# THE PERSISTENCE OF EARLY CHILDHOOD MATURITY: INTERNATIONAL EVIDENCE OF LONG-RUN AGE EFFECTS*

Kelly Bedard and Elizabeth Dhuey

## Abstract

A continuum of ages exists at school entry due to the use of a single school cut-off date – making the "oldest" children approximately twenty percent older than the "youngest" children. We provide substantial evidence that these initial maturity differences have long lasting effects on student performance across OECD countries. In particular, the youngest members of each cohort score 4-12 percentiles lower than the oldest members in grade four and 2-9 percentiles lower in grade eight. In fact, data from Canada and the United States shows that the youngest members of each cohort are even less likely to attend university.

## I. INTRODUCTION

Nearly all education systems have a single cut-off date for school eligibility. For example, a child may be allowed to enter kindergarten as long as he is five years old by September 1 of the relevant year. Cut-off dates are important because they cause some students to be older than others when they begin school. To put this in perspective, in an education system in which students must be five to start school, the oldest students are approximately twenty percent older than the youngest students at school entry. Given the magnitude of the age range on the first day of school, the oldest students are likely to be substantially more mature than the youngest students. As such, one would expect an age-based performance differential during the early grades. If this relative maturity effect is significant in early primary grades, but then dissipates with age, this phenomenon, while interesting, is not particularly important for the economy. On the other hand, if early relative maturity effects propagate themselves through the human capital accumulation process into later life, long after small differences in age are important in and of themselves, they may have important implications for adult outcomes and productivity.

Recent work by Heckman and various coauthors (see Cuhna, Heckman, Lochner, and Masterov [2006] for an excellent review) argues that skills accumulated in early childhood are complementary to later learning. Relative age differences at the start of formal schooling may therefore be long-lasting if relatively older students are better positioned to accumulate more skills in the early grades because their maturity advantage increases the likelihood that they are selected for more advanced curriculum groups or because they progress through a common curriculum at a faster rate.[1] Notice that relative age effects can persist either because students are separated into programs with different rates of human capital accumulation during the early

primary grades or because stronger students are encouraged to continue progressing through the curriculum while weaker students are simply allowed to lag farther and farther behind. As such, relative age effects can be generated by education systems using ability-specific curriculum groups during the primary grades like in England and the United States as well as by countries that employ social promotion and claim to have no ability specific tracking, such as Japan.

Uncovering the causal impact of early relative maturity on later outcomes is difficult because age enters into educational decisions in at least four important ways. First, school cut-off dates only determine relative age if the rules are strictly followed. For example, a significant fraction of American children defer school entry by a year, making them the oldest students [Datar, 2006]. This is problematic because these children are not a random draw. To distinguish between observed relative age and the relative age at which a child should be observed, based on their birth date relative to the school cut-off date, we refer to the latter measure as <u>assigned relative age</u>. Secondly, children who are young at school entry are more likely to repeat a grade. Thirdly, relative maturity may at least partially determine academic program placement during elementary school. Finally, at young ages, relatively older students will be more mature and hence score higher on achievement tests, independent of program placement.

While relative age evaluated at any point in the educational process is endogenous, the initial timing of births is arguably exogenous. We therefore compare the test scores of children with older and younger assigned relative ages at the fourth and eighth grade levels across OECD countries using data from the Trends in International Mathematics and Science Study (TIMSS). Following from the previous discussion, the impact of assigned relative age on test scores reflects both differential school entry and grade retention/failure across the assigned relative age distribution, as well as differences in program placement and skill acquisition, and is therefore a

net, or reduced form, effect. Given that we know both observed and assigned relative age, we can also estimate the causal (within grade) impact of relative age using assigned relative age as an instrument for observed age. Finally, since some of the countries participating in TIMSS employ "clean" education systems, in the sense that essentially all children enter on time and pass from one grade to the next on schedule, we can also compare the clean countries to the rest of the countries to get a sense of the relative importance of failure and within grade differences in relative maturity in determining the net (reduced form) relative age effects. Our ability to compare countries with different education systems in TIMSS is one of the important novelties of this paper. Using the data from nineteen countries allows us to ascertain the pervasiveness of relative age effects across countries and education systems.

Overall, we find that the youngest students score substantially lower than the oldest students at both the fourth and eighth grade levels. In grade four, the youngest students score 1.2-3.5 points lower on nationally standardized tests with a mean of 50 and a standard deviation of 10. To put this in perspective, this translates into a 4-12 percentile disadvantage for eleven months of relative age. While the age premium enjoyed by the oldest students declines between grades four and eight, there remains a 0.8-2.6 point difference, or 2-9 percentiles, between the oldest and the youngest students at the eighth grade level. These results clearly show the persistence of relative age into adolescence, and are therefore suggestive of a longer run impact.

In order to confirm the existence of long-run relative age effects, the last section of the paper examines the impact of relative maturity on the probability of participating in a pre-university program during the final year of high school in British Columbia, Canada and the probability of writing the SAT and enrolling in an accredited four year college in the United States – the only two jurisdictions for which we have been able to obtain micro-level university-

stream and month of birth data. In British Columbia, individuals born in the relatively youngest month are underrepresented in the pre-university program by 9.8 percent. The results for the United States are similar: Individuals born in the first assigned relative month are underrepresented in the pre-university stream (as measured by taking the SAT or ACT) by 7.7 percent. Further, individuals born in the first relative month are underrepresented in accredited four-year college/universities enrollments by 11.6 percent. Taken as a whole, the results from TIMSS and the university-bound/enrollment results from British Columbia and the United States clearly point to substantial long-run relative age effects that have important implications for the distribution of adult skills.

The remainder of the paper is as follows. Section II discusses the possible avenues through which relative maturity may be propagated into adulthood. Section III describes the econometric framework. Section IV discusses the data used in the analysis and looks for birth date targeting. Section V reports the relative age estimates at the fourth and eighth grade levels. Section VI analyzes the impact of relative age on pre-university program and college enrollment. Section VII concludes.

## II. THE PROPAGATION OF EARLY MATURITY DIFFERENCES
## INTO THE LONG-RUN

Despite the myriad of educational structures used across countries, essentially all nations have two common features: (1) A single annual cut-off date,[2] which generates at least a one-year age range within each grade cohort and (2) Ability-based sorting into curriculum groups or classes. In some countries sorting takes the form of strict program based streaming (i.e. academic versus vocational) while in others it takes the more flexible form of ability grouping

(i.e. reading groups).  In fact, even countries that employ social promotion (automatic promotion from one grade to the next) and claim to have only one track are implicitly streaming to the extent that the weakest students are allowed to fall progressively farther behind.  These types of educational structures are importantly related to relative age because skill-based curriculum usually begins during the primary grades when relative maturity likely plays a large part in determining skill differences between young and old students.  The interaction between the timing of skill-based curriculum commencement and relative maturity may therefore play a central role in determining program/group placement, and may hence affect skill accumulation throughout the educational process, even after relative age is irrelevant in and of itself.

Allen and Barnsley [1993] provide us with a simple example.  They report that 72 percent of boys between the ages of 16-20 playing in the highest level of minor hockey in Canada are born in the first half of the year (the cut-off date for Canadian hockey is January 1).  This results because the substantial variation in maturity within young cohorts makes it more likely that older boys are selected for more competitive (rep) teams.  Since rep teams attract the best coaches, practice more, and play against higher caliber opponents, rep team members accumulate more hockey skills and thereby increase their probability of being selected for rep teams in the future. While it would be surprising to find such enormous relative age differences for long-run educational outcomes, the combination of the Allen and Barnsley results and the extensive use of skill-based curriculum in schools (see Appendix Table I) certainly suggest that there is the potential for substantial differences in educational success across the relative age distribution.

Given the discussion so far, it is tempting to hypothesize that countries that rigidly stream students into academic and vocational streams at young ages will have the strongest propagation of relative age effects into adolescence and adulthood.  This is particularly appealing since this

type of streaming is observable and hence seems testable. However, this type of streaming does not usually occur until adolescence, long after less rigid forms of streaming such as reading and math groups and enrichment programs have already begun. As such, adolescent stream placement is at least partly an outcome of early maturity differences that influence primary level program placement and subsequent academic progress. Since most countries use such structures, it is difficult to determine, a priori, which countries we should expect to have the largest relative age effects. Further, there will even be relative age differences in student performance in countries that use social promotion as long as stronger students are encouraged to continue moving ahead even as weaker students fall behind. This is of course what one would expect unless students are only allowed to accumulate a specific set of skills in each grade and are then forced to wait for the rest of the class to catch up before progressing.

## III. ECONOMETRIC FRAMEWORK

We begin with a simple model of the relationship between student outcomes and observed age.

$$(1) \ S_{cgi} = \alpha_{cg} + \beta_{cg} A_{cgi} + X_{cgi} \gamma_{cg} + \varepsilon_{cgi}$$

where $S_{cgi}$ denotes student outcomes, usually a test score, for student $i$ in country $c$ in grade $g$, $A$ is observed age, $X$ is the vector of controls described in Section IV, and $\varepsilon$ is the usual error term. All models are estimated separately for each grade and country. The parameter of interest is $\beta_{cg}$ – the causal impact of relative age. However, the causal interpretation rests on the assumption that unobservables do not confound the observed age effect, which is clearly untrue given non-random grade retention. Since children who enter kindergarten early tend to score worse than they otherwise would and there are many more children who are old because they are retained

6

(who tend to score poorly) than children who are old because their parents hold them out of school so that they enter kindergarten a year late (who are positively selected), OLS estimates are downward biased.  In fact, in countries where a large fraction of students repeat at least one primary grade, such as in the United States, the OLS estimates can even be negative (see Section V).

We propose an instrumental variables (IV) solution to this problem using birth month relative to the school cut-off date, assigned relative age ($R$), as an exogenous determinant of observed age.  More specifically, we estimate the parameters of equation (1) using TSLS based on the following the first-stage equation for observed age:

(2) $A_{cgi} = \pi_{1cg} + \pi_{2cg} R_{cgi} + X_{cgi} \pi_{3cg} + v_{cgi}$ .

While the reduced form relationship between observed age and assigned relative age (the first stage) is not particularly interesting, the reduced form relationship between test scores and assigned relative age is important from a policy perspective.

(3) $S_{cgi} = \theta_{1cg} + \theta_{2cg} R_{cgi} + X_{cgi} \theta_{3cg} + u_{cgi}$

In particular, $\theta_{2cg}$ measures the impact of assigned relative age net of grade repetition and late entry.  Stated somewhat differently, $\theta_{2cg}$ is the overall, net, or reduced form impact of assigned age on test scores at a given grade level $g$ (we return to this point in Section V.A).

For the IV estimator to provide a consistent estimate two conditions must be satisfied. First, assigned relative age must be correlated with observed relative age.  Since most students enter school on time and are never retained this is easily satisfied (see Tables III and IV).  In other words, assigned relative age is clearly an important determinant of observed age.  The second condition requires that assigned relative age be uncorrelated with the unobserved determinants of test scores.  This assumption is violated if, for example, children born at

different times of the year have higher or lower unobserved ability levels. We explore this issue

in two ways. First, we study birth month patterns to check for birth date targeting by different

socioeconomic groups (see Section IV.D). Secondly, we control for season of birth effects

directly by estimating an alternative specification that includes both assigned relative age and

month of birth using data pooled across countries with cut-off dates in different months (see

Appendix Table I).[3] For example, the school cut-off date is September 1 in England, January 1

in France, and April 1 in Greece. As such, children born in the same calendar month (season of

birth) have different relative ages if they live in different countries. This is helpful because it

make us more confident that we are not confounding season of birth and relative age (see Bound

and Jaeger [2000]).

For concreteness, the main equation becomes,

$$(4) \quad S_{cgi} = \varphi_{1cg} + \varphi_{2g} + \varphi_{3g} A_{cgi} + X_{cgi} \varphi_{4g} + \nu_{cgi}$$

where $S_{cgi}$ is an internationally standardized test score, $\varphi_{1cg}$ is a vector of country indicators, and

$\varphi_{2g}$ is a vector of month of birth indicators (to capture season of birth effects that are common

across countries). Unlike equation (1), this TSLS model is estimated using data pooled across

countries, but separately by grade.

Our ability to separate the effects of relative age and season of birth using test data across

countries with different school start dates is one of the unique features of this study. While we

are not the only economists to have examined the impact of relative age on student performance,

ours is the only study capable of separating relative age and season of birth. This is due to the

fact that recent work by Datar [2006] for the United States, Fredriksson and Öckert [2004] for

Sweden, and Puhani and Weber [2005] for Germany are all based on samples from a single

country with a single school cut-off date, or in the case of the United States almost all fall/early winter cut-off dates.[4]

## IV. DATA

The data used in this study come from five sources. Our primary sources are the 1995 and 1999 Trends in International Mathematics and Science Study (TIMSS), which we supplement with the Early Childhood Longitudinal Study (ECLS) and the National Education Longitudinal Study (NELS) data for the United States in order to examine fourth and eighth grade test scores. In addition, we also use administrative data for British Columbia, Canada and NELS data for the United States to estimate the long-run impact of relative age on pre-university program and college participation. To avoid confusion, this section focuses on our primary (TIMSS) and supplementary (ECLS and NELS) data sources. The secondary data sets used to examine university outcomes are described in Section VI.

IV.A. The Trends in International Mathematics and Science Study (TIMSS)

TIMSS is an excellent source for studying the impact of relative age on student performance it includes nationally representative mathematics and science achievement results for third and fourth graders in 26 countries in 1995 and seventh and eighth graders in 41 and 38 countries in 1995 and 1999, respectively. We restrict the sample to OECD countries with unambiguous nationwide school starting age rules (cut-off dates). Australia, Germany, Hungary, Ireland, the Netherlands, Switzerland, and the United States are excluded because their rules regarding the school cut-off date either differ across regions, which are not reported in TIMSS, or the cut-off date is at the discretion of educators and/or parents. In addition, Korea and Turkey are excluded because their birth date data is of questionable quality.[5] These exclusions leave us

with a sample of 10 countries for third and fourth graders and 18 countries for seventh and eighth graders: For a total of 228,629 observations across all ages and countries. However, students who do not report their sex, birth month, birth year, test year, or test month are excluded, reducing the sample by 2,857 students. The country-grade specific sample sizes are reported in Table I.

It is important to clarify exactly who is being tested. The 1995 TIMSS includes test scores for two different grade groups. The first set of scores is for students enrolled in the two adjacent grades that contain the largest proportion of nine-year-olds – third and fourth graders in most countries. For expositional ease, we refer to these students as fourth graders. The second set of scores is for students enrolled in the two adjacent grades that contain the largest proportion of thirteen-year-olds – seventh and eighth graders in most countries.[6] We refer to these students as eighth graders. In contrast, the 1999 TIMSS includes only one age group in a single grade. While the 1999 TIMSS uses the 1995 definition to target the two adjacent grades containing the most thirteen-year-olds, only students in the upper of the two grades were tested – eighth graders in most countries. We again refer to these students as eighth graders.

The TIMSS test scores used in all analyses are standardized to mean 50 and a standard deviation 10. Two distinct standardizations are used. First, all country specific models are estimated using test scores that are standardized by test book within each country. Within test book standardization is required because each student wrote only one of eight possible exams, and within country standardization allows us to include United States estimates based on ECLS and NELS data in Tables III and IV.[7] Second, the cross-country models use data that is standardized within test book across all TIMSS participants. In both cases, the data is population weighted. Summary statistics for the internationally standardized scores are reported in Table I

by country. As one would expect, the country-specific internationally standardized scores means are generally above 50 because we are focusing on the subset of OECD countries. We do not report the within country standardized scores since the means and standard deviations only differ from 50 and 10 due to a small number of sample exclusions necessitated by missing data.

Measuring assigned relative age requires knowledge of the cut-off date for children to begin school. For example, if a child is allowed to enter kindergarten as long as he has reached the age of five by September 1 of the relevant year, then September 1 is the cut-off date. We determined these cut-off dates using the empirical distribution of birth months in each country. The beginning of the first month of the twelve consecutive months that contains the largest percentage of student birth dates is defined as the cut-off date. We then confirmed each of these dates using Eurydice (see www.eurodice.org), an information network on education in Europe, established by the European Commission, or by using an individual country's Department of Education website. We then use the cut-off date and month of birth, which is reported in TIMSS, to construct a linear measure of assigned <u>relative</u> age ($R$). More specifically, $R = 0$ for students born in the last eligible month and $R = 11$ for students born the first eligible month. For example, if the cut-off date is January 1, December babies are the youngest ($R = 0$) and January babies are the oldest ($R = 11$). Actual age in months ($A$) is constructed using the test date and birth date, both of which are reported in months.

All test score models include a basic set of socio-economic controls. These include indicator variables for sex, grade, test year (in eighth grade models only), rural residential locations, native born mother, native born father, child living with both parents, child has a calculator, child has a computer, child has more than 100 books, and parental education (in eighth grade models only), and a continuous measure for the number of people residing in the

child's household.  Unfortunately, there is a fair amount of non-reporting for some of the socioeconomic controls, and as we do not want to lose observations due to missing socioeconomic information, we replace the missing control variable observations with zeros and include a set of dummy variable indicating missing data.[8]  It is also important to point out that the grade dummy captures both across grade learning and absolute age.  Accordingly, these two effects cannot be statistically separated in this framework.[9]

IV.B. ECLS Data for 3rd Grade

As explained in the previous sub-section, the TIMSS data for the United States are not usable because it does not include state of residence.  We therefore use the ECLS to examine the impact of relative age on math and science test scores in the United States.  However, one drawback of the ECLS is that the sampling frame is different from the TIMSS sampling structure.  Where TIMSS test scores are for a particular grade, the ECLS test scores are for a particular age.  More specifically, the ECLS sample includes children enrolled in kindergarten in 1998.  The ECLS then tracks these children through 2002, at which time most children are in grade three, but children who have failed a grade are in grade two (7.27 percent of the sample), and children who have failed two grades are in grade one (0.03 percent of the sample).  In order to make the ECLS and TIMSS samples as similar as possible, we restrict the sample to children who entered kindergarten for the first time in 1998 for whom we have complete information regarding sex, birth date, school cut-off date, school year start date, test scores, race, number of siblings, rural status, parental education, whether the child lives with both mother and father, owns over 100 books, and owns a calculator.  This leaves us with 6,091 observations.  Finally, to make to ECLS results as comparable as possible to the TIMSS results, the math and science scores are standardized to have a mean of 50 and a standard deviation of 10.

IV.C. <u>NELS Data for 8<sup>th</sup> Grade</u>

Since the available ECLS waves currently end at grade three, we turn to NELS to examine the impact of relative age on academic achievement at the eighth grade level.  NELS is a nationally representative sample of eighth-graders who were first surveyed in the spring of 1988.  Unlike the ECLS, the NELS sample structure is identical to the TIMSS sample structure. We restrict the sample to students with complete test score, birth date, sex, race, family size, living with both mother and father, urban, owning over 50 books, owning a calculator, and parental education data.  We further restrict the sample to states in which there exists a single precisely defined school cut-off date during the sample period.[10]  Finally, states with a mid-month cut-off are also excluded since we know the month of birth but not the day of birth.  This leaves us with 15,155 observations.  Again, to make the NELS and TIMSS estimates as comparable as possible, the math and science scores are standardized to have a mean of 50 and standard deviation of 10.

IV.D. <u>Is Assigned Relative Age Random or Do Some Parents Target "Old" Relative Ages?</u>

Before proceeding to the results section of the paper, it is important to examine the potential endogeneity of relative age due to parental birth date targeting aimed at ensuring that their child is the oldest in their class.  To investigate this possibility, Table II reports the fraction of children in TIMSS (years and age groups are combined) born in each calendar and school quarter.  Focusing first on the calendar quarter of birth, i.e. January – March is quarter 1 and October – December is quarter 4, it is clear that births are evenly distributed across calendar quarters.  Further, to the extent that any quarter is slightly favored, it is either 2 or 3 (spring or summer).

More interesting for our purposes are the relative school age patterns. While it might seem that we could detect birth date targeting by looking for elevated birth rates in the fourth school quarter, this is not possible due to the mild seasonality in birth patterns shown in the first four columns of Table II. Even mild seasonality makes detecting birth date targeting difficult, or impossible, because strictly speaking we expect parents who target to have a somewhat higher fourth school quarter birth rate than parents who do not target age at school entry.

The last four columns of Table II therefore explores the possibility that more educated mothers may target their child's birth date more than less educated mothers. More specifically, these columns report the difference in the percentage of children born in each relative quarter for more and less educated mothers. These results are based on the eighth grade sample because maternal education is not reported for fourth graders. More educated mothers are defined as those with education levels in the top 30-40 percent of the education distribution within their country.[11] Japan and England are excluded because they do not report maternal education information. The results reported in columns 9-12 reveal only a few (7 out of 60) statistically significant differences across maternal education groups. Further, the significant differences that do exist provide no support for the hypothesis that mothers that are more educated target birth dates in order to ensure that their children are the oldest in their class. In fact, if there is a pattern, albeit a very weak one, it is that more educated mothers are slightly more likely to target the first relative age quarter in many countries. A possible explanation for this pattern is that a small fraction of more educated mothers target summer birth dates, likely for work reasons.[12]

V.A. <u>Grade 4</u>

We begin the analysis by focusing on individual countries at the fourth grade level. Table III reports the results for all equations of interest for mathematics (columns 1-4) and science (columns 5-8). For comparative purposes, columns 1 and 5 report the OLS results for the impact of observed age on test scores, equation (1). Columns 2 and 6 (labeled RF) and 3 and 7 (labeled FS) report the reduced form and first stage estimates for $\theta_{2gc}$ and $\pi_{2cg}$, equations (3) and (2), respectively. Lastly, columns 4 and 8 report the IV estimates for $\beta_{cg}$ using assigned relative age as an instrument for observed age.

It is easiest to begin with the four countries for which there is little or no evidence of early/late starting or grade retention: England, Iceland, Japan, and Norway (these countries are shaded). For these countries, there are no, or at least very few, confounding factors to contend with when estimating relative age effects as essentially all children follow the entry rules and progress on time. Stated somewhat more formally, since the mapping from assigned relative age to observed age is almost exact for these countries, the reduced form estimate ($\theta_{2gc}$) and the IV estimate ($\beta_{gc}$) should be almost identical. Comparing columns 2 and 4 (or 6 and 8) shows this to be the case. More interestingly, the point estimates for the impact of relative age are large for all countries. Referring to IV estimates reported in column 4, one month of additional relative age increases the average math test score by 0.330, 0.258, 0.291, and 0.255 in England, Iceland, Japan, and Norway, respectively. These coefficients translate into average math test score premiums for the relatively oldest ($R=11$) compared to the relatively youngest ($R=0$) of 3.6, 2.8, 3.2, and 2.8 points on a standardized test with a mean of 50 and a standard deviation of 10. To put these numbers into perspective, they imply 12, 11, 12, and 11 percentile test score ranking

premiums for eleven months of relative age, for the respective countries – a massive advantage

by any metric (see Figure I). The reduced form percentile premium is the difference between the

mean predicted percentile evaluated at $R=11$ and the mean predicted percentile evaluated at $R=0$,

where the predicted percentiles are based on the empirical country-grade specific test score

distributions. The IV age premium is analogously defined, except that $R=11$ is replaced by

$A$=oldest on-time age and $R=0$ is replaced by $A$=youngest on-time age.

The remaining (not shaded) countries in Table III all have a sizeable minority of students

either ahead or, more likely, behind their assigned grade at the time of testing (see Appendix

Table I). Further, students who are behind are more likely to be relatively young and those who

are ahead of their assigned grade are more likely to be relatively old.[13] As such, assigned

relative age also affects test scores through grade acceleration and retention/late entry. While it

is impossible to distinguish between late entry and retention in TIMSS, evidence from the ECLS

in the United States shows that by grade three, 12.2 percent of children are behind and that 40.6

percent of this fraction is due to late entry and 59.4 percent is due to retention. In contrast, if we

restrict attention to the relatively youngest children in the United States (R=0), 41.4 percent of

children are behind and that 57.4 percent of this fraction is due to late entry and 42.6 percent is

due to retention. These percentages clearly show that the relatively youngest children are both

more likely to be held back by their parents and more likely to repeat a grade during primary

school, at least in the United States.

While there is substantial variation in the amount of retention and acceleration across

countries, there is a common pattern across assigned relative ages: Younger children are more

likely to be behind their assigned grade and older children are more likely to be ahead of their

assigned grade. As such, the results reported in Table III are the overall, or net, impact of

assigned relative age on test scores. They are net in the sense that children who have been retained have had an extra year of education at the time of the test, and their score includes this investment. Accordingly, as relatively younger children are more likely to be retained, relative age affects test scores both through within grade differences as well as through retention. Note that the reverse is not generally true for those who have been accelerated. As most acceleration occurs at school entry, these children are younger, and hence may score worse as a result, but they have not lost a year of training. The reduced form estimates are important from a policy perspective because they measure the impact of relative age at a specific grade and therefore incorporate both within grade differences in age and retention differences across relative ages. In other words, if retention is partly determined by relative age, and if retention is effective at raising human capital, then the reduced form relative age effects estimates should be smaller than the IV estimates because the latter exclude the impact of retention since the age of children observed ahead or behind their assigned grade cannot be predicted by assigned relative age.

Several features of Table III warrant comment. First, as discussed in Section III, the OLS estimates are substantially downward biased in countries with high behind rates because children who are old due to retention (who tend to score poorly) substantially out number children whose parents hold them out of school so that they enter kindergarten a year late (who are positively selected). The exception is the United States, which has a strong positive OLS estimate despite a reasonably high behind rate. This reflects the difference in the sampling frame. In the ECLS, students are observed in the spring of their fourth year of school regardless of grade. As such, observed young students – who are the most likely to be retained – score poorly both because they are young and because they have not yet completed as much curriculum and the small number of observed old students whose parents entered them into kindergarten a year late tend to

be positively selected and hence score highly. Second, all of the reduced form and IV estimates are statistically significant at the conventional level. Third, in all but the "clean" countries, the IV estimates are larger than the reduced form estimates, although in some cases the difference is not statistically significant. These findings suggest that retention may partly ameliorate the disadvantage of being relatively young, at least in some countries, and at least in terms of test scores at the fourth grade level.[14] Fourth, and most importantly, the size of the relative age premium is large in all countries at the fourth grade level. Finally, the Table III results are consistent with the Puhani and Weber [2005] estimates showing that relatively old German students score 0.4 of a standard deviation higher on the Progress in International Reading Literacy Study (PIRLS) exam at the end of grade four.[15]

In order to more easily describe the magnitude of the relative age premium, panel A in Figure I graphs the reduced form and IV estimates of the math test score percentile premium for the oldest students ($R=11$) compared to the youngest students ($R=0$). There are two general groups – countries with low failure rates and countries with high failure rates. In general, the low failure rate countries (England, Iceland, Japan, and Norway) have large relative age effects, with similar reduced form and IV estimates. For these countries, the oldest children score 11-12 percentiles higher than the youngest children. Among the high failure rate countries the pattern is quite different: The reduced form estimates are smaller than the IV estimates and the relative age effects tend to be somewhat smaller than in the low failure rate countries. The IV point estimates generally range from a 2-3 point advantage (6-10 percentiles) for the oldest children and a reduced form point estimate advantage generally ranging from 1-2 points (4-8 percentiles). The one anomaly is New Zealand. The IV estimate for New Zealand is extremely large; a 4.7-point (17 percentile) advantage for the oldest children, but the reduced form estimate is only 2.2

(an 8 percentile difference) – by far the biggest differential between the two estimates. The anomaly is that the behind rate in New Zealand is only 7 percent. However, New Zealand has a high acceleration rate, 5 percent, which is likely driving the large difference in the two estimates.

V.B. <u>Grade 8</u>

As discussed in Section I, our primary objective is to estimate the persistence of relative age effects. While one might expect a few months of relative maturity to impact performance during the primary grades, it is less clear how important this might be at older ages. Its magnitude depends on the interaction of relative age and the structure of the education system, the degree to which human capital accumulated in early childhood is complementary to later human capital accumulation, and the extent to which young students and their teachers exert additional effort to bring the achievement of young students up to that of their older peers.

Table IV replicates Table III for the eighth grade sample. The main finding for the eighth grade sample is that relative age continues to be an important determinant of test scores even at the end of middle school and the beginning of secondary school. In general, the IV point estimates range from approximately 0.13 to 0.38, which is consistent with the statistically significant positive secondary school results reported by Fredriksson and Öckert [2004] for Sweden. Unlike the fourth grade sample, a few of the coefficient estimates are statistically insignificant. In particular, there is no evidence of relative age effects in Denmark or Finland. However, one should expect weak relative age effects in countries where formal curriculum-based education begins later because initial age differences will be less important. In contrast to almost all other countries in the sample, compulsory education does not begin until age seven in Finland and even then, the initial grades emphasize play and personal development rather than curriculum-based activities. And in Denmark, differentiation on the basis of ability is officially

prohibited before the age of sixteen [OFSTED 2003]. Given these educational features, it is not be surprising that we fail to find relative age effects in these countries.

The other substantive difference between Tables III and IV is that the United States sampling frame in NELS matches that of TIMSS. As such, all United States point estimates reported in Table IV are comparable to those of other countries. As one would expect, the United States OLS estimates are now negative, the reduced form and IV estimates are now similar to the Canadian estimates, and the gap between the reduced form and IV estimate is fairly large. Based on the reduced form (IV) coefficient, the relatively oldest United States students score 1.1 (2.6) points higher in mathematics than the relatively youngest students, or 4 (8) percentiles higher.

Again for interpretive ease, panel B in Figure I graphs the reduced form and IV estimates of the mathematics test score percentile premium for the oldest students ($R$=11) compared to the youngest students ($R$=0). The results are similar to those shown in panel A in the sense that countries with lower failure rates tend to have larger relative age effects (one exception is Portugal), the difference between the reduced form and IV estimates are largest in countries with high failure rates, and New Zealand is again an outlier. The main difference between the fourth and eighth grade estimates is the magnitude of the relative age effect. While the test score premium enjoyed by the relatively oldest students falls in all countries, it remains economically important at the eighth grade level in almost all countries. The oldest students continue to score 4 or more percentiles higher than the youngest students (based on the reduced form) in fourteen of the nineteen countries.

Overall, the results presented in Table IV point to the persistence of sizeable relative age effects into adolescence across almost all countries. This finding is important because it shows

that early relative age effects are propagated by a wide range of educational structures: From the Japanese system of automatic promotion, to the accomplishment oriented French system, to the supposedly more flexible skill-based program models used in Canada and the United States.

The results reported above also speak to the question of whether relative age, birth date combined with school starting and stopping rules, is a legitimate instrument for education in adult outcome regressions (i.e. wages, employment, or even offspring health). An identification strategy of this type requires that there is no direct association between birth date and the outcome of interest (see Angrist and Krueger [1991] or Bound, Jaeger, and Baker [1995]). In contrast to this requirement, the results presented in this paper indicate that relative age has a direct impact on test scores as late as the end of middle school or early high school, as well as on college enrollment. In subsequent work, Dhuey and Lipscomb [2005] also show that relatively old children are more likely to be leaders in high school, which has in turn been shown to increase wages in adulthood [Kuhn and Weinberger 2005]. As such, relative age appears to have a direct effect on human capital accumulation holding educational attainment constant and is therefore likely to have a direct impact on adult outcomes such as wages, independent of its effect through educational attainment. While the age effects listed above should bias IV estimates of the return to education downward it is certainly possible that there are other existing age effects working in the opposite direction. Further complicating matters is the fact that the relationship between relative age and educational attainment is non-monotonic. While older students may be more likely to drop out of high school because they reach the compulsory school age at lower levels of educational attainment, older students are also more likely to attend academic universities (see Section VI).

V.C. Potential Season of Birth Effects

Although the fact that school starting rules differ across countries gives us substantial confidence that we are not confounding relative age effects with season of birth effects, in this section we take advantage of the cross-national dimension of TIMSS to control for season of birth effects directly by estimating an alternative specification that includes both assigned relative age and month of birth using data pooled across countries with cut-off dates in different months (equation (4)).[16] The results are reported in Table V. In all cases, the reported coefficients are for relative age (RF and FS) or observed age (OLS and IV). In addition, the set of controls ($X$) listed in Section IV are included in all models. For comparative purposes, row A reports a base model that includes only relative age and country indicators. The point estimates in this row reflect the average reduced form and IV estimates across the entire international sample with no controls for season of birth. In all cases, these estimates are about what one would have guessed by averaging the country specific estimates reported in Tables III and IV.

In order to control for season of birth effects, row B further includes a vector of birth month indicators. A comparison of rows A and B reveals that both the reduced form and IV estimates are remarkably robust to the inclusion of season of birth. The point estimates reported in rows A and B are very close in all cases. Referring to IV estimates reported in column 4, one month of additional relative age increases the average math test score by 0.243, which translates into an average math test score premium for the relatively oldest ($R$=11) compared to the relatively youngest ($R$=0) of 2.7 points on an internationally standardized test with a mean of 50 and a standard deviation of 10.

Finally, row C further adds the age at which each child is eligible for school entry to the list of controls. This variable is calculated using the school entry age and cut-off dates reported

22

in Appendix Table I combined with month of birth. This variable captures the absolute age at which children are eligible to enter school, which differs across birth dates and countries. It is important to separate absolute and relative age if, for example, human capital accumulation rates depend on age. Since there is no theoretical reason to impose any particular functional form on absolute age, it enters the row C specification as unconstrained months of absolute age indicators. While the point estimates in row C are slightly smaller compared to the specification that includes relative age and month of birth but not absolute age at school eligibility (row B), the differences are never statistically significant.

## VI. THE IMPACT OF RELATIVE AGE ON PRE-UNIVERSITY BEHAVIOR
## IN THE UNITED STATES AND BRITISH COLUMBIA

The strength of the relative age effects estimated across a wide range of countries at the fourth and eighth grade levels lead one to wonder how far these effects propagate themselves forward. In fact, the continuing strength of relative age in the eighth grade sample, when most children are thirteen or fourteen years of age, suggests that relative age may play a role in determining educational success throughout the educational process – even into college. Unfortunately, it is difficult to find data sources that contain completed education and birth dates. One exception is the 1980 United States Census, which reports quarter of birth. However, it is difficult to analyze the long run relative age effects during the eras covered by this data because they are confounded by the large fraction of students dropping out immediately upon reaching the school leaving age as well as by the G.I. Bills instituted after World War II and the Korean War. We are, however, aware of two recent studies examining the impact relative age and educational attainment. Consistent with the estimates discussed in the next two sub-sections,

both Fredriksson and Öckert [2004] and Puhani and Weber [2005] find that relatively old students attain higher levels of education in Sweden and Germany, respectively.

Although we do not have representative micro data that include birth dates and higher education choices for all of the countries included in the analysis so far, we have obtained micro-level data that includes birth dates and pre-university choices for the Canadian province of British Columbia and data on birth dates, pre-university choices, and university enrollment for the United States. The British Columbia data is a comprehensive restricted-use administrative data set including every student enrolled in a public school on an annual basis compiled by the Ministry of Education and maintained by Edudata Canada at the University of British Columbia. The United States data comes from the second and third waves of the restricted-use version of NELS data discussed earlier in the paper.

VI.A. Pre-University Behavior in British Columbia

The best available data on pre-university academic behavior during the final years of high school comes from British Columbia, Canada. The Ministry of Education tracks student enrollment by grade each October. We therefore have panel data reporting each student's grade level as of October for a nine-year period (1995-2003). In addition, we also know each student's birth date, sex, and school of enrollment in each year that they are enrolled.[17] Moreover, since these data include all students enrolled in the British Columbia education system, non-random selection is not a concern. Further, since all Canadian provinces have a January 1 cut-off date, we can be quite confident that we have correctly measured relative age in almost all cases.[18] However, as this is administrative data and not a panel survey, we have no information for students who leave the British Columbia school system, either because they move to a different province or because they drop out of school. This does not cause a problem as long as the

probability that a youth leaves the British Columbia school system because they move to another province is independent of their birth month.[19]

Unfortunately, given the structure of the available data, we cannot track individual students over their entire public school career. We can, however, track individuals who entered grade nine for the first time from 1996-1998 for five years; this allows one extra year to finish high school so that we do not misclassify students who take an extra year complete high school. We restrict the sample to students observed in grade nine to avoid non-random sampling. Under this sampling strategy, we have the universe of students who enter grade nine regardless of whether or not they finish high school.[20] This is particularly important because it is possible that relative age affects high school dropout behavior.

We then define students as "university-bound" if we observe them in grade 12 during the appropriate sample period, if they report having graduated by June of the fifth year after they enter grade nine and have taken at least four examinable subjects and earned a 75 percent average or higher.[21] It is important to point out that this definition reflects the fact that we are interested in identifying highly successful students who are bound for purely academic institutions of higher learning rather than vocational or lower-end academic programs. Guided by this objective, the aforementioned "university-bound" definitions capture two important features of the British Columbia education system. First, most academic, or pre-university, courses require students to take a province-wide final examination that is written and graded by the Ministry of Education. During the period of interest, the grade on these exams constituted 40 percent of the final grade for these courses, with the other 60 percent being assigned at the school-level by teachers. Secondly, in order to gain admission into one of the flagship provincial universities students must generally take at least four examinable subjects and score in excess of

75 percent – this is just an approximation, not a formal rule set out by the universities.  As a sensitivity check, we also alternatively define an individual as university-bound if they if we observe them in grade 12 during the appropriate sample period, they report having graduated by June of the fifth year after they enter grade nine, they take at least four examinable subjects, and they earn a 75 percent or higher average across all attempted provincial exams, rather than 75 percent across examinable courses.

Since we have few control variables, but can identify schools, we estimate the model using school fixed effects.  The reduced form is given by:

(5) $C_{syi} = \psi_{1s} + \psi_{2y} + \psi_3 R_{syi} + \psi_4 F_{syi} + v_{syi}$

where $\psi_{1s}$ is a vector of school fixed effects, defined by the school of record in grade nine, $\psi_{2y}$ is a vector of ninth grade year indicators, and $F$ is a female indicator.  Following the analysis in all previous sections, we report the OLS, first stage, reduced form, and IV estimates in Table VI. Whenever actual age is used, it refers to the individual's age the first time they are observed in grade nine.  In all cases, the models are estimated using a linear probability model and the reported standard errors are heteroskedastic consistent.

Column 1 reports the fraction of students defined as university-bound, and columns 2-5 report the OLS, first stage, reduced form, and IV estimates for relative age.  Based on our primary definition of "university-bound," 18 percent of British Columbia students are defined as university-bound.  To give the reader some perspective, the university participation rate among 18-21 year olds in British Columbia was 13.3 percent in 2002/2003.[22]  More importantly, the reduced form estimates reported in column (3) reveal that the relatively oldest students are 1.8 percentage points, or 9.8 percent, more likely to be university-bound than the relatively

youngest.  As one would expect, the IV estimates are somewhat larger – the relatively oldest are 12.8 percent more likely to be university-bound than the relatively youngest.

VI.B. <u>Pre-University and University Enrollment Behavior in the United States</u>

To examine pre-university and university enrollment behavior in the United States, we return to the restricted-use NELS data discussed earlier in the paper.  In contrast to the eighth grade results, which are based on the first wave of the survey, in this section we use the second and third waves conducted in 1992 and 1994, respectively.  While the waves are different, the sample exclusions and the control variables are identical to those described in Section IV.C.[23]

The second wave of NELS was conducted in 1992 while the respondents were in grade twelve.  As we are interested in the impact of relative age on "academic" success, we are consequently interested in separating all college going, including vocational and/or quasi-academic programs, from high-level academic degree pursuits.  The most obvious indicator of such aspirations reported in the second wave of NELS is whether the respondent has taken the SAT exam, ACT exam, or both.  This is an appealing measure because most flagship colleges and universities require a SAT or ACT score as part of their application process.  In a similar vein, data from the third wave of NELS, conducted in 1994, almost two years after the on-time high school graduation date, reports whether respondents have enrolled in a four-year accredited college/university during the previous two years.  Again, this is an appealing measure because it isolates academic post-secondary enrollment.

The results for the same set of models reported for British Columbia are also reported for the United States university measures in rows 3 and 4 in Table VI.  Column 1 reports the fraction of students designated as belonging to the group of interest (having taken the SAT or ACT or enrolling in an accredited four-year postsecondary institution), and columns 2-5 report the OLS,

first stage, reduced form, and IV estimates for relative age. The reduced form estimates reported in column (3) reveal that the relatively oldest students are 4.6 percentage points, or 7.7 percent more likely to have taken the SAT or ACT, and 4.6 percentage points, or 11.6 percent, more likely to enroll in an accredited four-year college/university. In addition, as one would expect, the IV estimates are somewhat larger – the relatively oldest are 11.1 (10.7) percentage points more likely to have taken the SAT or ACT (enrolled in college) than the relatively youngest.

## VII. CONCLUSION

The large relative age effects reported in this paper are the result of what many might consider a necessary and benign educational construct: a single school starting date. Despite the theoretical possibility of long-run relative age effects, one might have expected these effects to exist early in the educational process, but then dissipate with age, making the single start date seem both easy and innocuous. However, the results reported in this paper cast serious doubt on this view. In particular, we find that the oldest students score 4-12 percentiles higher than the youngest students at the fourth grade level and 2-9 percentiles higher at the eighth grade level across a wide range of countries. The TIMSS results are corroborated by the fact that relatively older students in British Columbia and the United States are more likely to participate in pre-university academic programs during the final years of high school, and more likely to enter a flagship postsecondary institution in the United States. Further, the findings reported in this paper are consistent with recent work by Heckman and various coauthors (see Cuhna, Heckman, Lochner, and Masterov [2006]) showing that over the childhood lifecycle skills beget skills through a multiplier process: Skills accumulated early in childhood are complementary to later learning.

The fact that early advantages held by relatively old children persist into adulthood through differences in skill accumulation, college preparation, and the accumulation of softer skills, such as leadership (see Dhuey and Lipscomb [2005]), suggests that we should pay close attention to early school structures that affect skill acquisition. For example, we should learn more about the long-run differences in educational outcomes in countries with no, or at least limited, ability differentiated learning groups during the primary grades like Denmark and Finland [OFSTED 2003], compared to countries where children are separated into ability-specific curriculum groups during the early primary grades, such as England and the United States [Mullis et al. 2002 and 2003].

The finding of long-run relative age effects also has potentially important implications for the distribution of skills across socioeconomic groups in countries that allow parents to defer school entry. In the United States, for example, 5 percent of children enter kindergarten a year later than they are eligible to do so. While this may seem innocent enough, 77 percent of deferrals are for students born in the last relative quarter, and 30 percent of these children are from the top quarter of the socioeconomic distribution. As such, low socioeconomic children now make up a disproportionately large share of the relatively youngest quarter, which is a cause for concern, since these children are now at two substantial disadvantages.

Department of Economics, University of California at Santa Barbara

Department of Economics, University of California at Santa Barbara

**REFERENCES**

Allen, Jeremiah and Roger Barnsley, "Streams and Tiers: The Interaction of Ability, Maturity, and Training in Systems with Age-Dependent Recursive Selection," *The Journal of Human Resources*, XXVIII (1993), 649-659.

Angrist, Joshua A. and Alan B. Krueger, "Does Compulsory School Attendance Affect Schooling and Earnings?" *The Quarterly Journal of Economics*, CVI (1991), 979-1014.

Arnove, Robert F. and Enid Zimmerman, "Dynamic Tensions in Ability Grouping: A Comparative Perspective and Critical Analysis," *Educational Horizons*, LXXVII (1999), 120-127.

Bound, John and David A. Jaeger, "Do Compulsory School Attendance Laws Alone Explain the Association Between Quarter of Birth and Earnings?" *Research in Labor Economics*, XIX (2000), 83-108.

Bound, John, David A. Jaeger and Regina M. Baker, "Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, XC (1995), 443-450.

Cascio, Elizabeth and Ethan Lewis, "Schooling and the Armed Forces Qualifying Test: Evidence from School Entry Laws," *Journal of Human Resources,* XLI (2006), forthcoming.

Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov, "Interpreting the Evidence on Life Cycle Skill Formation" *Handbook of the Economics of Education*, (E. Hanushek and F.Welch, editors, North Holland, 2006), forthcoming.

Datar, Ashlesha, "Does Delaying Kindergarten Entrance Give Children a Head Start?' *Economics of Education Review*, XXV (2006), 43-62.

de Cos, Patricia, "Readiness for Kindergarten: What Does It Mean?" California Research Bureau, California State Library, CRB-97-014, (1997).

Dhuey, Elizabeth and Stephen Lipscomb, "What Makes a Leader? Relative Age and High School Leadership" University of California, Santa Barbara Working Paper, (2005).

Fredricksson, Peter and Björn Öckert, "Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labor Market Performance," Working Paper, (2004).

Jacob, Brian A. and Lars Lefgren, "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis," *Review of Economics and Statistics* LXXXVI (2004), 226-244.

Kuhn, Peter and Catherine Weinberger, "Leadership Skills and Wages," *Journal of Labor Economics*, XXIII (2005), 395-436.

McCrary, Justin and Heather Royer, "The Effect of Maternal Education on Fertility and Infant Health: Evidence from School Entry Laws Using Exact Date of Birth," Working Paper, (2005).

Mullis, Ina, Michael Martin, Ann Kennedy, and Cheryl Flaherty, *PIRLS 2001 Encyclopedia*. (Boston, MA: International Study Center-Boston College, 2002).

Mullis, Ina, Michael Martin, Eugenio Gonzalez and Steven Chrostowski, *TIMSS 2003 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades,* (TIMSS and PIRLS International Study Center, 2003).

OECD, *OECD Handbook for Internationally Comparative Education Statistics: Concepts, Standards, Definitions and Classifications,* (OECD, 2004).

Office for Standards in Education (OFSTED), *The Education of 6 year olds in Denmark, England, and Finland, an International Comparison*, (OFSTED Publications Centre, http://www.ofsted.gov.uk/, 2003).

Puhani, Patrick and Andrea Weber, "Does the Early Bird Catch the Worm? Instrumental Variable Estimates of Educational Effects of Age of School Entry in Germany." IZA Discussion Paper # 1827, (2005).

Stevenson, Harold W. and James W. Stigler, *The Learning Gap,* (New York, NY: Summit Books, 1992).

---

[1] Early performance gaps by any identifiable group, e.g. race, socioeconomic status, or gender, can similarly propagate themselves through time via the same multiplier process.

[2] England and New Zealand have multiple kindergarten entry dates, but a single first grade entry point. Although United States states set their own cut-off dates, and hence there are many cut-off dates in the United States, in most jurisdictions a single date applies to all residents (see Appendix Table II).

[3] We thank two anonymous referees for making this suggestion.

[4] There is also a large education literature on age effects (see de Cos [1997] and the references therein).

[5] Both countries have an unreasonably large percentage of births occurring in January and February. The percentage of births occurring in December, January, February, and March are: 8.3, 10.4, 10.5, and 7.9 in Korea and 5.7, 12.0, 8.0, and 8.2 in Turkey. According to colleagues from Korea it is not uncommon for parents with relatively old children (born from March-June) to bribe officials to enter an incorrect birth date on the birth certificate so that their child are eligible for school early. This means that some relative fourth quarter births are erroneously reclassified as relative first quarter births. Conversely, according to colleagues from Turkey, it not uncommon for the parents of children born in December to wait until January to register them so that they will not start school early with older children (the reverse of the motivation in Korea).

[6] In all cases, except the Italian eighth grade sample, even youth who have been retained once have not yet reached the compulsory schooling age (see www.right-to-education.org). As youth in Italy can leave school at age fourteen, it is possible that we have a non-random eighth grade sample for Italy. However, this possibility does not seem to be concern since all estimates are similar if we restrict the Italian sample to seventh graders.

[7] All results are similar if internationally standardized test scores are used instead.

[8] The results are not specification specific. All results are similar if we exclude observations with missing data or if we include all observations but only sex, grade, and test year indicators, or no covariates at all.

[9] Cascio and Lewis [2006] isolate the impact of an additional year of schooling on AFQT scores using NLSY data.

[10] States are excluded if they change school start dates during the period of interest, list their cut-off date as the start of the school year because we do not have a complete historical listing of school start dates, or allow local education authorities to set their own cut-off date (see Appendix Table II).

[11] The fraction of mothers in the more and less educated groups is not always evenly split because TIMSS reports maternal education in six categories, and in some cases a single category includes a large percentage of mothers. More detail is available in the working paper version of the paper.

[12] One potential concern with Table II is that the small sample sizes in TIMSS (2,920 to 25,063) might make it difficult to detect small differences in birth patterns across maternal education groups. We therefore repeated the analysis using the United States Natality Detail Files for 1997-1999 and find similar results. (These results are available in the working paper version of the paper.)

[13] See the working version of the paper for a detailed analysis of the impact of relative age on retention and acceleration.

[14] These results are consistent with the finding of Jacob and Lefgren [2004] that retention in Chicago public schools successfully increases third grade test scores. Further consistent with Jacob and Lefgren [2004], we will see in Section V.B, that in contrast to the fourth grade results, we can rarely reject that the IV and reduced form results are statistically different at the eighth grade level – Jacob and Lefgren [2004] similarly find that sixth grade test scores are not increased by retention. Given these results, it seems safe to say that in the short run primary grade retention increases student achievement, but that the long run impact is much less clear.

[15] The results for specifications including non-linear relative measures and male-female sub-samples are available in the working paper version of the paper.

[16] New Zealand is excluded from the cross-country analysis due to the reversal of seasons in the southern hemisphere. However, all results are similar if New Zealand is included, and are available upon request. The United States is excluded because the data do not come from the TIMSS sample.

[17] Both the student and school identifiers included in the data are encoded to ensure confidentiality. Further, due to confidentiality concerns, Aboriginal and ESL students were removed from the data before we received it.

[18] The only exception is foreign born students who enter the British Columbia education system prior to grade nine, but as we exclude ESL students, this only leaves English speaking foreign born students who entered the British Columbia education system between kindergarten and grade nine, which is a small number.

[19] However, we make no such assumption regarding the probability that a student drops out of high school; recent studies by Angrist and Krueger [1991] Cascio and Lewis [2006] and McCrary and Royer [2005] show that educational attainment and particularly dropping out are a function of birth month due to school entry age cut-offs.

Table I
Summary Statistics

| | Fourth Grade | | | | | Eighth Grade | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Math (1) | Science (2) | Relative Age (3) | Observed Age (4) | Sample (5) | Math (6) | Science (7) | Relative Age (8) | Observed Age (9) | Sample (10) |
| Austria | 52.35 (9.27) | 52.57 (8.55) | 5.38 (3.45) | 119.61 (8.50) | 5,045 | 52.69 (8.85) | 53.28 (9.20) | 5.42 (3.47) | 165.63 (8.50) | 5,528 |
| Belgium | | | | | | 55.21 (8.45) | 51.27 (8.81) | 5.52 (3.41) | 165.45 (9.57) | 15,650 |
| Canada | 50.23 (9.30) | 51.36 (8.89) | 5.58 (3.41) | 114.61 (9.00) | 15,588 | 52.69 (8.57) | 52.26 (8.89) | 5.46 (3.40) | 165.13 (8.60) | 25,062 |
| Czech Republic | 53.18 (9.31) | 51.80 (8.53) | 5.44 (3.44) | 119.00 (8.11) | 6,523 | 54.30 (8.65) | 54.52 (8.15) | 5.40 (3.44) | 168.31 (7.69) | 10,123 |
| Denmark | | | | | | 49.07 (8.51) | 46.68 (8.94) | 5.68 (3.40) | 160.99 (7.70) | 4,251 |
| England | 47.88 (9.76) | 51.29 (9.78) | 5.54 (3.45) | 114.53 (6.90) | 6,066 | 50.21 (9.25) | 53.25 (9.65) | 5.58 (3.48) | 165.56 (7.17) | 6,515 |
| Finland | | | | | | 53.96 (7.33) | 54.07 (7.94) | 5.66 (3.37) | 166.11 (4.23) | 2,920 |
| France | | | | | | 51.90 (8.09) | 47.76 (8.31) | 5.49 (3.40) | 165.56 (10.91) | 5,610 |
| Greece | 46.48 (9.92) | 47.25 (8.95) | 5.80 (3.40) | 109.53 (7.46) | 5,952 | 46.93 (9.10) | 47.67 (9.03) | 5.65 (3.45) | 157.26 (8.84) | 7,800 |
| Iceland | 44.32 (8.70) | 46.16 (9.07) | 5.43 (3.36) | 109.50 (6.92) | 3,431 | 48.00 (7.97) | 48.03 (8.44) | 5.59 (3.40) | 157.54 (7.03) | 3,718 |
| Italy | | | | | | 49.31 (9.06) | 49.99 (8.77) | 5.57 (3.40) | 163.96 (8.54) | 8,163 |
| Japan | 55.44 (8.03) | 53.87 (7.43) | 5.60 (3.41) | 118.70 (6.89) | 8,536 | 58.00 (8.15) | 54.60 (8.31) | 5.58 (3.41) | 168.51 (6.63) | 14,889 |
| New Zealand | 47.22 (9.70) | 49.76 (9.78) | 5.43 (3.47) | 113.93 (7.27) | 4,916 | 50.14 (9.13) | 50.99 (9.42) | 5.43 (3.47) | 168.27 (10.93) | 10,476 |
| Norway | 46.37 (9.10) | 48.69 (9.65) | 5.60 (3.37) | 112.17 (7.14) | 4,300 | 48.80 (8.56) | 50.55 (8.82) | 5.74 (3.35) | 160.25 (7.23) | 5,644 |
| Portugal | 45.51 (9.28) | 45.57 (9.55) | 5.45 (3.42) | 117.26 (12.92) | 5,451 | 44.59 (7.10) | 46.04 (8.21) | 5.47 (3.42) | 167.56 (13.36) | 6,600 |
| Slovak Republic | | | | | | 53.83 (8.69) | 52.64 (8.36) | 5.42 (3.46) | 167.04 (7.23) | 10,597 |
| Spain | | | | | | 47.48 (8.07) | 49.95 (8.32) | 5.49 (3.42) | 164.99 (10.31) | 7,595 |
| Sweden | | | | | | 52.05 (8.97) | 52.68 (9.16) | 5.75 (3.38) | 167.21 (10.56) | 8,823 |
| United States* | -- | -- | 5.90 (3.56) | 102.59 (4.09) | 6,091 | -- | -- | 5.36 (3.47) | 170.19 (7.20) | 15,155 |

Test scores are internationally standardized to mean 50 and standard deviation 10. Sample means are population weighted. *U.S. data are from the ECLS for grade 3 and NELS for grade 8. All other data are from TIMSS.

Table II
Quarter of Birth Patterns

| | Calendar Quarter of Birth | | | | School Quarter of Birth | | | | Maternal Education Difference (More-Less) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Q1 (1) | Q2 (2) | Q3 (3) | Q4 (4) | Q1 (5) | Q2 (6) | Q3 (7) | Q4 (8) | Q1 (9) | Q2 (10) | Q3 (11) | Q4 (12) |
| Austria | 0.250 | 0.250 | 0.257 | 0.244 | 0.262 | 0.247 | 0.251 | 0.241 | -- | -- | -- | -- |
| Belgium | 0.244 | 0.263 | 0.253 | 0.241 | 0.241 | 0.253 | 0.263 | 0.244 | 0.016 | 0.000 | -0.009 | -0.007 |
| Canada | 0.244 | 0.258 | 0.256 | 0.243 | 0.243 | 0.256 | 0.258 | 0.244 | 0.009 | 0.005 | 0.001 | **-0.015** |
| Czech Republic | 0.257 | 0.260 | 0.255 | 0.229 | 0.254 | 0.266 | 0.242 | 0.238 | **0.023** | 0.015 | **-0.028** | -0.010 |
| Denmark | 0.256 | 0.273 | 0.244 | 0.227 | 0.227 | 0.244 | 0.273 | 0.256 | 0.010 | **-0.040** | 0.005 | 0.025 |
| England | 0.246 | 0.250 | 0.254 | 0.251 | 0.241 | 0.257 | 0.245 | 0.257 | -- | -- | -- | -- |
| Finland | 0.251 | 0.273 | 0.256 | 0.220 | 0.220 | 0.256 | 0.273 | 0.251 | 0.002 | 0.006 | -0.025 | 0.016 |
| France | 0.237 | 0.268 | 0.254 | 0.241 | 0.241 | 0.254 | 0.268 | 0.237 | 0.000 | 0.033 | -0.011 | -0.022 |
| Greece | 0.229 | 0.266 | 0.268 | 0.237 | 0.241 | 0.250 | 0.267 | 0.242 | -0.008 | 0.021 | -0.008 | -0.006 |
| Iceland | 0.240 | 0.261 | 0.261 | 0.238 | 0.238 | 0.261 | 0.261 | 0.240 | 0.005 | 0.000 | -0.020 | 0.015 |
| Italy | 0.248 | 0.268 | 0.246 | 0.237 | 0.237 | 0.246 | 0.268 | 0.248 | 0.006 | -0.019 | 0.020 | -0.008 |
| Japan | 0.239 | 0.247 | 0.271 | 0.244 | 0.239 | 0.244 | 0.271 | 0.247 | -- | -- | -- | -- |
| New Zealand | 0.244 | 0.254 | 0.247 | 0.255 | 0.256 | 0.246 | 0.257 | 0.241 | -0.008 | 0.015 | 0.001 | -0.007 |
| Norway | 0.249 | 0.283 | 0.246 | 0.222 | 0.222 | 0.246 | 0.283 | 0.249 | 0.007 | -0.019 | -0.001 | 0.013 |
| Portugal | 0.238 | 0.257 | 0.259 | 0.246 | 0.246 | 0.259 | 0.257 | 0.238 | -0.018 | -0.007 | **0.027** | -0.002 |
| Slovak Republic | 0.254 | 0.255 | 0.263 | 0.228 | 0.255 | 0.259 | 0.245 | 0.240 | 0.012 | 0.007 | -0.001 | -0.018 |
| Spain | 0.241 | 0.257 | 0.254 | 0.248 | 0.248 | 0.254 | 0.257 | 0.241 | 0.022 | 0.005 | 0.004 | **-0.031** |
| Sweden | 0.265 | 0.268 | 0.252 | 0.215 | 0.215 | 0.252 | 0.268 | 0.265 | -0.001 | **0.027** | -0.002 | -0.024 |

All data are pooled across grades 4 and 8 and are from TIMSS. All proportions are population weighted. The shaded values in columns 1-8 indicate the quarter with the highest fraction of births. The bold coefficients in columns 9-12 indicate that the difference between the fraction of births to "more" educated mothers differs from the fraction of births to "less" educated mothers at the 5 percent level or better. The results reported in columns 9-12 are based on the eighth grade sample (maternal education is not reported in the fourth grade sample), and Austria, England, and Japan are excluded because they do not report maternal education in the eighth grade sample.

Table III

The Impact of Relative Age on Test Scores at the Fourth Grade Level

| | Math | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| | OLS (1) | RF (2) | FS (3) | IV (4) | OLS (5) | RF (6) | FS (7) | IV (8) |
| Austria | **-0.262** | **0.108** | **0.540** | **0.201** | **-0.196** | **0.149** | **0.540** | **0.276** |
| | (0.031) | (0.049) | (0.031) | (0.097) | (0.031) | (0.048) | (0.031) | (0.096) |
| Canada | 0.010 | **0.115** | **0.607** | **0.190** | -0.011 | **0.110** | **0.607** | **0.181** |
| | (0.018) | (0.036) | (0.025) | (0.059) | (0.019) | (0.036) | (0.025) | (0.060) |
| Czech Republic | **-0.200** | **0.111** | **0.498** | **0.223** | **-0.127** | **0.180** | **0.498** | **0.362** |
| | (0.023) | (0.034) | (0.019) | (0.069) | (0.024) | (0.034) | (0.019) | (0.070) |
| England | **0.274** | **0.315** | **0.955** | **0.330** | **0.250** | **0.267** | **0.955** | **0.280** |
| | (0.034) | (0.036) | (0.008) | (0.038) | (0.033) | (0.036) | (0.008) | (0.038) |
| Greece | **0.131** | **0.196** | **0.894** | **0.219** | **0.130** | **0.245** | **0.894** | **0.274** |
| | (0.031) | (0.040) | (0.016) | (0.044) | (0.030) | (0.040) | (0.016) | (0.045) |
| Iceland | **0.183** | **0.254** | **0.984** | **0.258** | **0.180** | **0.247** | **0.984** | **0.251** |
| | (0.046) | (0.050) | (0.008) | (0.051) | (0.047) | (0.050) | (0.008) | (0.051) |
| Japan | **0.244** | **0.280** | **0.962** | **0.291** | **0.306** | **0.353** | **0.962** | **0.367** |
| | (0.031) | (0.031) | (0.005) | (0.032) | (0.032) | (0.031) | (0.005) | (0.032) |
| New Zealand | **0.081** | **0.202** | **0.469** | **0.430** | **0.078** | **0.173** | **0.469** | **0.369** |
| | (0.033) | (0.039) | (0.023) | (0.086) | (0.033) | (0.040) | (0.023) | (0.086) |
| Norway | **0.178** | **0.238** | **0.933** | **0.255** | **0.164** | **0.248** | **0.933** | **0.265** |
| | (0.039) | (0.041) | (0.011) | (0.044) | (0.040) | (0.042) | (0.011) | (0.046) |
| Portugal | **-0.096** | **0.195** | **0.758** | **0.258** | **-0.073** | **0.191** | **0.758** | **0.252** |
| | (0.012) | (0.038) | (0.042) | (0.054) | (0.013) | (0.039) | (0.042) | (0.054) |
| United States* | **0.277** | **0.230** | **0.774** | **0.298** | **0.341** | **0.267** | **0.774** | **0.345** |
| | (0.038) | (0.040) | (0.016) | (0.051) | (0.034) | (0.038) | (0.016) | (0.048) |

*U.S. data are from the ECLS and all other data are from TIMSS. All models are population weighted and heteroskedastic-consistent standard errors are in parentheses. Bold coefficents are significant at the 5 percent level or better. Countries with minimal late entry, failure, or early entry are shaded. RF stands for reduced form and reports the coefficient estimate for assigned relative age from equation (3). FS stands for first stage and reports the coefficient estimates for assigned age from equation (2). F-statistics for the first stage range from 295-35,475. IV estimates use assigned age to instrument for observed age in equation (1). All TIMSS models include controls for: sex, grade, rural residential locations, native born mother, native born father, child lives with both parents, child has a calculator, child has a computer, child has more than 100 books, a continuous measure for the number of people residing in the child's household, and a complete set of indicators for controls with missing information. The U.S. models using ECLS data includes controls for: sex, race, number of siblings, rural status, parental education, whether child lives with both parents, has over 100 books, and has a calculator.

| | Math | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| | OLS (1) | RF (2) | FS (3) | IV (4) | OLS (5) | RF (6) | FS (7) | IV (8) |
| Austria | **-0.227** | **0.099** | **0.617** | **0.160** | **-0.172** | **0.115** | **0.617** | **0.186** |
| | (0.024) | (0.040) | (0.024) | (0.066) | (0.025) | (0.039) | (0.024) | (0.065) |
| Belgium | **-0.379** | **0.097** | **0.719** | **0.134** | **-0.218** | **0.123** | **0.719** | **0.171** |
| | (0.015) | (0.031) | (0.021) | (0.045) | (0.015) | (0.030) | (0.021) | (0.043) |
| Canada | **-0.060** | **0.082** | **0.509** | **0.162** | **-0.127** | **0.097** | **0.509** | **0.190** |
| | (0.014) | (0.026) | (0.020) | (0.052) | (0.015) | (0.026) | (0.020) | (0.051) |
| Czech Republic | **-0.299** | **0.077** | **0.560** | **0.137** | **-0.209** | **0.108** | **0.560** | **0.193** |
| | (0.021) | (0.033) | (0.017) | (0.060) | (0.022) | (0.033) | (0.017) | (0.059) |
| Denmark | **-0.200** | 0.011 | **0.559** | 0.020 | **-0.158** | 0.027 | **0.559** | 0.049 |
| | (0.036) | (0.045) | (0.025) | (0.080) | (0.035) | (0.045) | (0.025) | (0.080) |
| England | **0.068** | **0.166** | **0.951** | **0.175** | **0.091** | **0.159** | **0.951** | **0.167** |
| | (0.031) | (0.034) | (0.010) | (0.036) | (0.031) | (0.034) | (0.010) | (0.036) |
| Finland | **-0.130** | 0.049 | **0.918** | 0.054 | -0.059 | 0.106 | **0.918** | 0.115 |
| | (0.045) | (0.057) | (0.019) | (0.062) | (0.047) | (0.056) | (0.019) | (0.061) |
| France | **-0.298** | **0.075** | **0.599** | 0.125 | **-0.192** | **0.093** | **0.599** | **0.155** |
| | (0.015) | (0.038) | (0.038) | (0.066) | (0.015) | (0.038) | (0.038) | (0.067) |
| Greece | **-0.109** | **0.170** | **0.795** | **0.214** | **-0.087** | **0.163** | **0.795** | **0.206** |
| | (0.016) | (0.030) | (0.020) | (0.039) | (0.017) | (0.030) | (0.020) | (0.039) |
| Iceland | -0.013 | 0.092 | **0.959** | 0.096 | 0.063 | **0.179** | **0.959** | **0.187** |
| | (0.047) | (0.052) | (0.016) | (0.055) | (0.047) | (0.054) | (0.016) | (0.056) |
| Italy | **-0.147** | **0.101** | **0.792** | **0.128** | **-0.125** | **0.117** | **0.792** | **0.148** |
| | (0.016) | (0.031) | (0.021) | (0.040) | (0.018) | (0.031) | (0.021) | (0.040) |
| Japan | **0.228** | **0.234** | **0.983** | **0.238** | **0.226** | **0.235** | **0.983** | **0.239** |
| | (0.024) | (0.024) | (0.003) | (0.025) | (0.023) | (0.024) | (0.003) | (0.024) |
| New Zealand | **-0.166** | **0.138** | **0.391** | **0.353** | **-0.161** | **0.149** | **0.391** | **0.380** |
| | (0.022) | (0.027) | (0.016) | (0.072) | (0.021) | (0.026) | (0.016) | (0.071) |
| Norway | 0.056 | **0.200** | **0.847** | **0.236** | **0.095** | **0.228** | **0.847** | **0.269** |
| | (0.036) | (0.041) | (0.014) | (0.048) | (0.036) | (0.039) | (0.014) | (0.046) |
| Portugal | **-0.204** | **0.146** | **0.588** | **0.248** | **-0.164** | **0.142** | **0.588** | **0.241** |
| | (0.009) | (0.034) | (0.043) | (0.065) | (0.010) | (0.034) | (0.043) | (0.064) |
| Slovak Republic | **-0.182** | **0.179** | **0.804** | **0.222** | **-0.163** | **0.144** | **0.804** | **0.179** |
| | (0.022) | (0.028) | (0.012) | (0.035) | (0.022) | (0.028) | (0.012) | (0.035) |
| Spain | **-0.262** | **0.159** | **0.766** | **0.207** | **-0.207** | **0.144** | **0.766** | **0.188** |
| | (0.012) | (0.033) | (0.028) | (0.045) | (0.014) | (0.032) | (0.028) | (0.044) |
| Sweden | -0.033 | **0.155** | **0.921** | **0.168** | -0.010 | **0.163** | **0.921** | **0.177** |
| | (0.028) | (0.030) | (0.009) | (0.033) | (0.027) | (0.030) | (0.009) | (0.033) |
| United States* | **-0.257** | **0.103** | **0.438** | **0.235** | **-0.187** | **0.098** | **0.438** | **0.223** |
| | (0.011) | (0.024) | (0.019) | (0.057) | (0.011) | (0.024) | (0.019) | (0.057) |

*U.S. data are from the NELS and all other data are from TIMSS. All models are population weighted and heteroskedastic-consistent standard errors are in parentheses. Bold coefficents are significant at the 5 percent level or better. Countries with minimal late entry, failure, or early entry are shaded. RF stands for reduced form and reports the coefficient estimate for assigned relative age from equation (3). FS stands for first stage and reports the coefficient estimates for assigned age from equation (2). F-statistics for the first stage range from 189-125,577. IV estimates use assigned age to instrument for observed age in equation (1). All TIMSS models include controls for: sex, grade, test year, rural residential locations, native born mother, native born father, child lives with both parents, maternal education, child has a calculator, child has a computer, child has more than 100 books, a continuous measure for the number of people residing in the child's household, and a complete set of indicators for controls with missing information. The U.S. models using NELS data includes controls for: sex, race, family size, living with both parents, urban, has over 50 books, has a calculator, and parental education.

Table V
The Impact of Relative Age on Test Scores - Pooled Country Sample

| | Math | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| | OLS (1) | RF (2) | FS (3) | IV (4) | OLS (5) | RF (6) | FS (7) | IV (8) |
| Grade 4, Additional Variables: | | | | | | | | |
| (A) None | **0.053** | **0.213** | **0.869** | **0.245** | **0.066** | **0.228** | **0.869** | **0.263** |
| | (0.010) | (0.015) | (0.005) | (0.018) | (0.010) | (0.014) | (0.005) | (0.017) |
| (B) Month of Birth | **0.036** | **0.209** | **0.859** | **0.243** | **0.046** | **0.217** | **0.859** | **0.253** |
| | (0.010) | (0.016) | (0.006) | (0.018) | (0.010) | (0.015) | (0.006) | (0.018) |
| (C) Month of Birth and Eligible School Entry Age | **-0.071** | **0.203** | **0.829** | **0.245** | **-0.062** | **0.198** | **0.829** | **0.239** |
| | (0.010) | (0.029) | (0.013) | (0.035) | (0.010) | (0.027) | (0.013) | (0.032) |
| Grade 8, Additional Variables: | | | | | | | | |
| (A) None | **-0.122** | **0.133** | **0.814** | **0.163** | **-0.089** | **0.129** | **0.814** | **0.159** |
| | (0.005) | (0.009) | (0.006) | (0.012) | (0.005) | (0.010) | (0.006) | (0.012) |
| (B) Month of Birth | **-0.135** | **0.132** | **0.836** | **0.158** | **-0.103** | **0.120** | **0.836** | **0.144** |
| | (0.005) | (0.011) | (0.006) | (0.013) | (0.005) | (0.011) | (0.006) | (0.014) |
| (C) Month of Birth and Eligible School Entry Age | **-0.207** | **0.128** | **0.867** | **0.147** | **-0.163** | **0.132** | **0.867** | **0.152** |
| | (0.006) | (0.020) | (0.008) | (0.023) | (0.006) | (0.020) | (0.008) | (0.024) |

All data are from TIMSS. All models are population weighted and heteroskedastic-consistent standard errors are in parentheses. Bold coefficents are significant at the 5 percent level or better. RF stands for reduced form and reports the coefficient estimate for assigned relative age. FS stands for first stage and reports the coefficient estimates for assigned age. F-statistics for the first stage range from 4,084-27,701. IV estimates use assigned age to instrument for observed age. All models include controls for: sex, grade (for eighth grade models only), test year, rural residential locations, native born mother, native born father, child lives with both parents, maternal education (for eighth grade models only), child has a calculator, child has a computer, child has more than 100 books, a continuous measure for the number of people residing in the child's household, and a complete set of indicators for controls with missing information.

Table VI
The Impact of Relative Age on College Preparation

| | Mean (1) | OLS (2) | RF (3) | FS (4) | IV (5) | Sample Size (6) |
|---|---|---|---|---|---|---|
| British Columbia | | | | | | |
| 75%+ in Examinable Courses* | 0.18 | **-0.0032** | **0.0016** | **0.7570** | **0.0021** | 106,917 |
| | [0.39] | (0.0004) | (0.0003) | (0.0051) | (0.0004) | |
| 75%+ Across All Exams | 0.16 | **-0.0030** | **0.0013** | **0.7570** | **0.0018** | 106,917 |
| | [0.37] | (0.0004) | (0.0003) | (0.0051) | (0.0004) | |
| United States | | | | | | |
| Took SAT or ACT | 0.60 | **-0.0114** | **0.0042** | **0.4181** | **0.0101** | 10,200 |
| | [0.49] | (0.0010) | (0.0016) | (0.0247) | (0.0040) | |
| Enrolled in a 4-year Accredited College | 0.40 | **-0.0090** | **0.0042** | **0.4325** | **0.0097** | 7,644 |
| | [0.49] | (0.0008) | (0.0017) | (0.0264) | (0.0041) | |

*75% average across all examinable courses taken. The B.C. data are restricted-use administrative data provided by the B.C. Ministry of Education and the U.S. data are from NELS. Heteroskedastic-consistent standard errors are in parentheses. Standard deviations in square parentheses. Bold coefficents are significant at the 5 percent level or better. RF stands for reduced form and reports the coefficient estimate for assigned relative age. FS stands for first stage and reports the coefficient estimates for assigned age. F-statistics for the first stage range from 269-21,880. IV estimates use assigned age to instrument for observed age in grade nine (eight) in British Columbia (the United States). All B.C. models inlcude school fixed effects, year indicators, and a female dummy variable. All U.S. models are population wieghted and include school fixed effects as well as controls for: sex, race, family size, living with both parents, owning more than 50 books, owning a calculator, and parental education.

Appendix Table I
Differences in International Educational Systems

| | School Cut-off Date | Age at School Entry | Ability Grouping in Primary School | First Grade with Formal Streaming | Percent Academic at Grade 10 | Age at End of Compulsory Education | Percent Behind as of Grade 3/4 | Percent Ahead as of Grade 3/4 | Percent Behind as of Grade 7/8 | Percent Ahead as of Grade 7/8 |
|---|---|---|---|---|---|---|---|---|---|---|
| Austria | January 1 | 6 | yes | 8 | 17 | 15 | 17.9 | 2.6 | 18.1 | 1.8 |
| Belgium-Flemish | January 1 | 6 | yes | 8 | 38 | 18 | | | 24.3 | 1.3 |
| Belgium-French | January 1 | 6 | yes | 6 | 53 | 18 | | | 24.3 | 1.3 |
| Canada | January 1 | 5-6 | yes | 11 | 100 | 16-18 | 15.0 | 2.7 | 19.5 | 1.5 |
| Czech Republic | September 1 | 6 | yes | 9 | 19 | 15 | 18.2 | 0.6 | 17.5 | 0.3 |
| Denmark | January 1 | 6 | no | 10 | 48 | 15 | | | 9.4 | 4.3 |
| England | September 1 | 5 | yes | 9 | -- | 16 | 1.1 | 0.9 | 1.1 | 1.0 |
| Finland | January 1 | 7 | no | 9 | 43 | 16 | | | 4.1 | 0.8 |
| France | January 1 | 6 | yes | 9 | 48 | 16 | | | 37.0 | 3.7 |
| Greece | April 1 | 6 | yes | 9 | 61 | 16 | 4.3 | 0.9 | 12.6 | 1.1 |
| Iceland | January 1 | 6 | yes | 10 | 37 | 16 | 0.8 | 0.5 | 0.6 | 1.0 |
| Italy | January 1 | 6 | yes | 8 | 33 | 14 | | | 11.6 | 5.0 |
| Japan | April 1 | 6 | no | 9 | 75 | 15 | 0.5 | 0.6 | 0.3 | 0.3 |
| New Zealand | May 1 | 5 | yes | 11 | 100 | 16 | 7.2 | 5.0 | 9.1 | 5.9 |
| Norway | January 1 | 6 | yes | 10 | 31 | 16 | 1.0 | 0.9 | 2.1 | 1.9 |
| Portugal | January 1 | 6 | -- | 6 | 71 | 15 | 21.7 | 0.4 | 37.8 | 0.5 |
| Slovak Republic | September 1 | 6 | yes | 9 | 46 | 16 | | | 6.1 | 0.9 |
| Spain | January 1 | 6 | yes | 8 | 65 | 16 | | | 27.5 | 0.3 |
| Sweden | January 1 | 7 | yes | 9 | 87 | 16 | | | 3.1 | 0.7 |
| United States | Appendix 2 | 5 | yes | none | 100 | 16-18 | 12.2 | 0.3 | 25.3 | 3.1 |
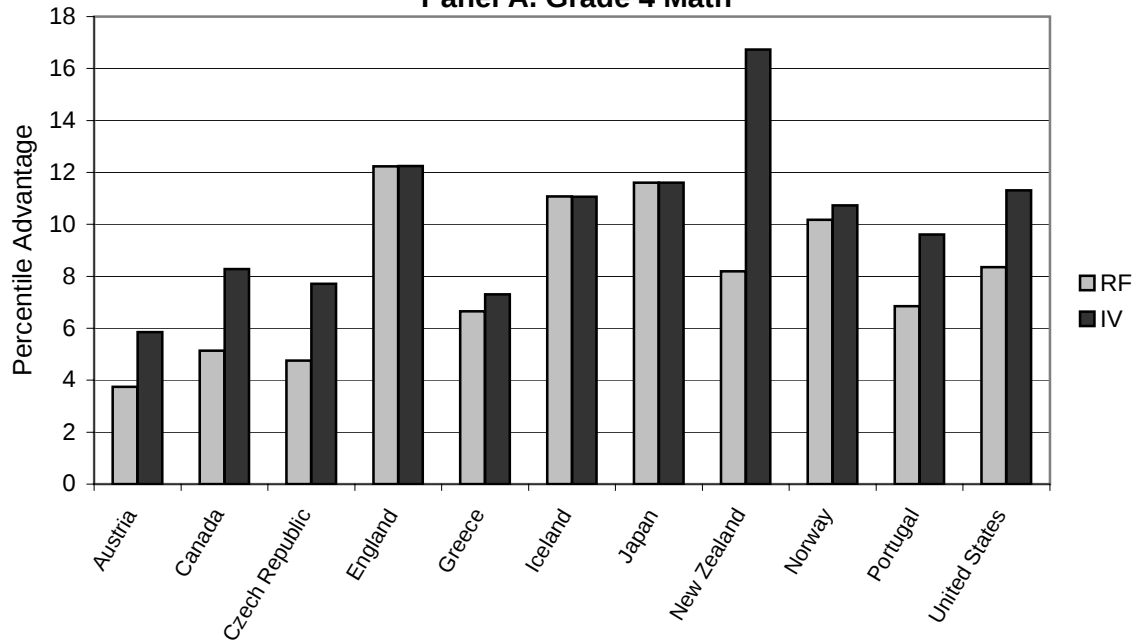
Age at school entry (age at ISCED 1 entry), first grade with formal streaming, and percent academic at grade 10 data from OECD (2004). Age at end of compulsory education is from right-to-education.org. Ability grouping data from Arnove and Zimmerman (1999), Mullis et al. (2002 and 2003), OFSTED (2003), Stevenson and Stigler (1992), and www.eurydice.org. First grade with formal streaming indicates the first grade in which explicit vocational tracks are offered. Percent academic at grade 10 is the percentage of students enrolled in an academic track. Percent behind and ahead as of grade 3/4 and 7/8 are calculated from TIMSS for all countries except the U.S. which are from the ECLS for grade 3 and NELS for grade 8.

Appendix Table II
United States School Cut-off Dates

| | School Cut-off Date | | School Cut-off Date |
|---|---|---|---|
| AL | October 1 | MO | October 1 |
| AK | November 1 | MT | September 15 |
| AZ | January 1 (≤1978) | NE | October 15 |
| | December 1 (=1979) | NV | October 1 |
| | November 1 (=1980) | NH | October 1 |
| | October 1 (=1981) | NJ | LEA |
| | September 1 (≥1982) | NM | September 1 |
| AR | October 1 | NY | December 1 |
| CA | December 1 | NC | October 15 |
| CO | LEA | ND | November 1 (≤1977) |
| CT | January 1 | | October 1 (≥1978) |
| DE | January 1 | OH | October 1 |
| FL | February 1 | OK | November 1 |
| GA | January 1 | OR | November 15 |
| HI | January 1 | PA | February 1 |
| ID | October 15 | RI | January 1 |
| IL | December 1 | SC | November 1 |
| IN | LEA | SD | November 1 (≤1978) |
| IA | October 15 (≤1978) | | September 1 (≥1979) |
| | September 15 (≥1979) | TN | November 1 |
| KS | September 1 | TX | September 1 |
| KY | January 1 (≤1978) | UT | SSY |
| | October 1 (≥1979) | VT | January 1 |
| LA | January 1 | VA | October 1 |
| ME | October 15 | WA | LEA |
| MD | unknown | WV | November 1 |
| MA | LEA | WI | December 1 (≤1976) |
| MI | December 1 | | September 1 (≥1977) |
| MN | September 1 | WY | September 15 |
| MS | January 1 (≤1976) | | |
| | December 1 (=1977) | | |
| | November 1 (=1978) | | |
| | October 1 (=1979) | | |
| | September 1 (≥1980) | | |

Cut-offs change dates are defined as the school year in which the change went into effect. LEA denotes Local Education Authority. SSY denotes start of school year. All dates are from U.S. state statutes.

**Figure I.**
**Percentile Advantage of Oldest Versus Youngest**
**Panel A. Grade 4 Math**

**Panel B. Grade 8 Math**