k9fdfqffr28-en

Buckles, K. and D. Hungerman, (2012), Season of birth and later outcomes: old questions, new answers, *Review of Economics and Statistics*, forthcoming

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

Cahan, S. & D. Davis, (1987), A between-grade-levels approach to the investigation of the absolute effects of schooling on achievement, *American Educational Research Journal*, 24(1): 1-12.

Cascio, E. U., & Lewis, E. G. (2006). “Schooling and the armed forces qualifying test: Evidence from school entry laws,” *Journal of Human Resources*, 41(2), 294–318.

Currie, J. (2001), Early Childhood Education Programs, *Journal of Economic Perspectives*, 15(2): 213-38.

Dickert-Conlin, Stacy and Amitabh Chandra. (1999) “Taxes and the Timing of Births,” *Journal of Political Economy*, 107(1): 161–177.

Dickert-Conlin, Stacy and Todd Elder. (2010) “Suburban Legend: School Cutoff Dates and the Timing of Births,” *Economics of Education Review*, 29(5): 826–841.

Dobkin, C. and F. Ferreira, (2010), Do school entry laws affect educational attainment and labor market outcomes? *Economics of Education Review*, 29(1): 40-54.

Figlio, D. and S. Loeb, (2011), School accountability, In Eric A. Hanushek, Stephen Machin and Ludger Woessmann (Eds.), *Handbook of the Economics of Education, Vol. 3, Amsterdam: North Holland*, 2011, pp. 383-421.

Fredriksson, Peter, and Björn Öckert. (forthcoming), “Life-cycle Effects of Age at School Start,” *Economic Journal*

Gormley, W. T., & Gayer, T. (2005) “Promoting school readiness in Oklahoma: An evaluation of Tulsa’s pre-K program,” *Journal of Human Resources*, 60, 533–558.

Hansen, K., Heckman, J., & Mullen, K. (2004) “The effect of schooling and ability on achievement test scores,” *Journal of Econometrics*, 121(1–2), 39–98.

Hanushek, E.A., Kain J.F., O’Brien, D.M., Rivkin, S.G, (2005), The market for teacher quality, NBER Working Paper 11154.

Hanushek, E.A., Rivkin, S.G., (2006), Teacher quality. In: Hanushek, E., Welch, F. (Eds.), *Handbook of Economics of Education*, vol 2. Elsevier.

Hanushek, E.A., Woessmann, L., (2008) “The role of cognitive skills in economic development,” *Journal of Economic Literature*, 46 (3), 607–668.

Hanushek, E.A., Woessmann, L., (2011) “The economics of international differences in educational achievement,” In Eric A. Hanushek, Stephen Machin and Ludger Woessmann (Eds.), *Handbook of the Economics of Education, Vol. 3, Amsterdam: North Holland*, 2011, pp. 89-200.

Harris, D.N., Sass, T.R., (2011) “Teacher training, teacher quality and student achievement,” *Journal of Public Economics*, 95, 798-812.

IEA, 2011, TIMSS and PIRLS—Informing Educational Policy for Improved Teaching and Learning, document.

Jacob, B. and Lefgren, L, (2009), The Effect of Grade Retention on High School Completion. American Economic Journal: Applied Economics. 1(3): 33-58.

Jürges, H, Schneider K. and F. Büchel, 2005, The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMSS Germany, *Journal of the European Economic Association*, 3 (5), 1134-1155.

Krueger, A.B. (1999). “Experimental estimates of education production functions,” *The Quarterly Journal of Economics*, 114 (2): 497-532.

Leuven, Edwin, Mikael, Lindahl, Hessel, Oosterbeek, and Dinand, Webbink. (2010) “Expanding schooling opportunities for 4-year-olds,” *Economics of Education Review*, 29: 319–328.

Lee, D., Lemieux T., (2010) “Regression Discontinuity Designs in Economics,” *Journal of Economic Literature* 48(2), 281-355.

Manacorda, M, (2012), The Cost of Grade Retention, *Review of Economics and Statistics*, 94 (2): 596–606.

Meghir, C., Rivkin, S., (2010) “Econometric methods for research in education,” In Eric A. Hanushek, Stephen Machin and Ludger Woessmann (Eds.), *Handbook of the Economics of Education, Vol. 3, Amsterdam: North Holland*, 2011, pp. 89-200.

McCrary, Justin, and Heather Royer. (2011) “The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth,” *American Economic Review*, 101(1): 158–195.

McEwan, Patrick J., Joseph S. Shapiro. (2008) “The benefits of delayed primary school enrollment: Discontinuity estimates using exact birth dates,” *Journal of Human Resources*, 43(1): 1–29.

Neal, D.A. and Johnson, W.R, (1996), The role of premarket factors in black-white differences, *Journal of Political Economy*, 104, 869-895.

Rivkin, S.G., Hanushek, E.A., and Kain, J.F. (2005). “Teachers, schools and academic achievement,” *Econometrica*,Vol. 73, No. 2: 417-458.

Rothstein, J. (2010), Teacher Quality in Educational Production: Tracking, Decay and Student Achievement, *Quaterly Journal of Economics*, 125(1): 129-174

Schwerdt & West M.R, (2012), The Effects of Test-based Retention on Student Outcomes over Time:
Regression Discontinuity Evidence from Florida, Working Paper.

TIMSS & PIRLS, (2011) TIMSS and PIRLS—Informing Educational Policy for Improved Teaching and Learning, International Study Center, http://timssandpirls.bc.edu/home/pdf/TP_Impact_Statement.pdf

Todd, P.E., Wolpin, K.I., (2003) "On the specification and estimation of the production function for cognitive achievement," *Economic Journal*, 113 (485), F3-33.

Von Davier, M., Gonzalez, E., & Mislevy, R.J. (2009). What are plausible values and why are they useful? IERI Monograph Series. Issues and Methodologies in Large-Scale Assessments, Vol. 2, 9-36.

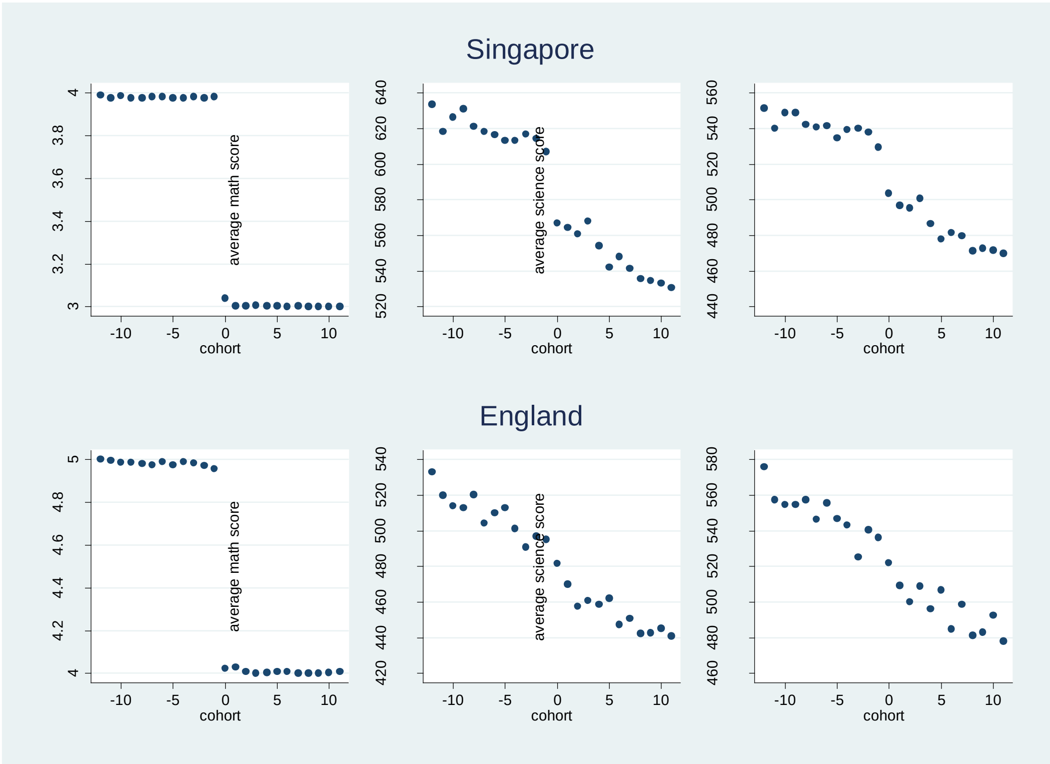
West, M.R. and L. Woessmann, (2010), Every Catholic child in a Catholic school: Historical resistance to state schooling, comtempory school competition, and student achievement, *Economic Journal*, 120 (546), F229-255.

Woessman, L, (2003), Schooling resources, educational institutions and student performance: The international evidence, *Oxford Bulletin of Economics and Statistics*, 65 (2), 117-170.

Woessmann, L. and M. West, (2006), Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS, *European Economic Review*, 50 (3): 695-736.

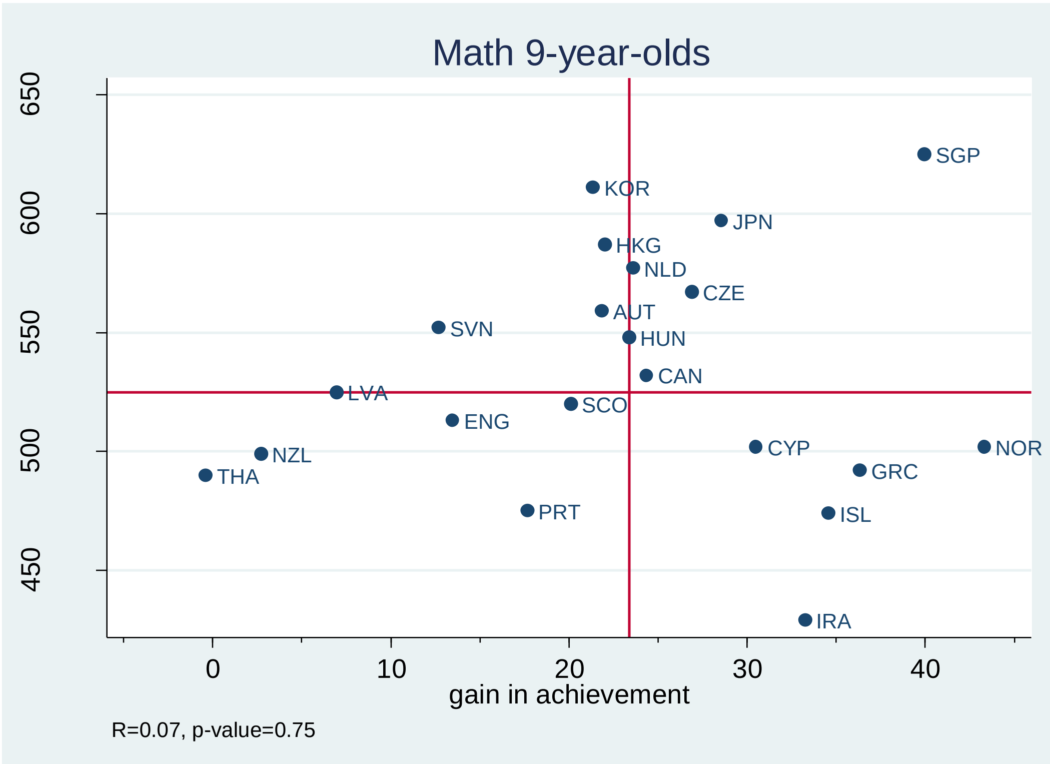
Figures

Figure 1: Grade level and math/science scores around the cut-off date for 9-year-olds from Singapore and England (TIMSS 1995)



Notes: Each dot represents a monthly average of the grade level or the test score. Students born in month 0 are born in the first month after the cutoff. See Table A.1 for the cutoff dates per country.

Figure 2: The association between country level scores and the effect of one year of school time on cognitive skills for 9-year-olds.



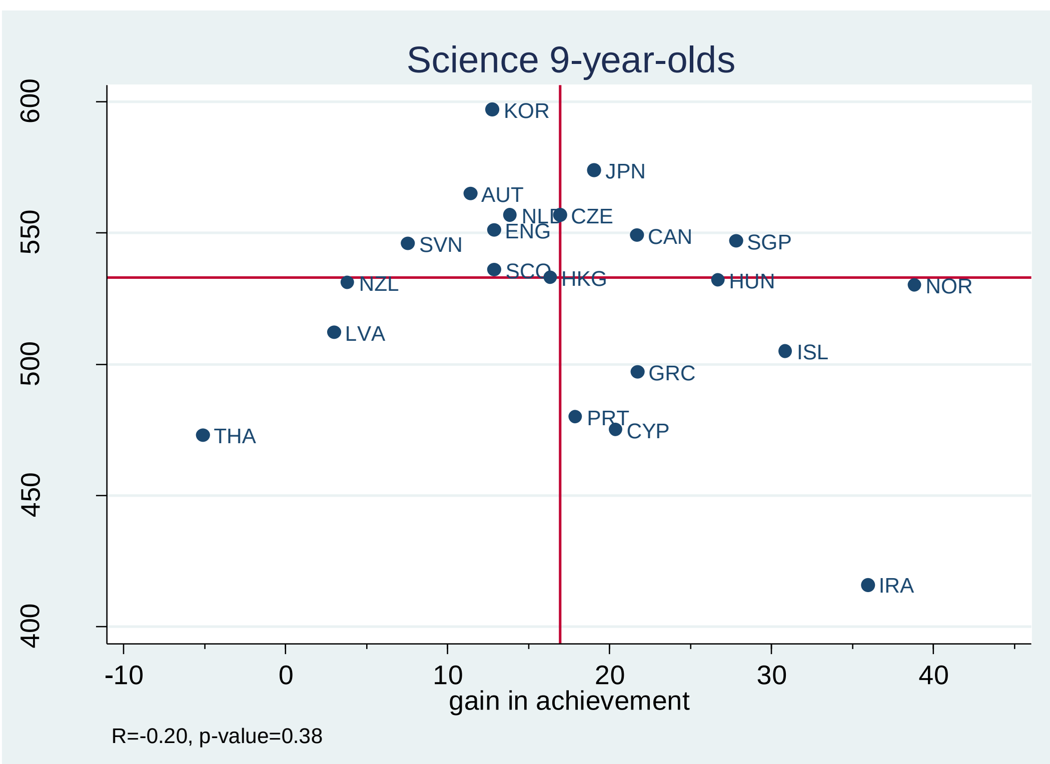
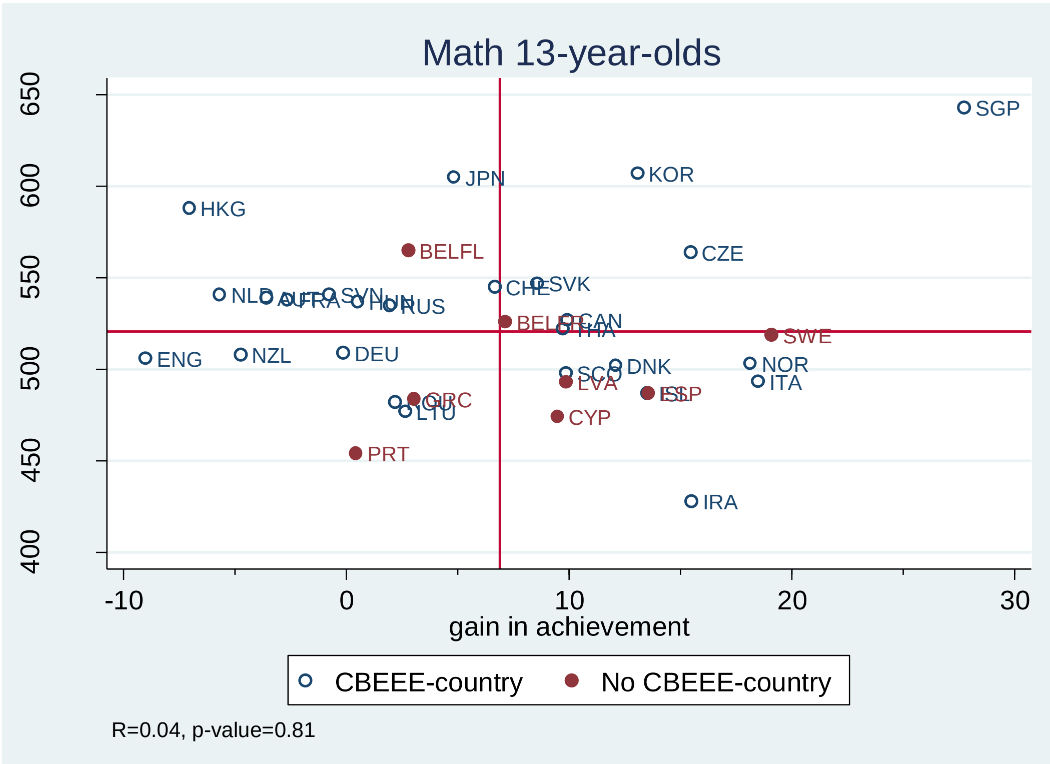


Figure 3: The association between country level scores and the effect of one year of school time on cognitive skills for 13-year olds.



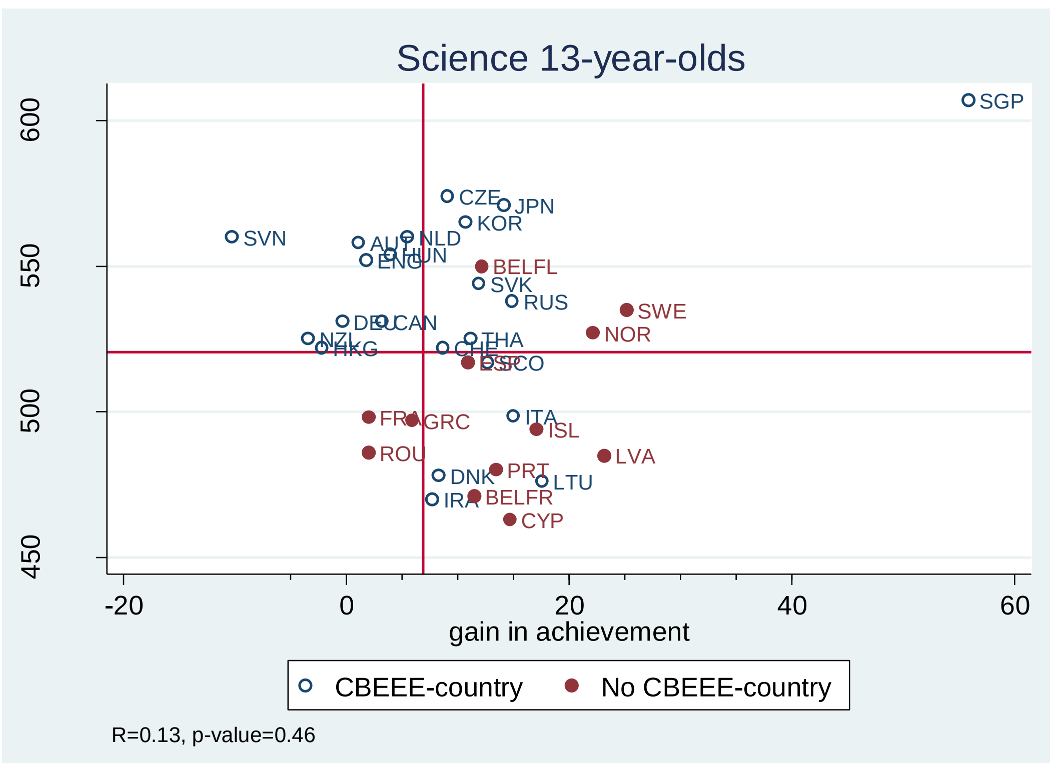
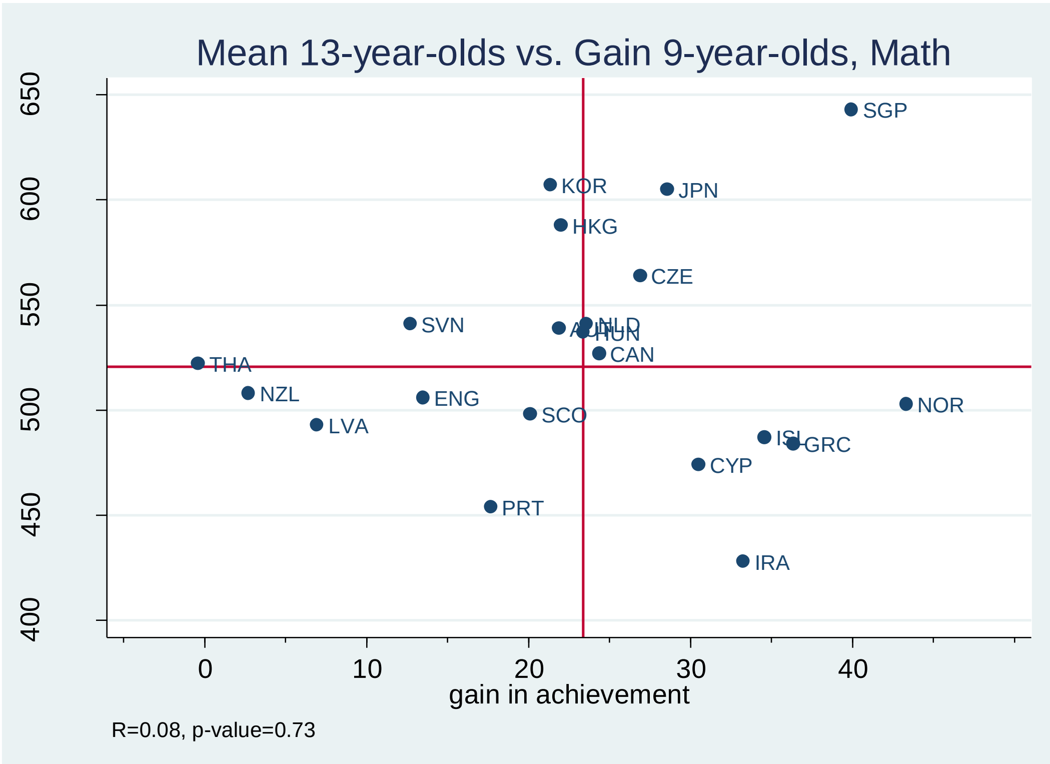


Figure 4: The association between country level scores of 13-year-olds and the effect of one year of school time on cognitive skills for 9-year-olds.



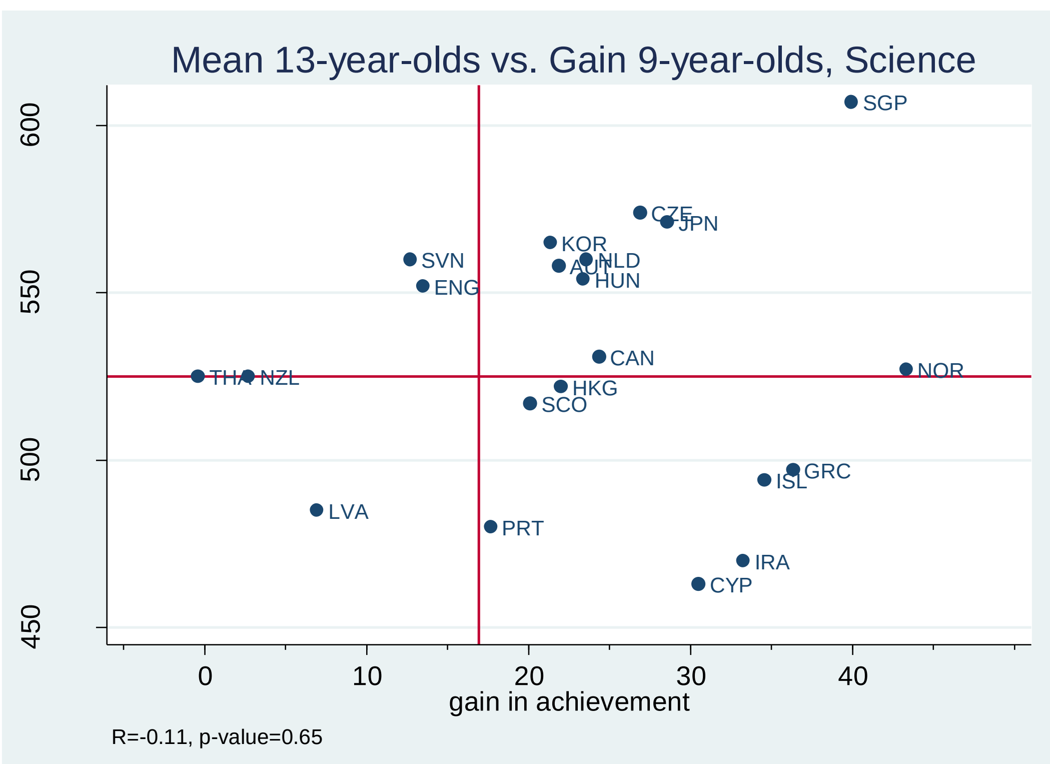


Table 1: Estimates of the effect of one year of school time by subject and age group based on pooled regression for all countries

	Reduced form estimates (OLS)				Sensitivity analysis using ± 6 months around the cutoff							
	± 3 months		± 6 months		Functional form		Day of birth sample [column (9) excludes 3 days around the cutoff]			BD-sample	Countries with first stage>0.75	Sampling bias adjustment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: TIMMS 9 (21 countries)												
math	25.8	25.1	27.0	26.4	29.9	34.5	27.5	27.9	28.2	23.8	29.7	30.4
	(2.2)	(2.2)	(1.5)	(1.4)	(2.1)	(2.8)	(1.5)	(1.4)	(1.5)	(1.9)	(2.0)	(1.4)
N	33615	33615	66803	66803	66803	66803	47782	47782	47001	31874	29861	68869
Coefficient first stage	*0.66*	*0.65*	*0.70*	*0.70*								
science	18.9	18.1	20.0	19.4	21.3	23.2	19.1	19.5	19.4	18.3	21.8	23.5
	(2.1)	(2.0)	(1.4)	(1.4)	(2.0)	(2.6)	(1.6)	(1.5)	(1.6)	(2.1)	(2.1)	(1.3)
N	33615	33615	66803	66803	66803	66803	47782	47782	47001	31874	29861	68869
Coefficient first stage	*0.66*	*0.65*	*0.70*	*0.70*								
Panel B: TIMMS 13 (34 countries)												
math	6.0	6.4	7.0	7.2	7.1	5.7	6.6	6.7	6.5	7.8	10.3	12.3
	(1.7)	(1.6)	(1.1)	(1.0)	(1.6)	(2.4)	(1.2)	(1.5)	(1.2)	(1.3)	(1.7)	(1.0)
N	51908	51908	104316	104316	104316	104316	85161	85161	83811	55651	36760	109942
Coefficient first stage	*0.56*	*0.56*	*0.62*	*0.62*								
science	10.9	11.3	10.8	11.0	11.9	11.4	11.7	11.6	11.6	10.1	17.6	16.4
	(1.5)	(1.4)	(1.0)	(0.9)	(1.5)	(2.3)	(1.0)	(1.0)	(1.0)	(1.5)	(1.8)	(0.9)
N	51908	51908	104316	104316	104316	104316	85161	85161	83811	55651	36760	109942
Coefficient first stage	*0.56*	*0.56*	*0.62*	*0.62*								
Birth month controls	linear	linear	linear	linear	square	cubic	linear	no	no	linear	linear	linear
Birth day controls	no	no	no	no	no	no	no	linear	linear	no	no	no
Additional controls	no	yes	no	yes	yes	yes	yes	yes	yes	yes	yes	yes

Notes: All models include country dummies and interactions of these dummies with the polynomial in birth month/day that differs at either side of the cutoff. The models in columns (2), (4), and (5)-(12) include additional controls. For 9-year-olds we control for gender, born in country of test, lives with mother/father, language of test spoken at home and number of books at home. For 13-year-olds we additionally control for parental education. These variables were not available 9-year-olds. Standard errors in parentheses are adjusted for using plausible values.

Table 2. Reduced form estimates of the gain in cognitive skills and mean upper grade by country for 9-year-olds

		Math			Science				
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ranking	Country	gain	corr. gain	mean	gain	corr. gain	mean	first stage	N
1	Norway	43.3 (7.4)	43.1 (7.4)	502	38.8 (8.5)	38.5 (8.5)	530	0.89 (0.02)	2133
2	Singapore	40.0 (5.2)	41.8 (5.2)	625	27.8 (4.9)	29.7 (4.9)	547	0.96 (0.01)	6986
3	Greece	36.3 (6.7)	38.0 (6.7)	492	21.7 (6.5)	23.2 (6.5)	497	0.89 (0.02)	2981
4	Iceland	34.6 (9.5)	35.9 (9.5)	474	30.9 (8.4)	32.4 (8.4)	505	0.96 (0.01)	1702
5	Iran	33.3 (7.1)	44.8 (6.6)	429	35.9 (6.8)	48.9 (6.4)	416	0.56 (0.04)	2716
6	Cyprus	30.5 (6.4)	33.9 (6.4)	502	20.4 (5.8)	23.2 (5.8)	475	0.79 (0.02)	3230
7	Japan	28.6 (5.4)	28.3 (5.4)	597	19.1 (5.3)	18.9 (5.3)	574	0.94 (0.01)	4343
8	Czech Republic	26.9 (6.9)	30.3 (6.9)	567	16.9 (6.5)	20.5 (6.5)	557	0.43 (0.03)	3108
9	Canada	24.4 (5.1)	29.9 (4.9)	532	21.7 (4.8)	28.7 (4.7)	549	0.70 (0.02)	7436
10	Netherlands	23.6 (6.8)	29.3 (6.9)	577	13.9 (6.7)	17.3 (6.6)	557	0.48 (0.04)	2241
11	Hungary	23.4 (7.8)	29.2 (7.8)	548	26.7 (7.0)	32.2 (7.0)	532	0.41 (0.04)	2743
12	Hong Kong	22.0 (5.6)	26.6 (5.5)	587	16.3 (5.0)	21.0 (4.9)	533	0.69 (0.02)	3851
13	Austria	21.9 (7.4)	27.9 (7.4)	559	11.4 (8.1)	17.6 (8.1)	565	0.55 (0.03)	2315
14	South-Korea	21.3 (6.5)	25.4 (6.2)	611	12.8 (6.4)	17.2 (6.1)	597	0.66 (0.03)	2636
15	Scotland	20.1 (6.6)	21.1 (6.6)	520	12.9 (7.2)	14.0 (7.2)	536	0.77 (0.03)	3089
16	Portugal	17.7 (7.3)	22.6 (7.3)	475	17.9 (8.9)	24.9 (8.9)	480	0.79 (0.03)	2310
17	England	13.5 (7.1)	13.6 (7.1)	513	12.9 (9.2)	13.4 (9.2)	551	0.93 (0.02)	3087
18	Slovenia	12.7 (7.3)	15.9 (7.3)	552	7.6 (7.1)	10.3 (7.1)	546	0.53 (0.03)	2484
19	Latvia	6.9 (7.8)	14.9 (7.4)	525	3.0 (9.0)	9.0 (8.5)	512	0.24 (0.05)	2116
20	New Zealand	2.7 (7.6)	2.3 (7.5)	499	3.8 (8.4)	3.6 (8.4)	531	0.40 (0.04)	2459
21	Thailand	-0.4 (6.2)	14.6 (5.7)	490	-5.1 (6.0)	6.9 (5.5)	473	0.41 (0.05)	2837

Notes: 21 out of 26 participating countries have been included. See section 3 for the exclusion of five countries. The estimation sample consists of students born in the period 6 months before and after the cut-off date. The countries in grey are from the sample used by Bedard & Dhuey (2006). Standard errors in parentheses are adjusted for using plausible values.

Table 3. Reduced form estimates of the gain in cognitive skills and mean upper grade by country for 13-year-olds

		Math			Science				
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ranking	Country	gain	corr. gain	mean	gain	corr. gain	mean	first stage	N
1	Singapore	27.8 (4.6)	27.8 (4.6)	643	55.9 (5.9)	56.6 (5.9)	607	0.96 (0.01)	3567
2	Sweden	19.1 (5.4)	22.0 (5.4)	519	25.2 (5.5)	28.3 (5.5)	535	0.91 (0.01)	3403
3	Italy	18.5 (8.3)	29.4 (8.1)	493	15.0 (8.0)	24.7 (7.8)	498	0.71 (0.03)	2186
4	Norway	18.2 (6.2)	18.5 (6.1)	503	22.1 (6.7)	22.8 (6.7)	527	0.84 (0.02)	2751
5	Iran	15.5 (7.1)	24.9 (6.5)	428	7.7 (6.9)	14.1 (6.3)	470	0.36 (0.04)	2495
6	Czech Republic	15.5 (5.6)	20.1 (5.5)	564	9.1 (5.6)	13.2 (5.5)	574	0.52 (0.02)	3248
7	Spain	13.6 (5.7)	27.0 (5.5)	487	10.9 (5.9)	23.4 (5.6)	517	0.73 (0.03)	3157
8	Iceland	13.5 (7.2)	13.9 (7.2)	487	17.1 (7.4)	17.5 (7.4)	494	0.94 (0.02)	1834
9	South-Korea	13.1 (6.5)	14.6 (6.5)	607	10.7 (6.2)	12.0 (6.1)	565	0.82 (0.02)	2912
10	Denmark	12.1 (7.3)	13.8 (7.3)	502	8.3 (8.2)	10.5 (8.2)	478	0.51 (0.04)	2096
11	Canada	9.9 (3.6)	16.0 (3.6)	527	3.2 (4.4)	9.3 (4.2)	531	0.55 (0.03)	7733
12	Latvia	9.9 (7.4)	17.7 (7.2)	493	23.2 (7.6)	33.1 (7.4)	485	0.57 (0.04)	2334
13	Scotland	9.9 (6.6)	9.6 (6.5)	498	12.7 (7.1)	13.1 (7.1)	517	0.75 (0.03)	2824
14	Thailand	9.7 (4.9)	14.9 (4.3)	522	11.1 (3.8)	18.6 (3.4)	525	0.46 (0.04)	5232
15	Cyprus	9.5 (8.1)	16.2 (8.0)	474	14.7 (7.8)	21.2 (7.8)	463	0.80 (0.02)	2837
16	Slovak Republic	8.6 (5.5)	12.5 (5.5)	547	11.9 (5.8)	15.7 (5.8)	544	0.79 (0.02)	3475
17	Belgium (French)	7.1 (6.8)	20.2 (6.5)	526	11.5 (7.7)	26.9 (7.4)	471	0.55 (0.04)	1872
18	Switzerland	6.7 (4.3)	19.0 (4.0)	545	8.7 (5.5)	21.2 (5.0)	522	0.30 (0.04)	3727
19	Japan	4.8 (4.6)	5.0 (4.6)	605	14.1 (4.5)	14.1 (4.5)	571	0.98 (0.01)	5158
20	Greece	3.0 (6.3)	11.2 (6.2)	484	5.9 (6.8)	13.0 (6.7)	497	0.75 (0.02)	3543
21	Belgium (Flemish)	2.8 (4.8)	7.1 (4.8)	565	12.2 (5.4)	17.2 (5.3)	550	0.77 (0.03)	2622
22	Lithuania	2.6 (7.3)	7.5 (7.3)	477	17.6 (7.5)	23.5 (7.5)	476	0.55 (0.04)	2590
23	Romania	2.2 (6.3)	8.8 (6.0)	482	2.0 (6.8)	7.6 (6.5)	486	0.16 (0.03)	3504
24	Russia	1.9 (5.7)	4.8 (5.6)	535	14.9 (6.6)	19.2 (6.5)	538	0.47 (0.03)	3855
25	Hungary	0.5 (5.8)	3.2 (5.9)	537	4.0 (6.0)	6.3 (6.0)	554	0.30 (0.03)	2887
26	Portugal	0.4 (5.8)	18.8 (5.6)	454	13.4 (7.5)	34.5 (7.2)	480	0.62 (0.03)	2579
27	Germany	-0.1 (6.4)	9.8 (6.2)	509	-0.4 (7.4)	11.6 (7.1)	531	0.36 (0.04)	2447
28	Slovenia	-0.8 (6.3)	4.6 (6.2)	541	-10.3 (6.5)	-6.4 (6.4)	560	0.43 (0.03)	2688
29	France	-2.6 (6.7)	21.0 (6.3)	538	2.0 (6.6)	25.3 (6.2)	498	0.43 (0.04)	2159
30	Austria	-3.6 (6.0)	4.8 (5.9)	539	1.1 (6.8)	10.9 (6.7)	558	0.53 (0.03)	2569
31	New Zealand	-4.7 (6.0)	-4.6 (6.0)	508	-3.4 (6.7)	-3.2 (6.7)	525	0.29 (0.04)	3432
32	Netherlands	-5.7 (6.5)	4.1 (6.5)	541	5.5 (7.2)	14.0 (7.1)	560	0.39 (0.04)	1770
33	Hong Kong	-7.1 (6.3)	-2.5 (6.1)	588	-2.2 (6.7)	2.7 (6.5)	522	0.59 (0.03)	2996
34	England	-9.0 (9.4)	-8.3 (9.3)	506	1.8 (9.0)	2.6 (8.9)	552	0.94 (0.02)	1834

Notes: 34 out of 41 countries have been included (see Section 3). Estimation sample for estimating gains in achievement scores consists of students born in the period 6 months before and after the cut-off date. The countries in grey are from the sample used by Bedard & Dhuey (2006). Standard errors in parentheses are adjusted for using plausible values.

Table 4: Correlations between countries gain score and mean upper grade by age group and subject

Panel A		**Total sample using different discontinuity samples**								
		± 3 months			**± 6 months**			**± 9 months**		
test	**subject**	**correlation**	**p-value**	**N**	**correlation**	**p-value**	**N**	**correlation**	**p-value**	**N**
TIMSS 9	math	0.05	0.83	21	0.07	0.75	21	0.09	0.71	21
TIMSS 9	science	-0.23	0.32	21	-0.20	0.38	21	-0.18	0.44	21
TIMSS 13	math	0.08	0.65	34	0.04	0.81	34	0.04	0.84	34
TIMSS 13	science	0.05	0.80	34	0.13	0.46	34	0.14	0.43	34
Panel B		**Addressing birth selection around the cutoff using ± 6 months sample and different assignment variable:**								
		birth day			**birth day excluding 3 days before and after the cutoff**			**birth month**		
test	**subject**	**correlation**	**p-value**	**N**	**correlation**	**p-value**	**N**	**correlation**	**p-value**	**N**
TIMSS 9	math	0.30	0.24	17	0.28	0.28	17	0.29	0.25	17
TIMSS 9	science	-0.13	0.61	17	-0.08	0.77	17	0.06	0.82	17
TIMSS 13	math	0.14	0.46	30	0.13	0.49	30	0.15	0.44	30
TIMSS 13	science	0.14	0.45	30	0.13	0.51	30	0.14	0.47	30
Panel C		**Addressing rolling admissions & sample selection using ± 6 months sample:**								
		Bedard & Dhuey sample			**countries with first stage>0.75**			**gains corrected for sample selection**		
test	**subject**	**correlation**	**p-value**	**N**	**correlation**	**p-value**	**N**	**correlation**	**p-value**	**N**
TIMSS 9	math	0.01	0.99	10	0.22	0.56	9	-0.01	0.96	21
TIMSS 9	science	-0.28	0.44	10	-0.07	0.86	9	-0.38	0.09	21
TIMSS 13	math	-0.08	0.76	18	0.33	0.30	12	-0.14	0.43	34
TIMSS 13	science	-0.08	0.76	18	0.47	0.13	12	-0.08	0.66	34

Table 5: The effect of curriculum-based external exams on cognitive skills of 13-year and 9-year olds

	Country level data			Micro-level data		
	(1)	(2)	(3)	(4)	(5)	(6)
	mean upper grade	gain	corr. gain	mean upper grade	gain	corr. gain
Panel A: 13-year-olds						
Math	28.5*	-1.9	-6.8**	24.5***	-1.8	-3.1
	(15.6)	(3.0)	(3.0)	(0.6)	(2.5)	(2.5)
Observations	34	34	34	116235	104316	109942
Science	33.0***	-2.5	-8.8**	29.5***	-0.9	-1.4
	(11.7)	(3.8)	(3.8)	(0.6)	(2.5)	(2.5)
Observations	34	34	34	116235	104316	109942
Panel B: 9-year-olds						
Math	41.0**	-0.0	-0.7	51.4***	3.9	6.0
	(17.0)	(6.6)	(5.5)	(1.1)	(3.8)	(3.8)
Observations	21	21	21	71874	66803	68869
Science	33.9**	-4.3	-3.9	35.9***	-2.3	0.86
	(15.2)	(6.0)	(5.3)	(0.8)	(3.6)	(3.6)
Observations	21	21	21	71874	66803	68869

Notes: In columns (1)-(3) the country's mean/gain has been regressed on a dummy for CBEEE-country. In column (4) individual test scores have been regressed on the CBEEE-dummy. The model in columns (5) and (6) also includes a dummy for CBEEE-countries and the interaction of CBEEE with the dummy for being born before the cutoff date (see Equation 5). For the country level data robust standard errors are used. For the micro-level data standard errors are adjusted for using plausible values.

Appendix

Table A.1. Cutoff dates per country & source and data availability

Country	Cutoff date	Source	TIMSS 9	TIMSS 13
Austria[1]	September 1	Bedard & Dhuey	yes	yes
Belgium-Flemish	January 1	Bedard & Dhuey	no	yes
Belgium-French	January 1	Bedard & Dhuey	no	yes
Canada	January 1	Bedard & Dhuey	yes	yes
Czech Republic	September 1	Bedard & Dhuey	yes	yes
Denmark	January 1	Bedard & Dhuey	no	yes
England	September 1	Bedard & Dhuey	yes	yes
France	January 1	Bedard & Dhuey	no	yes
Greece	April 1	Bedard & Dhuey	yes	yes
Iceland	January 1	Bedard & Dhuey	yes	yes
Italy	January 1	Bedard & Dhuey	no	yes
Japan	April 1	Bedard & Dhuey	yes	yes
New Zealand	May 1	Bedard & Dhuey	yes	yes
Norway	January 1	Bedard & Dhuey	yes	yes
Portugal	January 1	Bedard & Dhuey	yes	yes
Slovak Republic	September 1	Bedard & Dhuey	no	yes
Spain	January 1	Bedard & Dhuey	no	yes
Sweden	January 1	Bedard & Dhuey	no	yes
Germany[2]	July 1	Internet/TIMMS Data	no	yes
Singapore[3]	January 1	Internet/TIMMS Data	yes	yes
South-Korea[4]	March 1	Internet/TIMMS Data	yes	yes
Latvia[5]	September 1	Internet/TIMMS Data	yes	yes
Scotland[6]	March 1	Internet/TIMMS Data	yes	yes
Lithuania[7]	September 1	Internet/TIMMS Data	no	yes
Romania[8]	September 1	Internet/TIMMS Data	no	yes
Hungary[9]	June 1	Internet/TIMMS Data	yes	yes
Slovenia[10]	January 1	Internet/TIMMS Data	yes	yes
Netherlands[11]	October 1	Internet/TIMMS Data	yes	yes
Iran	October 1	TIMSS Data	yes	yes
Thailand	January 1	TIMSS Data	yes	yes
Cyprus	March 1	TIMSS Data	yes	yes
Switzerland	January 1	TIMSS Data	no	yes
Russia	October 1	TIMSS Data	no	yes
Hong Kong	January 1	TIMSS Data	yes	yes

Notes: All cutoff dates have been checked in our data and show a (sharp) discontinuity in average grade around the given cutoff. The column 'Source' shows whether the cutoff was also shown in other sources. Cutoff dates refer to the situation in 1995. The columns 'TIMSS 9' and 'TIMSS 13' show whether the country was included for these tests.

1 Bedard & Dhuey use January 1 as the cutoff date, we deviate based on the data and: http://virtuelleschule.bmukk.gv.at/fileadmin/folder/Folder_Basisinformationen/school_system_Austria_EN.pdf
2 http://de.wikipedia.org/wiki/Schulpflicht_(Deutschland)
3 http://www.moe.gov.sg/education/
4 http://en.wikipedia.org/wiki/Education_in_South_Korea#Elementary_school
5 http://www.viaa.gov.lv/files/news/1808/educ_in_latvia.pdf
6 http://en.wikipedia.org/wiki/Education_in_Scotland
7 http://en.wikipedia.org/wiki/Education_in_Lithuania
8 http://en.wikipedia.org/wiki/Romanian_educational_system
9 http://en.wikipedia.org/wiki/Education_in_Hungary
10 http://eacea.ec.europa.eu/education/eurydice/documents/eurybase/national_summary_sheets/047_SI_EN.pdf
11 http://www.onderwijsinspectie.nl/actueel/vraagantwoord#Wie_bepaalt_of_een_kind_overgaat_naar_groep_3_

Table A.2a: Means of test scores and covariates by age relative to the cutoff for the pooled sample of 9-year-olds

relative age	math	science	female	born in country	speaks language of test at home	living with mother	living with father	number of books at home	N
-12	549.21	538.68	0.51	0.93	0.75	0.96	0.85	95.94	4999
-11	547.29	536.93	0.51	0.93	0.76	0.95	0.84	96.77	5137
-10	545.45	535.48	0.52	0.93	0.74	0.96	0.84	97.78	5613
-9	543.34	532.30	0.50	0.92	0.76	0.95	0.84	97.69	5664
-8	542.74	532.14	0.51	0.93	0.76	0.95	0.84	97.41	5763
-7	538.04	528.83	0.50	0.93	0.75	0.96	0.84	95.76	5676
-6	538.46	528.05	0.50	0.93	0.76	0.96	0.84	97.62	5707
-5	533.71	522.10	0.49	0.92	0.74	0.96	0.84	95.01	5812
-4	533.02	521.58	0.51	0.93	0.74	0.96	0.84	97.58	5535
-3	529.74	516.15	0.50	0.93	0.72	0.96	0.85	92.75	6099
-2	526.90	515.43	0.50	0.92	0.73	0.95	0.84	95.14	5678
-1	517.61	506.14	0.50	0.93	0.73	0.96	0.85	93.07	5911
0	498.54	494.00	0.51	0.92	0.75	0.96	0.85	97.37	5400
1	489.24	487.23	0.49	0.92	0.74	0.96	0.84	97.33	5266
2	488.91	484.65	0.51	0.92	0.75	0.96	0.84	96.59	5261
3	485.70	483.74	0.49	0.93	0.75	0.96	0.85	98.58	5348
4	481.92	479.15	0.52	0.93	0.74	0.95	0.85	93.91	5431
5	479.01	476.73	0.51	0.92	0.75	0.95	0.84	96.18	5355
6	479.60	475.20	0.49	0.92	0.74	0.95	0.84	96.72	5372
7	474.85	471.19	0.50	0.93	0.72	0.94	0.84	92.73	5200
8	474.75	470.37	0.48	0.92	0.72	0.95	0.84	92.43	5099
9	475.07	466.54	0.49	0.92	0.71	0.95	0.85	93.02	4957
10	473.42	466.33	0.49	0.92	0.70	0.95	0.85	91.21	4436
11	468.58	460.85	0.49	0.93	0.70	0.95	0.84	89.33	4623

Note: The relative age of the oldest students is -12; relative age 0 means born in the first month at the right side of the cut-off data.

Table A.2b: Means of test scores and covariates by age relative to the cutoff for the pooled sample of 13-year-olds

relative age	math	science	female	born in country	speaks language of test at home	living with mother	living with father	number of books at home	high educated mother	high educated father	N
-12	518.765	513.228	0.502	0.940	0.827	0.952	0.826	102.198	0.261	0.315	7752
-11	517.974	514.059	0.498	0.939	0.838	0.952	0.832	105.008	0.286	0.342	7701
-10	517.791	513.232	0.501	0.944	0.842	0.955	0.829	106.183	0.282	0.337	8666
-9	518.192	513.094	0.501	0.945	0.849	0.954	0.838	106.671	0.297	0.350	9126
-8	518.042	512.474	0.502	0.947	0.841	0.958	0.827	106.556	0.284	0.334	9329
-7	513.727	508.215	0.505	0.946	0.834	0.959	0.833	104.163	0.285	0.341	9222
-6	515.361	509.274	0.499	0.943	0.837	0.955	0.837	104.889	0.294	0.340	9168
-5	514.537	508.856	0.494	0.946	0.837	0.959	0.841	106.206	0.300	0.354	9256
-4	512.667	505.842	0.497	0.944	0.832	0.954	0.840	106.218	0.293	0.340	9233
-3	511.510	504.443	0.492	0.949	0.831	0.959	0.838	105.772	0.295	0.345	9160
-2	509.452	503.850	0.491	0.948	0.835	0.954	0.838	104.219	0.285	0.333	8747
-1	505.535	500.905	0.498	0.950	0.822	0.957	0.843	102.909	0.289	0.354	9191
0	500.315	489.859	0.490	0.949	0.839	0.954	0.837	104.547	0.307	0.357	8563
1	499.595	488.652	0.499	0.952	0.838	0.957	0.844	106.647	0.307	0.351	7883
2	496.908	487.421	0.486	0.954	0.845	0.958	0.843	107.520	0.316	0.368	8364
3	495.877	484.989	0.490	0.955	0.841	0.958	0.843	107.507	0.313	0.362	8422
4	493.663	482.046	0.490	0.949	0.836	0.962	0.844	106.720	0.297	0.358	8403
5	493.269	482.782	0.494	0.959	0.842	0.962	0.846	108.773	0.316	0.364	7926
6	493.782	480.746	0.479	0.956	0.838	0.961	0.846	107.214	0.313	0.362	8156
7	492.918	481.231	0.483	0.954	0.836	0.964	0.843	106.992	0.325	0.383	7962
8	492.970	480.117	0.488	0.954	0.824	0.963	0.847	108.596	0.324	0.377	7906
9	493.018	478.652	0.481	0.954	0.823	0.967	0.851	107.079	0.312	0.361	7316
10	491.793	479.242	0.477	0.955	0.831	0.959	0.852	106.122	0.309	0.361	7052
11	488.953	478.188	0.467	0.959	0.815	0.967	0.858	105.302	0.308	0.368	6634

Note: The relative age of the oldest students is -12; relative age 0 means born in the first month at the right side of the cut-off data. High educational level is defined as having some vocational education or more.

Table A.3a: Balancing tests for 9-year-olds

	Effects on variable		Effects on dummy=1 if variable is missing	
	± 3 months	± 6 months	± 3 months	± 6 months
Effect of being born left of the cutoff date on:	(1)	(2)	(3)	(4)
Female	-0.00644	0.00189	0.00106	0.000600
	(0.0117)	(0.00757)	(0.00123)	(0.000932)
N	33502	66560	33615	66803
Born in country of test	0.00681	0.00311	-0.00254	-0.000786
	(0.00603)	(0.00432)	(0.00289)	(0.00221)
N	30798	61120	33615	66803
Language at home is language of test	0.0121	0.00229	0.00277	0.00351
	(0.00994)	(0.00687)	(0.00664)	(0.00513)
N	24666	49050	33615	66803
Living with father	-0.00613	0.000481	0.00213	-0.000342
	(0.00855)	(0.00576)	(0.00326)	(0.00240)
N	30575	60559	33615	66803
Living with mother	0.000436	-0.00503	0.000197	-0.000594
	(0.00464)	(0.00309)	(0.00277)	(0.00219)
N	30702	60793	33615	66803
Number of books at home	-0.525	-0.914	-0.00340	-0.00851**
	(1.733)	(1.254)	(0.00492)	(0.00340)
N	29706	58884	33615	66803

Notes: Each cell shows the estimation results of a separate regression of the covariate on a dummy for being born at the left side of the cut-off and a linear function of age. All models include country dummies and interactions of these dummies with the age function which differs at either side of the cutoff. The dependent variable in columns (3) and (4) is a dummy for having a missing value on the relevant covariate.

Table A.3b: Balancing tests for 13-year-olds

	Effects on variable		Effects on dummy=1 if variable is missing	
	± 3 months	± 6 months	± 3 months	± 6 months
Effect of being born left of the cutoff date on:	(1)	(2)	(3)	(4)
Female	0.00222	0.00194	-0.00193***	-0.00125***
	(0.00996)	(0.00636)	(0.000690)	(0.000452)
N	51839	104162	51908	104316
Born in country of test	-0.00279	-0.00257	0.00118	0.000752
	(0.00414)	(0.00273)	(0.00185)	(0.00127)
N	47727	95783	51908	104316
Language at home is language of test	-0.000758	-0.00226	-0.00927*	-0.00736**
	(0.00613)	(0.00424)	(0.00485)	(0.00370)
N	45561	91554	51908	104316
Living with father	-0.00368	-0.00240	0.00180	-0.00121
	(0.00693)	(0.00467)	(0.00231)	(0.00157)
N	48507	97408	51908	104316
Living with mother	-0.000945	0.000630	-0.000108	-0.00186
	(0.00365)	(0.00255)	(0.00187)	(0.00128)
N	48725	97833	51908	104316
Number of books at home	-1.159	-0.871	-0.00105	-0.00117
	(1.362)	(0.913)	(0.00219)	(0.00148)
N	48536	97463	51908	104316
High educated mother	-0.0102	-0.0150***	-0.00740	-0.0173***
	(0.00872)	(0.00576)	(0.00786)	(0.00498)
N	35683	71869	51908	104316
High educated father	0.00549	-0.00569	-0.00229	-0.0102**
	(0.0103)	(0.00694)	(0.00745)	(0.00505)
N	34732	69875	51908	104316

Notes: Each cell shows the estimation results of a separate regression of the covariate on a dummy for being born at the left side of the cut-off and a linear function of age. All models include country dummies and interactions of these dummies with the age function which differs at either side of the cutoff. The dependent variable in columns (3) and (4) is a dummy for having a missing value on the relevant covariate.

Appendix A.4: Sampling bias adjustment

Table A.4 illustrates the sampling bias adjustment. For instance, we do not observe retained students born in the first month at the right side of the cut-off date (relative age=0). However, we do observe retained students born one year earlier (relative age = -12). We adjust the scores of these students and include them in the main estimation sample. The adjusted score is obtained by:

$$Y^{i}_{delayed(missing)} = Y^{i}_{ontrack} * \frac{Y^{i-12}_{delayed}}{Y^{i-12}_{ontrack}}$$

Hence, for the missing students with relative age=0 we get: 492.40*439.79/553.60=492.40*0.79=391.17. Adjusted Scores for missing accelerated students are similarly obtained:

$$Y^{i}_{accelerated(missing)} = Y^{i}_{ontrack} * \frac{Y^{i+12}_{accelerated}}{Y^{i+12}_{ontrack}}$$

To obtain estimates that are adjusted for sampling bias we perform this adjustment for each separate month $i \in [0,5]$ for missing delayed students, and $i \in [-6,-1]$ for missing accelerated students.

Table A.4: Fraction on track and average test scores of those on track and not on track (delayed or accelerated) for TIMMS 9. l=left of the cutoff, r=right of the cutoff

rel. age (l)	*on track (l)*	*math on track (l)*	*math delayed (l)*	*science on track (l)*	*science delayed (l)*	*N*
-12	0.96	553.60	439.79	543.00	431.11	4999
-11	0.95	552.02	447.80	541.93	431.64	5137
-10	0.96	550.33	435.40	540.32	426.11	5613
-9	0.95	548.72	450.66	537.28	446.69	5664
-8	0.94	548.59	449.64	537.74	443.10	5763
-7	0.94	543.89	449.54	534.36	445.10	5676
-6	0.93	546.16	442.95	535.30	438.11	5707
-5	0.91	540.77	461.40	528.43	457.27	5812
-4	0.90	541.75	453.36	529.38	450.50	5535
-3	0.86	542.77	450.16	527.61	446.10	6099
-2	0.85	538.12	464.04	524.60	464.07	5678
-1	0.79	532.37	462.32	518.05	461.56	5911
rel. age (r)	*on track (r)*	*math on track (r)*	*math accelerated (r)*	*science on track (r)*	*science accelerated (r)*	*N*
0	0.88	492.40	544.92	488.86	532.84	5400
1	0.94	485.58	541.89	484.26	530.05	5266
2	0.96	487.53	525.59	483.54	514.22	5261
3	0.98	484.87	518.83	483.14	507.47	5348
4	0.98	481.18	520.73	478.69	503.52	5431
5	0.98	478.50	507.47	476.18	507.22	5355
6	0.99	479.26	507.59	474.97	494.04	5372
7	0.99	474.46	502.53	470.97	486.76	5200
8	0.99	474.63	483.30	470.19	483.32	5099
9	0.98	474.50	507.13	465.99	497.18	4957
10	0.99	473.45	471.19	466.38	461.11	4436
11	0.99	468.39	483.51	460.52	487.32	4623

Table A.5a. Estimated gains in achievement by assignment variable (birth month/birth day) for 9-year olds. 3 days around the cutoff excluded when using birth day as assignment variable

		Math				Science			
		birth month		birth day		birth month		birth day	
Ranking	Country	gain	N	gain	N	gain	N	gain	N
1	Norway	43.3 (7.4)	2133	-	-	38.8 (8.5)	2133	-	-
2	Singapore	40.0 (5.2)	6986	38.3 (5.2)	6855	27.8 (4.9)	6986	26.3 (4.9)	6855
3	Greece	36.3 (6.7)	2981	33.4 (11.2)	1055	21.7 (6.5)	2981	23.0 (11.0)	1055
4	Iceland	34.6 (9.5)	1702	34.2 (9.2)	1474	30.9 (8.4)	1702	26.6 (9.0)	1474
5	Iran	33.3 (7.1)	2716	-	-	35.9 (6.8)	2716	-	-
6	Cyprus	30.5 (6.4)	3230	34.3 (6.5)	3167	20.4 (5.8)	3230	23.7 (5.6)	3167
7	Japan	28.6 (5.4)	4343	31.6 (5.3)	4268	19.1 (5.3)	4343	21.3 (5.4)	4268
8	Czech Republic	26.9 (6.9)	3108	28.2 (7.4)	2777	16.9 (6.5)	3108	14.3 (7.4)	2777
9	Canada	24.4 (5.1)	7436	-	-	21.7 (4.8)	7436	-	-
10	Netherlands	23.6 (6.8)	2241	21.5 (6.6)	2195	13.9 (6.7)	2241	11.0 (6.9)	2195
11	Hungary	23.4 (7.8)	2743	28.8 (8.0)	2590	26.7 (7.0)	2743	30.1 (7.6)	2590
12	Hong Kong	22.0 (5.6)	3851	22.0 (5.5)	3724	16.3 (5.0)	3851	17.0 (5.1)	3724
13	Austria	21.9 (7.4)	2315	22.0 (7.8)	2263	11.4 (8.1)	2315	11.2 (8.1)	2263
14	South-Korea	21.3 (6.5)	2636	24.2 (6.5)	2570	12.8 (6.4)	2636	16.0 (6.8)	2570
15	Scotland	20.1 (6.6)	3089	18.5 (7.1)	2593	12.9 (7.2)	3089	7.3 (7.8)	2593
16	Portugal	17.7 (7.3)	2310	16.6 (7.5)	2185	17.9 (8.9)	2310	17.9 (8.8)	2185
17	England	13.5 (7.1)	3087	-	-	12.9 (9.2)	3087	-	-
18	Slovenia	12.7 (7.3)	2484	14.5 (7.3)	2416	7.6 (7.1)	2484	9.4 (7.1)	2416
19	Latvia	6.9 (7.8)	2116	5.4 (8.5)	1906	3.0 (9.0)	2116	3.0 (9.1)	1906
20	New Zealand	2.7 (7.6)	2459	4.4 (7.6)	2405	3.8 (8.4)	2459	5.9 (8.2)	2405
21	Thailand	-0.4 (6.2)	2837	13.5 (6.3)	2558	-5.1 (6.0)	2837	7.1 (6.1)	2558

Table A.5b: Estimated gains in achievement by assignment variable (birth month/birth day) for 13-year-olds. 3 days around the cutoff excluded when using birth day as assignment variable

		Math				Science			
		birth month		birth day		birth month		birth day	
Ranking	**Country**	**gain**	**N**	**gain**	**N**	**gain**	**N**	**gain**	**N**
1	Singapore	27.8 (4.6)	3567	26.8 (4.9)	3502	55.9 (5.9)	3567	55.0 (6.2)	3502
2	Sweden	19.1 (5.4)	3403	16.6 (5.5)	3235	25.2 (5.5)	3403	23.3 (5.7)	3235
3	Italy	18.5 (8.3)	2186	16.1 (8.4)	2142	15.0 (8.0)	2186	12.1 (8.0)	2142
4	Norway	18.2 (6.2)	2751	-	-	22.1 (6.7)	2751	-	-
5	Iran	15.5 (7.1)	2495	-	-	7.7 (6.9)	2495	-	-
6	Czech Republic	15.5 (5.6)	3248	14.4 (5.8)	3162	9.1 (5.6)	3248	8.0 (5.8)	3162
7	Spain	13.6 (5.7)	3157	14.6 (5.9)	3084	10.9 (5.9)	3157	11.6 (6.0)	3084
8	Iceland	13.5 (7.2)	1834	11.8 (7.6)	1641	17.1 (7.4)	1834	14.0 (7.9)	1641
9	South-Korea	13.1 (6.5)	2912	13.8 (6.8)	2857	10.7 (6.2)	2912	11.8 (6.3)	2857
10	Denmark	12.1 (7.3)	2096	15.2 (7.7)	2000	8.3 (8.2)	2096	10.8 (8.5)	2000
11	Canada	9.9 (3.6)	7733	-	-	3.2 (4.4)	7733	-	-
12	Latvia	9.9 (7.4)	2334	9.1 (7.6)	2263	23.2 (7.6)	2334	24.2 (8.0)	2263
13	Scotland	9.9 (6.6)	2824	8.5 (7.1)	2457	12.7 (7.1)	2824	9.4 (7.7)	2457
14	Thailand	9.7 (4.9)	5232	8.3 (4.8)	5120	11.1 (3.8)	5232	10.1 (4.2)	5120
15	Cyprus	9.5 (8.1)	2837	6.5 (8.4)	2772	14.7 (7.8)	2837	12.5 (8.1)	2772
16	Slovak Republic	8.6 (5.5)	3475	9.3 (5.5)	3385	11.9 (5.8)	3475	10.5 (5.9)	3385
17	Belgium (French)	7.1 (6.8)	1872	6.8 (7.2)	1778	11.5 (7.7)	1872	11.9 (7.9)	1778
18	Switzerland	6.7 (4.3)	3727	5.4 (4.4)	3595	8.7 (5.5)	3727	8.9 (6.0)	3595
19	Japan	4.8 (4.6)	5158	3.7 (4.7)	5075	14.1 (4.5)	5158	13.7 (4.8)	5075
20	Greece	3.0 (6.3)	3543	14.4 (10.2)	1484	5.9 (6.8)	3543	12.6 (10.4)	1484
21	Belgium (Flemish)	2.8 (4.8)	2622	2.8 (5.0)	2479	12.2 (5.4)	2622	13.0 (5.5)	2479
22	Lithuania	2.6 (7.3)	2590	2.2 (7.7)	2531	17.6 (7.5)	2590	19.2 (7.8)	2531
23	Romania	2.2 (6.3)	3504	-1.8 (6.4)	3305	2.0 (6.8)	3504	-3.8 (7.3)	3305
24	Russia	1.9 (5.7)	3855	1.1 (5.9)	3780	14.9 (6.6)	3855	14.5 (6.7)	3780
25	Hungary	0.5 (5.8)	2887	2.4 (6.0)	2827	4.0 (6.0)	2887	6.0 (6.1)	2827
26	Portugal	0.4 (5.8)	2579	1.5 (5.9)	2532	13.4 (7.5)	2579	15.2 (7.6)	2532
27	Germany	-0.1 (6.4)	2447	-	-	-0.4 (7.4)	2447	-	-
28	Slovenia	-0.8 (6.3)	2688	1.0 (6.4)	2646	-10.3 (6.5)	2688	-9.3 (6.6)	2646
29	France	-2.6 (6.7)	2159	-1.0 (7.0)	2121	2.0 (6.6)	2159	3.7 (6.7)	2121
30	Austria	-3.6 (6.0)	2569	-6.1 (6.2)	2498	1.1 (6.8)	2569	-0.4 (7.0)	2498
31	New Zealand	-4.7 (6.0)	3432	-3.4 (6.3)	3333	-3.4 (6.7)	3432	0.2 (7.1)	3333
32	Netherlands	-5.7 (6.5)	1770	-4.1 (6.8)	1677	5.5 (7.2)	1770	6.6 (7.6)	1677
33	Hong Kong	-7.1 (6.3)	2996	-6.7 (6.4)	2895	-2.2 (6.7)	2996	-0.6 (7.0)	2895
34	England	-9.0 (9.4)	1834	-11.3 (9.9)	1635	1.8 (9.0)	1834	-1.7 (9.7)	1635

Publisher:

CPB Netherlands Bureau for Economic Policy Analysis
P.O. Box 80510 | 2508 GM The Hague
T (070) 3383 380

September 2013 | ISBN 978-90-5833-615-6