The Persistence of Early Maturity: International Evidence of Long-Run Age Effects

Kelly Bedard
Elizabeth Dhuey

Department of Economics
University of California, Santa Barbara

October 2004

Preliminary

Abstract

School starting age/date rules mean that there is a continuum of ages within each starting class – with the "oldest" children being approximately twenty percent older than the "youngest" children at school entry. We provide substantial evidence that these initial maturity differences have long lasting effects on student performance across OECD countries. In particular, the oldest students score 3-14 percentiles higher than the youngest students in grade four, and 3-9 percentiles higher in grade eight, depending upon the country. In fact, data from Zealand shows that the oldest children are even more likely to complete college. Taken together, these findings point to important early relative maturity effects that propagate themselves into adulthood through the structure of education systems.

1. Introduction

Almost all education systems have a specified entry date. For example, a child might be allowed to begin kindergarten as long as he has reached the age of five by September 1st of the relevant year. Entry rules 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 with an age five start date, the oldest students are approximately twenty percent older than the youngest students at school entry. Given the size of the age range on the first day of school, the oldest students are likely to be substantially more mature than the youngest students, on average. As such, one would expect an age-based performance differential during the early grades. If this relative maturity effect is significant in early grades, but 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 structure of the education system into later life, long after small differences in age should matter, they may have potentially important implications for adult outcomes and productivity.

This leads us to a second common characteristic of education systems – skill-based curriculum/programs and streaming or tracking.[1] Almost all countries sort students into different classes or groups according to ability; in some countries this takes the form of strict program based streaming (i.e. academic versus vocational) while in others it takes the more flexible form of curriculum based program placement (i.e. reading groups within classrooms). 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.

[1] There are several studies that examine the impact of streaming on student outcomes. Examples include, Oakes (1985), Arnott and Rowse (1987), Gamoran and Mare (1989), Hoffer (1992), Allen and Barnsley (1993), Argys, Rees and Brewer (1996), Betts and Shkolnik (2000A and 2000B), and Figlio and Page (2002).

These types of educational structures are importantly related to relative age because skill-based curriculum and sorting often begin during the primary or intermediate grades when relative maturity likely plays an important role in determining skill differences between younger and older students. The interaction between the commencement of skill-based curriculum and relative maturity may therefore play a central role in determining program placement or curriculum choice, and hence affect skill accumulation throughout the educational process, even after relative maturity is irrelevant in and of itself.

Allen and Barnsley (1993) provide us with a good example. They show that relative maturity plays an important role in determining minor hockey team selection for young Canadian boys, and that the oldest boys within each cohort are more likely to have successful hockey careers. 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 these teams attract the best coaches, are allocated more practice time and play against higher caliber opponents, rep team members acquire more hockey skills and increase their probability of being selected for rep teams in the future. Rep team composition therefore continues to reflect initial age differences long after relative maturity is a relevant selection factor. In fact, Allen and Barnsley (1993) 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 minor hockey is January 1).

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 certainly hint at the potential for substantial within grade differences in educational success across the relative age distribution. However, 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 entry ages determine which children are relatively older and younger, but only if the rules are strictly followed. At least in the United States, there is substantial evidence that a significant fraction of young children defer school entry by a year, making them the oldest students (Datar 2003). This is problematic since the children who delay school entry are not a random draw.[2] To distinguish between actual or 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 later measure as assigned relative age. Secondly, children who are younger at school entry are more likely to repeat a grade during elementary school (see Section 5.1). 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 or advanced and hence one would expect them to score higher on achievement tests independent of program placement.[3]

Although education researchers have studied relative age effects for many years, they have failed to come to agreement regarding the impact of relative maturity on student outcomes. While many researchers find that relatively older children academically outperform younger students,[4] others find no evidence of relative age effects.[5] In addition, some researchers find

---

[2] The difficult task of dealing with this issue is part of the reason that many previous studies have come to contradictory conclusions. For example, Breznitz and Teltsch (1989) and Bisanz, Dunn and Morrison (1995) find a positive correlation between entry ages and student performance, while DeMeis and Stearns (1992) and Morrison, Griffith and Alberts (1997) find small initial differences that quickly dissipate with age.

[3] When evaluating advanced educational outcomes school leaving rules that depend on age may also be important as they can affect educational attainment (see Angrist and Krueger 1991).

[4] Alton and Massey (1998); Cahan and Davis (1987); Cameron and Wilson (1990); Crone and Whitehurst (1999); Crosser (1991); DiPasquale, Moule, and Flewelling (1980); Gullo and Burton (1992); Jones and Mandeville (1990); Kinard and Reinherz (1986); May and Welch (1986); McClelland et al. (2000); Morrison et al. (1997); Spi, Cupp and Parke (1995); Stipek and Byler (2001); Strom (2004); Sweetland and De Simon (1987)

[5] DeMeis and Strearns (1992); Dietz and Wilson (1985); Gullo (1990); Kinard and Reinherz (1986)

long-term relative age effects,[6] whereas others do not find this relationship,[7] and hence argue that relative age is only relevant during the first few years of schooling. The lack of agreement about the existence, magnitude, and duration of relative age effects stems from two factors: (1) most studies fail to adequately address non-random selection by taking age at school entry as exogenous, and (2) small samples that are difficult to generalize to the population at large.

While relative age evaluated at any point in the educational process may be endogenous, the initial timing of births is arguably exogenous (we provide evidence for this in Section 4.1). We therefore compare the test scores of children with older and younger assigned relative ages (based on month of birth and not age at school entry) at the fourth and eighth grade levels across OECD countries using data from the Trends in 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 relative assigned 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 age, we can also estimate the within year, or causal impact of relative age using assigned relative age as an instrument for observed age. And finally, since some of the countries participating in TIMSS employ very "clean" educational systems, in the sense that essentially all children enter on time and are passed 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 impact of failure and within grade differences in relative maturity in determining the overall (net or reduced form) relative age effects.

[6] Breznitz and Telsch (1989); Cameron and Wilson (1990); Crosser (1991)

[7] Bickel et al. (1991); DiPasquale, Moule and Flewelling (1980); Jones and Mandeville (1990); Kinard and Reinherz (1986); Langer, Kalk and Searls (1984); May and Welch (1986); McClelland et al. (2000); Spi, Cupp, and Parke (1995); Stipek and Byler (2001); Sweetland and De Simon (1987)

Overall for the TIMSS sample, we find that the oldest students score substantially higher than the youngest students at both the fourth and eighth grade levels. In grade four, the oldest students score 1-4 points higher on nationally standardized tests with a mean of 50 and a standard deviation of 10. To put this in perspective, this translates into approximately a 3-14 percentile advantage for eleven months of relative age. While the age premium enjoyed by the oldest students compared to the youngest students declines between grades four and eight, there remains a 1-2.5 point difference, or approximately 3-9 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 hence are suggestive of a long-run impact.

In order to confirm the existence of a long-run relative age effect, the last section of the analysis examines the impact of relative maturity on the probability of completing an undergraduate degree in New Zealand, the only country for which we have been able to obtain college data by month of birth. We show that in New Zealand children born in the second half of the school year are approximately 1 percentage point more likely to complete college than youth born in the first half of the school year. Taken as a whole, the fourth and eighth grade results from TIMSS and the college results from New Zealand clearly point to substantial long-run relative age effects that have important implications for the distribution of skills within and across countries.

The remainder of the paper is as follows. Section 2 presents a simple theoretical model of relative maturity and its propagation by various educational structures. Section 3 discusses the data used in the analysis. Section 4 describes the econometric model. Section 5 reports the relative age estimates at the fourth and eighth grade levels. Section 6 analyzes the impact of relative age on college attendance. Section 7 concludes.

2. The Propagation of Maturity Differences

Despite the myriad of educational structures used across countries, essentially all nations have one feature in common: a single annual cut-off date.[8] In this section, we lay out several simple theoretical models in which this seemingly benign, and seemingly necessary, educational convention may result in long-run human capital and school success rate differences across children only because they are born at different times of the year.

2.1 Additive Human Capital Accumulation

For simplicity, we assume that children are born with innate ability $\theta \sim N(\mu, \sigma^2)$, and that half the population is born in one period and half in the next period, but that both groups enter school on the same date (denoted by *k* for kindergarten) and accumulate human capital additively at the same rate, *H* units per year (section 2.3 explores the multiplicative human capital case). As a result, the first-born cohort (designated by *O*) is one period older at school entry than the young cohort (designated by *Y*).[9] We further assume that the distribution of innate (at birth, $t = 0$) ability is independent of birth cohort and that as children age, their innate skills are augmented by $A(t)$, where $A'(t) > 0$ and $A''(t) < 0$. The concavity of the aging or maturity function, $A(t)$, reflects the more rapid maturing process of children at younger ages. In other words, in early grades, the youngest students are much less mature than the oldest students, but over time the difference in age becomes less important. For example, the difference between the youngest and oldest students may be large in grade one, somewhat smaller by grade three, and completely

[8] England and New Zealand have multiple kindergarten entry dates, but a single first grade entry point.
[9] For expositional ease we have simplified the model to include only two groups (old and young). However, all predictions are qualitatively similar if the model is generalized to a continuum of ages.

inconsequential by grade seven. For simplicity we assume that the aging process is a discrete annual increment with $A(t) \geq A(t+1)$.

Observed skill, *X*, at age *t* for individual *i* is therefore given by,

$$X_{Yti} = \theta_i + \sum_{j=1}^{t} A(j) + (t-k)H \quad \forall Y \quad \text{and} \quad X_{Oti} = \theta_i + \sum_{j=1}^{t} A(j+1) + (t-k)H \quad \forall O. \tag{1}$$

And, the average observable skill levels for each group are similarly given by,

$$\bar{X}_{Yt} = \mu + \sum_{j=1}^{t} A(j) + (t-k)H \quad \text{and} \quad \bar{X}_{Ot} = \mu + \sum_{j=1}^{t} A(j+1) + (t-k)H. \tag{2}$$

Due to the concavity of $A(t)$, the mean skill difference is declining with age,

$$\bar{X}_{Ot} - \bar{X}_{Yt} = A(t+1). \tag{3}$$

2.2 The Impact of Early Skill-Based Curriculum

Thus far, student skills have grown by the natural maturing process and a common annual school-based additive human capital factor. However, since many education systems sort students into skill-based curriculum groups, it is of interest to ask how the skill differential across age groups is affected by such a structure. We therefore assume that in period *s,* students are separated into two groups, and that from that point forward, student ability is augmented by H_L in each period for students in the "lower" stream and H_U in each period for students in the "upper" stream, with $H_U > H_L$. We also assume that stream placement is determined by skill level at the time of streaming (i.e. the test to determine stream placement perfectly measures *X,* but cannot disentangle θ and *A*).[10] The bottom *L* students enter the lower stream and the

[10] For the moment we are abstracting from errors in streaming other than those directly related to age. Adding a random error component would not qualitatively alter the predictions of the model.

remaining students are placed in the upper stream. Given this structure, the lower stream contains the following students:

$$F(m_Y) + F(m_O) = L \tag{4}$$

where F( ) is a cumulative normal and m_Y and m_O are the ability levels of the marginal young and old students. Since the cut-off person in each cohort must have identical observable skills,

$$m_Y = m_O + A(s+1) \tag{5}$$

where s denotes the year in which streaming occurs.

In an environment where stream placement is permanent, four factors affect the post-streaming cohort skill differential: the shape of the aging function, the size of the streams, the time at which streaming occurs, and the relative skill accumulation rate of the streams.

$$E(X_{Os+n}) - E(X_{Ys+n}) = A(s+n+1) + n(H_U - H_L)[F(m_Y) - F(m_O)] \tag{6}$$

where n is the number of years since streaming occurred. It is easiest to first consider the cohort skill differential once the maturing process is exhausted, when $A(s+n+1) = 0$. At this point, the mean skill differential is simply the excess skills accumulated by the additional old students placed in the upper stream. At all points before the aging process is complete, the skill differential also reflects the maturity advantage of the old cohort. It is also clear from equation (6) that in the absence of streaming the inter-cohort skill differential goes to zero as $A(s+n+1)$ approaches zero, and that only in an environment with skill-based program placement or differential skill accumulation can early maturity differences propagate themselves forward.

2.3 Multiplicative Human Capital

We began with additive human capital accumulation because it isolates the impact of the concave aging function and highlights the interaction between early skill-based program placement and relative maturity. However, it might seem that human capital accumulation might alternatively be described by a multiplication function: Student skill augmentation depends on the current skill base. In this case, human capital accumulation is faster for those who begin with more skills even though we continue to assume a common human capital accumulation factor, H (assuming that $H > 1$).

Observed skill, X, at age t for individual i is then given by,

$$X_{Yti} = [\theta_i + \sum_{j=1}^{k} A(j)]H^{t-k} + \sum_{j=k+1}^{t} A(j)H^{t-j} \quad \forall Y$$

and (7)

$$X_{Oti} = [\theta_i + \sum_{j=1}^{k+1} A(j)]H^{t-k} + \sum_{j=k+1}^{t} A(j+1)H^{t-j} \quad \forall O$$

And, the average observable skill differential is given by,

$$\bar{X}_{Ot} - \bar{X}_{Yt} = A(k+1)H^{t} + \sum_{j=k+1}^{t} H^{t-j}[A(j+1) - A(j)]. \tag{8}$$

It is important to note that in this case the concavity of $A(t)$ does not imply that the skill difference is declining with age. For relatively large H's, the age differential is actually monotonically increasing with age due to the initial advantage held by the old. In contrast, for relatively small H's the skill differential initially declines if the aging process outweighs the differential human capital accumulation. However, even in this case, as the aging process draws to a close, $A(t) \cong 0$ the initial base advantage of the old students becomes the dominant factor and the skill advantage held by the old students begins to increase monotonically.

The main point is that with multiplicative human capital accumulation, initial relative age effects persist even in the absence of skill-based program placement or streaming. This has a similar feel to social promotion: Students are promoted to the next grade whether or not they successfully master the current grade material. As such, this educational structure causes a continually progressive falling behind of initially weaker students. The persistence of relative age effects in this simple multiplicative human capital model, with no apparent skill-based programs, highlights the fact that relative age effects can persist either because students are separated into programs with different rates of human capital accumulation or because stronger students are encouraged to continue moving ahead while those students who are behind are simply allowed to lag farther. Of course, there could also be multiplicative human capital accumulation and skill-based program placement, in which case the age effects would be even more pronounced as even fewer young students would be placed in the upper stream.

2.4 Educational Structures and the Propagation of Early Relative Age Effects

The simple models presented in this section highlight the interaction between educational structures and the propagation of relative maturity. In an environment where education did not commence until the aging process had rendered relative maturity irrelevant, we would expect no difference in the success rate of individuals born in different months. However, since schooling begins in most countries between the ages of 5 and 7, we should expect achievement differences across children born at different points in the school year. The open question is whether these effects should dissipate with age. Regardless of the structure of the educational model, relative age effects are found early in the educational process. However, it is the interaction between the structure of the education system and the maturation process that determines how initial relative age effects propagate themselves through time.

Given the discussion so far, it is tempting to hypothesize that students in countries that rigidly stream children in academic and vocational programs at younger ages will have the strongest propagation of relative age effects into adolescence and adulthood. This is particularly appealing as this type of streaming is observable and hence the hypothesis seems testable. However, this type of streaming does not occur until adolescence, long after less rigid forms of streaming such as reading and math groups or enrichment programs have begun. As such, adolescent stream placement is at least partly an outcome of early relative maturity differences that impact primary level program placement and hence academic progress. Since all countries use such structures, it is difficult to determine a priori which countries we should expect to have the largest relative age effects.

In addition, there will even be relative age differences in academic performance in countries that employ social promotion as long as the educational system is structured such that there is some degree of multiplicative human capital accumulation. 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 – which is clearly not the case.

3. Data

The data used in this study come from five sources. Our primary sources are the 1995 and 1999 Trends in Mathematics and Science Study (TIMSS). We also use the U.S. Natality data for 1997-1999 to examine potential birth date targeting. Finally, we use data from the United Kingdom and New Zealand to estimate the long-term impact of relative age on college entry and completion. To avoid confusion, this section focuses on our primary data sources. The supplementary data sets are described in the section in which they first appear.

3.1. The Trends in Mathematics and Science Study (*TIMSS*)

Our primary data sources are the 1995 and 1999 Trends in Mathematics and Science Study (TIMSS). These data are an excellent source for studying the impact of relative age on test performance as they include 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. In this study, we restrict the sample to OECD countries with nationwide unambiguous rules regarding the school starting age. Mexico and Luxembourg are excluded because they did not participate in TIMSS in either year. Australia, Germany, Hungary, Ireland, the Netherlands, Switzerland, and the United States are excluded because their rules regarding the school cut-off date either differs across regions, which are not reported in TIMSS, or the cut-off date is at the discretion of educators and/or parents. These exclusions leave us with a sample of 11 countries for third and fourth graders and 20 countries for seventh and eighth graders: for a total of 303,515 observations across all ages and countries. For all countries, students who do not report their sex, birth month, or birth year are eliminated, reducing the sample by 4,185 students. The sample sizes are reported in Table 1.

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 will 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. We will 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 will again refer to these students as eighth graders.

The TIMSS test scores used in all analyses are standardized within countries and test books to have a mean of 50 and a standard deviation of 10. This is necessary because each student wrote only one of eight possible exams. As we are interested in the impact of relative age within countries, this is the appropriate TIMSS score. The small deviations from 50 or a standard deviation 10 observed in Table 1 are due to the sample exclusions resulting from missing sex, birth month, or birth year information. The one exception is Denmark. As can be seen in Table 1, Denmark's test scores are incorrectly standardized. However, all results were similar when we re-standardized the scores to have a mean of 50 and standard deviation of 10.

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 1st of the relevant year, then September 1st is the cut-off date. We determined these cut-off dates by the empirical distribution of birth months in each country. The beginning of the first month of the twelve consecutive months that contained 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 individual country's Department of Education website. We then used the cut-off dates and month of birth, which is reported in TIMSS, to construct a linear measure of assigned relative age (A). In particular, A=0 for students born in the last eligible month and A=11 for students born the first eligible month. For example, if the cut-off date is January 1st, December babies are the youngest (A=0) and January babies are the oldest (A=11).[11] Actual age in months (M) is constructed using the test date and birth date, both of which are reported in months.

Unfortunately, the set of control variables that are well reported in TIMSS is limited; TIMSS includes a wide range of variables, but most are sporadically reported with many variables being completely blank for entire countries. As such, the base specifications (reported in Tables 5-8) include only the child's gender; the grade during which testing occurs; and, for the eighth grade sample, an indicator variable if the student was tested in 1999. We then check the sensitivity of the estimates to the inclusion of more control variables in Table 9, for the subset of individuals and countries where they are available. In particular, we add indicator variables for rural and suburban residential locations (urban is the excluded category), 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 a set of parental education indicators (including a category for missing parental education) for eighth graders.[12] This specification also includes a continuous measure for the number of people residing in the child's household. Due to non-reporting, the addition of these variables reduces the sample size for all countries and eliminates Japan from the grade four sample and Japan, England, and France from the grade eight sample.

4. Econometric Model

We begin with a simple cross-section model of the relationship between student performance and observed age in months (M).

$$S_{cgi} = \alpha_{cg} + \beta_{cg} M_{cgi} + X_{cgi}\gamma_{cg} + \varepsilon_{cgi} \tag{9}$$

[11] In some specifications linear assigned relative age is broken into a set of month or quarter of birth indicators.
[12] Parental education information is not reported for fourth graders.

where c denotes country, g denotes grade, i denotes individual, M denotes observed age, X is a vector of controls and ε is the usual error term. All models are estimated separately for each grade and country. The parameter of interest is β_{cg} – the causal impact of relative age. However, the causal interpretation rests on the assumption that unobservables do not confound the relative age effect, which is clearly untrue given non-random grade retention.

To illustrate, consider the mean difference for S_{cgi} in equation (9) when observed age is a single indicator for being older than the school year median age (i.e. born into the oldest half of the year for "on-time" students and all students who have been retained a grade or entered school late) and there are no other explanatory variables.

$$S_{cgi} = \alpha_{cg} + \beta_{cg} M^O_{cgi} + \varepsilon_{cgi} \tag{10}$$

In this case,

$$E[S_{cgi} \mid M^O_{cgi} = 1] - E[S_{cgi} \mid M^O_{cgi} = 0] = \beta_{cg} + E[\varepsilon_{cgi} \mid M^O_{cgi} = 1] - E[\varepsilon_{cgi} \mid M^O_{cgi} = 0]. \tag{11}$$

As equation (11) indicates, unless $E[\varepsilon_{cgi} \mid M^O_{cgi} = 1] = E[\varepsilon_{cgi} \mid M^O_{cgi} = 0]$, OLS estimates of β_{cg} will be biased. 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 (who tend to score poorly) than children who are held back and enter kindergarten a year late (who tend to be positively selected), it seems likely that $E[\varepsilon_{cgi} \mid M^O_{cgi} = 1] < E[\varepsilon_{cgi} \mid M^O_{cgi} = 0]$, which means that OLS estimates will be biased downwards. In fact, in countries with high failure rates the OLS estimates can be extremely downward biased (we will return to this issue in Section 5.1).

We propose an instrumental variables (IV) solution to this problem using birth month relative to the cut-off date, assigned relative age, as an exogenous determinant of observed age (Datar (2003) uses a similar estimation strategy for primary students in the U.S.). More specifically, we estimate the parameters of equation (9) using TSLS based on the following the first-stage equation for observed age:

$$M_{cgi} = \pi_{1cg} + \pi_{2cg} A_{cgi} + X_{cgi} \pi_{3cg} + v_{cgi} \,. \tag{12}$$

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 a vast majority of students enter school on time and are not retained in any grade this is clearly satisfied, the F-statistics for the first stage range from 63-121,801 (see Tables 5 and 7). The second condition requires that assigned relative age be uncorrelated with the unobserved determinants of test scores. In other words, the second assumption presumes that assigned relative age influences test scores only through its effect on observed age. 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, in the next section, we study birth month patterns to check for birth date targeting by different socioeconomic groups. Secondly, in Section 5 we break *A* into assigned relative birth quarters so that we can perform over-identification tests.

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 extremely important from a policy perspective.

$$S_{cgi} = \theta_{1cg} + \theta_{2cg} A_{cgi} + X_{cgi} \theta_{3cg} + u_{cgi} \tag{13}$$

In particular, θ_{2cg} measures the impact of assigned relative age net of grade repetition and late entry. Stated somewhat differently, θ_{2cg} is the overall, net, or reduced form impact of assigned age on test scores at a given grade level *g*.

4. Is Assigned Relative Age Random or Do Some Parents Target "Old" Relative Ages?

Relative age might be endogenous if parents attempt to target birth dates in order to ensure that their children are the oldest in their class. To investigate this possibility, Table 2 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. And, to the extent that any quarter is slightly favored, it is either 2 or 3 (i.e spring or summer). The only exceptions are Korea and Turkey. However, there are some data issues with Korea and Turkey that suggest that all results for these countries should be interpreted with caution.[13] More interesting for our purposes are the school relative age patterns. If parents are attempting to target birth dates, we should expect a higher percentage of children to be born in the fourth quarter. However, the only country in which this is the case is Turkey.

Table 3 extends the birth pattern analysis by examining the possibility that more educated mothers may target their child's birth date more than less educated mothers. As maternal education levels vary substantially across countries, we roughly define more educated mothers as those with education levels in the top half of the education distribution (see Appendix Table 1 for a detailed breakdown for each country).[14] Table 3 reports the percentage of children born in each relative age quarter for more and less educated mothers, as well as the differences and the Z-statistics for the differences across maternal education groups. The results reported in Table 4 reveal only a few (8 out of 68) statistically significant differences across maternal education

[13] In particular, both countries have an unreasonably large percentage of births during the first and/or second months of the year. For instance, in Turkey, 12.1 percent of births are in January but only 5.7 percent in December and 7.9 percent in February. This may be a signal of targeting but more likely a signal of data errors.

[14] The fraction of mothers in the more and less educated groups is not always evenly split because maternal education is reported in six categories in TIMSS, and in some cases a single category includes a large percentage of mothers.

groups. Further, the significant differences that do exist provide no support for the hypothesis that more educated mothers target birth dates in order to ensure that there 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 of this pattern is that a small fraction of more educated mothers target summer birth dates, likely for work reasons.

One potential concern with Table 3 is that the small sample sizes in TIMSS (2,920 to 25,063) might make it difficult to pick up small differences in birth patterns across maternal education groups. We therefore supplement the evidence presented in Table 3 with similar evidence for the United States using the Natality Detail Files for 1997-1999. These files include month of birth and basic maternal socioeconomic information for all births in the U.S. in each year, covering 11,792,986 births from 1997-1999. However, as these files only report month of birth, and not day of birth, we exclude the 13 states with school cut-off dates that do not fall within ±1 day of the first of the month,[15] as it is impossible to identify children with birth dates within these months as the relatively oldest or youngest; this leaves a sample of 9,804,941 births. Further, we exclude births occurring outside of the fifty states and the District of Columbia, records with missing maternal education, maternal birthplace, or child birth order information, and records with imputed month of birth, maternal marital status, or race. These exclusions render a final sample of 9,525,687 births.

While we only have natality records for the United States, this data has two advantages. First, it covers all births in the U.S. in each year. Secondly, the birth records include detailed

[15] The excluded states are: Alaska, Arkansas, Colorado, Iowa, Massachusetts, Montana, Nebraska, New Hampshire, New Jersey, North Carolina, Pennsylvania, and Wyoming. These states have mid-month cut-off dates or the ct-off date is at the discretion of local education agencies (see Datar 2003).

maternal education information, which allows us to examine month of birth patterns across narrowly defined maternal education groups.

Columns 1-4 in Table 4 report the linear probability estimates for the impact of maternal educational attainment on the probability that a child is in the oldest relative age quarter. For all models, we report the high school dropout mean (the omitted category) and the coefficients for high school graduates, mother's with some college, and mother's with an undergraduate degree or higher. Although they are not reported, all models also include continuous controls for maternal age and birth order, as well as indicator variables for black, Hispanic, other race, immigrant mother, child birth order, survey year 1998, survey year 1999, and an intercept. The first row in Table 4 reports the overall results for all included states and the subsequent rows report the results by school start dates. The estimates reveal a consistent pattern: more educated women are approximately 1 percentage point <u>less likely</u> to give birth to children in the relatively oldest quarter. Columns 5-12 replicate the analysis for summer and winter births, respectively. These results clearly show that the oldest birth quarter results (columns 1-4) are the result of a small excess of more educated women targeting summer births and avoiding winter births. These results are entirely consistent with those reported in Table 3 for TIMSS. As one would expect the children of more educated women to be more able than average, at worst, this small amount of birth date targeting will downwardly bias the estimated relative age effects reported in Tables 5-9 since these children make up a slightly higher percentage of the assigned young.

5. The Impact of Relative Age on Test Scores

5.1. Grade 4 Results

We begin the analysis at the fourth grade level. Table 5 reports the results for all of the equations of interest for mathematics (columns 1-6) and science (columns 7-12). For comparative purposes, columns 1 and 7 report the OLS results for the impact of observed age on test scores, equation (9). Columns 2 and 8 (labeled RF) and 3 and 9 (labeled FS) report the reduced form and first stage estimates for θ_{2gc} and π_{2cg} from equations (13) and (12), respectively. While not reported in detail for every model in Tables 5 and 7, the F-statistics for the first stage relationship between observed age and assigned relative age, either linear or in quarters, range from 63-121,801. As such, assigned relative age is clearly an important determinant of observed age. The fourth and tenth columns report the IV estimates for β_{gc} using linear assigned age as the instrument and the fifth and eleventh columns similarly report the IV estimates using assigned quarters of birth as instruments. Columns 6 and 12 report the OID test statistics with two degrees of freedom for the estimates reported in columns 5 and 11. Finally, given the high level of non-reporting for control variables, the base specifications reported in Tables 5-8 include only a male indicator and grade level indicators, and an indicator for the test being taken in 1999 for the eighth grade models (the controls are expanded in Table 9).

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 “clean” countries are shaded in Tables 5 and 7). For these countries, there are no confounding factors to contend with when estimating relative age effects as 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, θ_{2gc}, and the IV estimate, β_{gc} , should be very similar. Comparing columns 2, 4, and 5 (or 8, 10, and 11) shows this to be the case. More interestingly, the point estimates for the impact of relative age are large for all countries. Referring to linear IV estimates reported column 4, one month of additional relative age increase the average math test score by 0.380, 0.308, 0.291, and 0.278 in England, Iceland, Japan, and Norway, respectively, which translate into average math test score premiums for the relatively oldest (A=11) compared to the relatively youngest (A=0) of 4.2, 3.4, 3.2, and 3.1 points on a standardized test with a mean of 50 and a standard deviation of 10. To put this in perspective, these translate into approximately 10-14 percentile test score ranking improvements for eleven months of relative age – a substantial advantage by any metric.

The remaining (not shaded) countries in Table 5, all have a sizeable minority of students either ahead or more likely behind their assigned grade at the time of testing. Further, students who are behind are more likely to have been born late in the school year (be young) and those who are ahead of their assigned grade are more likely to have been born early in the school year (be old). As such, assigned relative age also affects test scores through grade acceleration and retention/late entry.[16] Table 6 examines these relationships. Columns 1 and 5 show the percentage of children in each country and grade who have either failed a grade or entered school late, we refer to these jointly as behind. Notice the substantial variation in the fraction of students who are behind their assigned grade level across countries. At one extreme, England, Iceland, Japan, and Norway all have 1 percent or fewer students behind their assigned grade in the fourth grade sample, while at the other extreme, more than 15 percent of the fourth grade sample are behind their assigned grade in Austria, Canada, the Czech Republic and Portugal.

[16] While it is impossible to distinguish between late entry and retention in TIMSS, evidence from the Early Childhood Longitudinal Study (ECLS) in the U.S. shows that by the third grade, approximately 15 percent of children are behind and that 20 percent of this fraction is due to late entry and 80 percent is due to retention between kindergarten and grade two, at least in the United States.

Columns 3 and 7 similarly show the percentage of children who started school a year early or were accelerated at some point. As one might expect, the percentage of students ahead of their assigned grade is much smaller than the percentage behind their assigned grade. In fact, in the fourth grade sample, only New Zealand has more than 3 percent of students ahead of their assigned grade.

While there is substantial variation in the amount of retention and acceleration across countries, there is a common pattern across assigned relative ages: Children with younger assigned relative ages are more likely to be behind their assigned grade and children with older assigned relative ages are more likely to be ahead of their assigned grade. For example, in the fourth grade sample, in Austria, 39.0 and 6.8 percent of the relatively youngest and oldest children are behind, and in England 2.5 and 0 percent of the relatively oldest and youngest are behind. Similarly, 0.3 and 19.4 percent of the relatively oldest Austrian children in the fourth grade sample are ahead of their assigned grade and 0.5 and 2.8 percent of the relatively youngest and oldest English children are ahead.

The results reported in Table 6 make it clear that the reduced form estimates reported in columns 2 and 8 in Table 5 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. And 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 and hence 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 as they measure the impact of relative age at a specific grade level and hence incorporate all factors, within grade difference in age and retention. In other words, if retention is partly determined by relative age and if retention is effective at raising human capital, then the reduced form estimates of the relative age effects should be substantially smaller than the within year, IV, estimates.

In contrast, the IV estimates in Table 5 reflect only the within year impact of relative age as the age of children observed ahead or behind their assigned grade can not be predicted by assigned relative age. We report two sets of IV estimates. The coefficients reported in columns 4 and 10 use a linear measure of assigned relative age as the instrument. As this model is exactly-identified, the β_{cg}'s are simply the ratio of the coefficients reported in columns 2/3 and 8/9. The second set of IV estimates, reported in columns 5 and 11, break relative age into assigned relative quarters (with the youngest quarter being the omitted category). In this case, each model is over-identified by two degrees of freedom. For these models, we also report the over-identification test statistics and for instruments-error orthogonality in columns 6 and 12. For the fourth grade sample, the null hypothesis of orthogonality is never rejected (the critical value for rejecting orthogonality is 5.99 at the 5 percent level).

As both the math/science and linear/quarter IV estimates are very similar, we focus on the reduced form (θ_{2gc}) and linear IV mathematics results. Several features of Table 5 warrant comment. First, all of the reduced form and IV estimates are statistically significant at the conventional level. Second, in all but the "clean" countries, the IV estimates are larger than the reduced form estimates, as would be expected, although in some cases the difference is not statistically significant. Finally, and most importantly, the size of the relative age premium is large in all countries in the fourth grade sample.

In order to more easily describe the magnitude of the relative age premium, Figure 1 graphs the reduced form and linear IV estimates of the mathematics test score premium for the oldest students (A=11) compared to the youngest students (A=0). There appear to be two general groups – countries with low failure rates and countries with high failure rates. In general, the low failure rate countries, England, Greece, Iceland, Japan and Norway, tend to have relatively large relative age effects, with similar reduced form and IV estimates. For these countries, the oldest children score approximately 3-4 points higher than the youngest children, or about 10-15 percentiles higher. 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, with the IV estimates ranging from 2-3 point advantage for the oldest children and reduced form estimates ranging from 1-2.5. The one anomaly is New Zealand. The IV estimate for New Zealand is extremely large, a 5-point advantage for the oldest children, but the reduced form estimate is only 2.5 – by far the biggest differential between the two estimates. The anomaly is that the failure 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.

5.2. Grade 8 Results

As discussed in Sections 1 and 2, one of our primary objectives is to estimate the long-run, or 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, as it depends on the interaction of relative age and the structure of the education system. In an environment with no streaming or advanced work for higher achieving students, relative age effects should dissipate over time, while in the presence of such sorting structures they can persist indefinitely.

Table 7 replicates Table 5 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 beginning of secondary school. Before discussing the size of the estimates, two features of Table 7 warrant discussion. First, unlike the fourth grade sample, some of the coefficient estimates are statistically insignificant in Table 7. In particular, there is little or no evidence of relative age effects in Denmark, and the evidence is somewhat weak for the Czech Republic, Finland, and Turkey. Secondly, in four out of forty cases, the null hypothesis of instrument-error orthogonality is rejected at the 5 percent level.

Again to more easily describe the magnitude of the relative age premium, Figure 2 graphs the reduced form and linear IV estimates of the mathematics test score premium for the oldest students (A=11) compared to the youngest students (A=0), the countries with statistically imprecise estimates are marked by a star. The results are similar to those shown in Figure 1 in the sense that countries with lower failure rates tend to have larger relative age effect estimates (one exception is Portugal), the difference between the reduced form and IV estimate 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. This is most easily seen in Figure 3, where the fourth and eighth grade IV estimates are graphed side-by-side for the countries that participated in both tests. While the test score premium enjoyed by the relatively oldest students falls in all countries, although the difference is not always statistically significant, it remains economically important at the eighth grade level in all countries. In all cases the oldest students continue to score at least 1 point, or 3.5 percentiles, higher than the youngest students, with the premium being much higher than this in most countries. For example, the oldest students in the eighth grade Norwegian sample score 3 points higher and the oldest students in the New Zealand sample score 4 points, or nearly 15 percentiles, higher.

Overall, the results presented in Table 7 point to the persistence of sizeable relative age effects into adolescence across almost all countries; the estimates for Denmark are too imprecise to make even tentative statements about relative age effects in that country. This general 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 “flexible” skill-based programs model used in countries like Canada (and the United States).

5.3. The Impact of Relative Age for Boys and Girls

Relative maturity at young ages is commonly assumed to be more of an issue for boys. In Table 8, we therefore replicate the Table 5 and 7 analyses for boys and girls separately. At the fourth grade level, there are statistically significant relative age effects for both girls and boys in all countries except Austria and Canada, where the coefficients for girls are statistically insignificant. In all countries except Greece the coefficients for boys and girls are of similar magnitude. In contrast, at the eighth grade level the patterns are more complicated. In Austria, Belgium, Canada, and Turkey, relative age effects appear to be a male phenomenon, in Iceland and Italy the only statistically significant relative age effects are for girls, and in Denmark, the Czech Republic, Finland, and France there is little evidence of relative age effects at all. In all other countries, there are statistically significant relative age effects for both boys and girls, which are generally of similar magnitude. Despite the small samples in some cases, overall the results reported in Table 8 are consistent with the earlier results. The main difference is that relative age appears to be a more important issue for boys in Austria, Belgium, and Canada.

5.4. Additional Controls

Table 9 adds additional control variables to the base specification. In particular, the list of regressors is expanded to include indicator variables for rural and suburban residential locations (urban is the excluded category), 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 a set of parental education indicators (including a category for missing parental education) for the eighth grade sample, and a continuous measure for the number of people residing in the child's household for both samples. As previously stated, due to non-reporting, the addition of these variables reduces the sample size for all countries and eliminates Japan from the fourth grade sample and Japan, England and France from the eighth grade sample. For the other countries, the sample size falls between 6-40 percent (sample sizes are reported in columns 5 and 10). Despite the additional controls and the sample size reductions, in all but one case, Greece at the fourth grade level, the reduced form and linear IV relative age point estimates are very similar to those reported in Tables 5 and 7.

6. The Long-Term Impact in New Zealand

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. 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 educational attainment and birth dates. The one exception is the 1980 U.S. 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 and the G.I. Bills instituted after World War II and the Korean War.

We have, however, been able to obtain summary data containing educational attainment and month of birth for New Zealand. Statistics New Zealand has provided us with the percentage of New Zealanders born between 1967-1976 who had completed a Bachelors degree or higher as of 2001, from their 2001 Census.

Column 1 in Table 10 reports the percentage of the 1967-1976 New Zealand born population who held a Bachelor's degree or higher as of 2001. For this cohort, 14.6 percent of individuals born into the youngest quarter have a BA or higher compared to 14.9, 15.6 and 15.2 percent for the second, third, and fourth quarters, respectively. In all cases, the college advantage for quarters two through four are statistically significant at the 5 percent level. The one oddity of the results reported in column 1 is the higher percentage university degrees for third quarter persons compared fourth quarter persons. This result likely reflects the relatively high rate of grade acceleration, or early school entry, in New Zealand. As we saw in Table 6, approximately 5 percent of New Zealand students are in a grade that is ahead of their scheduled grade. This feature of the New Zealand educational system may lead to a higher rate of university participation for third quarter babies compared to fourth quarter babies as acceleration is highly concentrated in the first few months surrounding the school start date,[17] and there is

[17] Based on data from the fourth grade TIMSS sample, 37 percent of children born in the oldest month are accelerated compared to 13 percent for the second oldest, 4 percent of the third oldest, and approximately 0.5-1.5

some evidence that students who enter school early tend to perform somewhat worse than they otherwise would have been expected to. For example, if approximately 15 percent of students continue on to college, then one might roughly expect that students who enter school a year early might be expected to eventually enter this group. Following this logic, we might further expect that all, or at least most, of the accelerated group would fall in the top 15 percent of the test score distribution. However, in the New Zealand fourth grade TIMSS sample the top 15 percent of on-time students score 61 points or above on the math exam, while the average accelerated student scores only 54, placing them only at the 62nd percentile of the on-time score distribution.

7. Conclusion

The large relative age effects reported in this paper are the result of what many might consider a necessary and benign educational construct: One school start date. Despite the theoretical possibility of long-run relative age effects, a priori one might have expected that these effects exist early in the educational process, but then dissipate with age, making the single start date seem both easiest and innocuous. However, the results reported in this paper cast serious doubt on this view. In particular, we find that the oldest students score 3-14 percentiles higher than the youngest students at the fourth grade level and 3-9 percentiles higher at the eighth grade level. Further, the TIMSS results are corroborated by the fact that relatively older students in the U.K. and New Zealand are ultimately more likely to attend and complete university. Taken as a whole, we believe that the evidence presented in this paper provide substantial evidence that a wide range of educational structures, from social promotion to skill-based grouping and streaming, propagate relative age effects into adulthood.

---

percent for all months thereafter. As noted earlier, the New Zealand acceleration rate is very high compared to the other countries in the sample. The next highest oldest month of birth acceleration rates are 19 percent in Austria and 9 percent in Korea.

These results suggest that we should think very carefully about our current single school start date format. While it is likely impossible to have a continuous school starting rule – starting on your fifth birthday, for example – as this would either simply result in some children staying in kindergarten longer than other, and hence pushing off the commencement of relative age effects by one year, or would require a continuum of grade progression dates (moving from one grade to the next) which seems intractable. A more tractable partial solution is to have two school starting dates; for example, some children start in September and move from grade to grade every September, while other children begin in February and progress from grade to grade in February.[18] While this would not entirely eliminate relative age differences, it would limit them to six months instead of twelve.

[18] This would be relatively easy to implement in schools with two or more classes per grade, although class size fluctuations due to variation in birth cohorts would be somewhat more severe. However, this would be more difficult to implement in schools with a single class at each grade level.

References

Allen, Jeremiah and Roger Barnsley. 1993. "Streams and Tiers: The Interaction of Ability, Maturity, and Training in Systems with Age-Dependent Recursive Selection," *The Journal of Human Resources*, 28(3): 649-659.

Angrist, Joshua A. and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *The Quarterly Journal of Economics*, 106(4): 979-1014.

Alton, Angus and Alf Massey. 1998. "Date of birth and achievement in GCSE and GCE A-level," *Educational Research*, 40(1): 105-109.

Argys, Laura M., Daniel I. Rees and Dominic J. Brewer. 1996. "Detracking America's Schools: Equity at Zero Cost?" *Journal of Policy Analysis and Management*, 15(4): 623-45.

Arnott, Richard and John Rowse. 1987. "Peer Group Effects and Educational Attainment," *Journal of Public Economics*, 32: 287-305.

Betts, Julian R. and Jamie L. Shkolnik. 2000. "The Effects of Ability Grouping on Student Math Achievement and Resource Allocation in Secondary Schools," *Economics of Education Review*, 19(1): 1-15.

Betts, Julian R. and Jamie L. Shkolnik. 2000. "Key Difficulties in Identifying the Effects of Ability Grouping on Student Achievement", *Economics of Education Review*, 19(1): 21-26.

Bickel, Donna, Naomi Zigmond and Joseph Strayhorn. 1991. "Chronological Age at Entrance to First Grade: Effects on Elementary School Success," *Early Childhood Research Quarterly*, 6: 105-117.

Bisanz, Jeffrey, Maria Dunn and Frederick Morrison. 1995. "Effects of Age and Schooling on the Acquisition of Elementary Quantitative Skills," *Developmental Psychology*, 31(2): 221-236.

Breznitz, Zvia and Tamar Teltsch. 1989. "The Effect of School Entrance Age on Academic Achievement and Socio-Emotional Adjustment of Children: Follow-Up Study on Fourth Graders," *Psychology in Schools*, 26: 62-68.

Cahan, Sorel and Daniel 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.

Cameron, Mary Bridget and Barry J. Wilson. 1990. "The Effects of Chronological Age, Gender, and Delay of Entry on Academic Achievement and Retention: Implications for Academic Redshirting," *Psychology in the Schools*, 27: 260-263.

Crone, D. A., and Whitehurst, G. J. 1999. "Age and schooling effects on emergent literacy and early reading skills," *Journal of Educational Psychology*, 91: 604-614.

Crosser, Sandra. 1991. "Summer Birth Date Children: Kindergarten Entrance Age and Academic Achievement," *Journal of Educational Research,* 84(3): 140-146.

Datar, Ashlesha. 2003. "The Effect of Kindergarten Entrance Age on Children's Academic Acheivement," Rand Working Paper.

DeMeis, Joseph and Eleanor Stearns. 1992. "Relationship of School Entrance Age to Academic and Social Performance," *Journal of Educational Research,* 86(1): 20-27.

Dietz, Carol and Barry Wilson. 1985. "Beginning School Age and Academic Achievement," *Psychology in the Schools*, 22: 93-94.

DiPasquale, Glenn, Allan Moule, and Robert Flewelling. 1980. "The Birthdate Effect," *Journal of Learning Disabilities,* 13(5); 234-248.

Figlio David N. and Marianne E. Page. 2002. "School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality?" *Journal of Urban Economics*, 51(3): 497-514.

Gamoran, Adam and Robert D. Mare. 1989. "Secondary School Tracking and Educational Inequality: Compensation, Reinforcement, or Neutrality?" *American Journal of Sociology*, 94(5): 1146-1183.

Gullo, D.F. 1990. "The Effects of Gender, At-Risk Status, and Number of Years in Preschool on Children's Academic Readiness," *Early Education and Development,* 2(1): January.

Gullo, D.F. and C.B. Burton. 1992. "Age of Entry, Preschool Experience, and Sex as Antecedents of Academic Readiness in Kindergarten," *Early Childhood Research Quarterly,* 7:175-186.

Hoffer, T. B. 1992. "Middle school ability grouping and student achievement in science and mathematics." *Educational Evaluation and Policy Analysis*, 14(3), 205-227.

Langer, Philip, John Michael Kalk, and Donald Searls. 1984. "Age of Admission and Trends in Achievement: A Comparison of Blacks and Caucasions." 21(1): 61-78.

Jones, Molly and Garrett Mandeville. 1990. "The Effect of Age at School Entry on Reading Achievement Scores Among South Carolina Students," *Remedial and Special Education,* 11(1): 56-62.

Kinard, E.M, and H. Reinherz. 1986. "Birthdate Effects on School Performance and Adjustment: A Longitudinal Study," *Journal of Educational Research,* 79(6):366-372.

May, D.C. and E. Welch. 1986. “Screening for School Readiness: The Influence of Birthdate and Sex,” *Psychology in the Schools,* 23: 100-105.

McClelland, M., F. Morrison, and D. Holmes. 2000. “Children at risk for early academic problems: The role of learning related social skills.” *Early Childhood Research Quarterly,* 7: 307-329.

Morrison, Frederick, Elizabeth Griffith and Denise Alberts. 1997. “Nature-Nurture in the classroom: Entrance Age, School Readiness, and Learning in Children,” *Developmental Psychology*, 33(2): 254-262.

Oakes, Jeannie. 1985. *Keep Tracking: How Schools Structure Inequality*, New Haven: Yale University Press.

Spi, S., R. Cupp, and R.D. Parke. 1995. “School Entrance Age, Social Acceptance and Self-Perceptions in Kindergarten and 1st Grade,” *Early Childhood Research Quarterly,* 10: 433-450.

Stipek, Deborah and Patricia Byler. 2001. “Academic achievement and social behaviors associated with age of entry into kindergarten,” *Applied Developmental Psychology,* 22: 175-189.

Strom, Bjarne. 2004. “Student Achievement and Birthday Effects,” Working Paper, Department of Economics Norwegian University of Science and Technology.

Sweetland, John and Philip De Simone, 1987. “Age of Entry, Sex, and Academic Achievement in Elementary School Children,” *Psychology in Schools,* 24: 406-412.

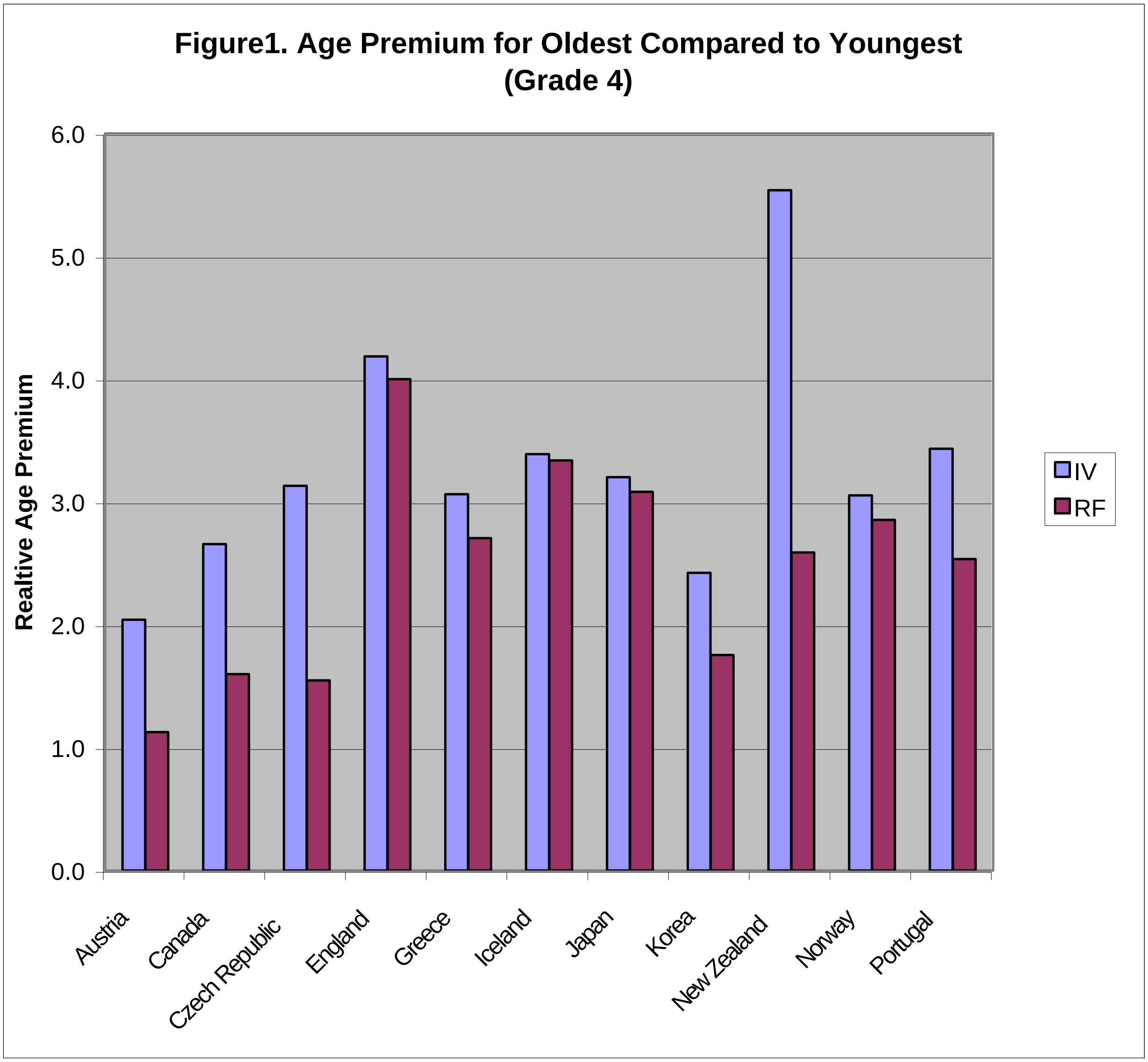
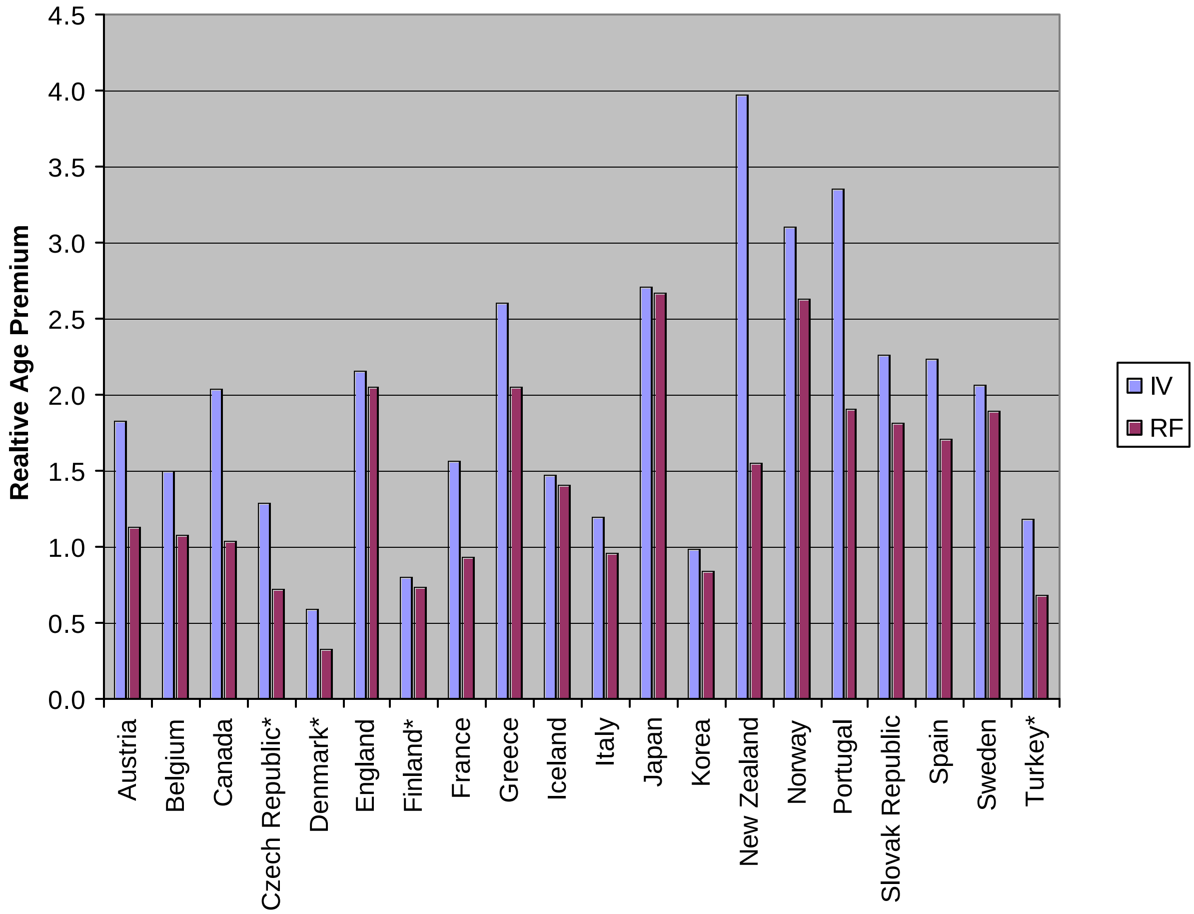


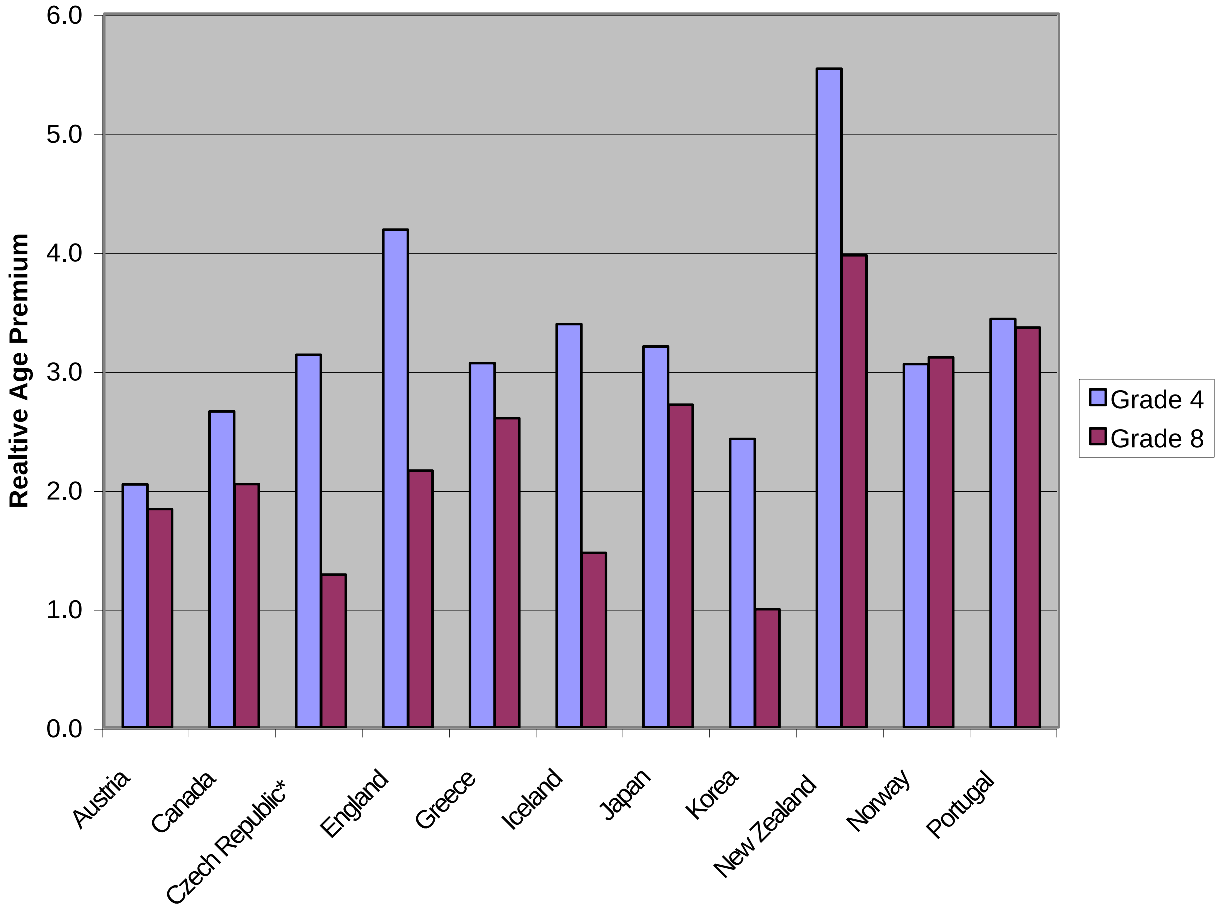
Figure1. Age Premium for Oldest Compared to Youngest
(Grade 4)
Realtive Age Premium
6.0
5.0
4.0
3.0
2.0
1.0
0.0
Austria
Canada
Czech Republic
England
Greece
Iceland
Japan
Korea
New Zealand
Norway
Portugal
IV
RF

Figure2. Age Premium for Oldest Compared to Youngest (Grade 8)



* indicates not statistically significant at the 5% level

Figure 3. Age Premium for Oldest Compared to Youngest (Linear IV Estimates)



* indicates not statistically significant at the 5% level

Table 1. TIMSS Summary Statistics

	Fourth Grade				Eighth Grade			
	Math (1)	Science (2)	% Female (3)	Sample (4)	Math (5)	Science (6)	% Female (7)	Sample (8)
Austria	50.03	50.01	0.51	5,045	50.01	50.01	0.51	5,528
	(9.98)	(9.97)	(0.50)		(9.98)	(9.98)	(0.50)	
Belgium					50.03	50.02	0.50	15,650
					(9.99)	(10.01)	(0.50)	
Canada	50.08	50.09	0.50	15,588	50.03	50.02	0.50	25,062
	(10.03)	(9.96)	(0.50)		(9.99)	(9.99)	(0.50)	
Czech Republic	50.00	50.00	0.48	6,523	50.00	50.00	0.50	10,123
	(10.00)	(10.00)	(0.50)		(10.00)	(10.00)	(0.50)	
Denmark					48.03	48.14	0.49	4,251
					(9.51)	(9.54)	(0.50)	
England	49.96	49.95	0.49	6,066	50.03	50.04	0.53	6,515
	(10.02)	(10.02)	(0.50)		(10.01)	(10.00)	(0.50)	
Finland					50.00	50.00	0.50	2,920
					(10.00)	(10.00)	(0.50)	
France					50.18	50.17	0.50	5,610
					(9.93)	(9.92)	(0.50)	
Greece	50.04	50.05	0.50	5,952	50.04	50.03	0.52	7,800
	(10.01)	(9.99)	(0.50)		(10.00)	(9.97)	(0.50)	
Iceland	50.10	50.09	0.49	3,431	50.01	50.00	0.51	3,718
	(9.97)	(9.96)	(0.50)		(10.00)	(10.00)	(0.50)	
Italy					50.00	50.00	0.49	8,163
					(10.00)	(10.00)	(0.50)	
Japan	50.06	50.06	0.50	8,536	50.02	50.02	0.51	14,889
	(9.93)	(9.90)	(0.50)		(9.99)	(9.98)	(0.50)	
Korea	50.00	50.00	0.51	5,589	50.00	50.00	0.55	11,941
	(10.00)	(10.00)	(0.50)		(10.00)	(10.00)	(0.50)	
New Zealand	50.02	50.02	0.49	4,916	50.00	50.00	0.52	10,476
	(9.99)	(9.99)	(0.50)		(10.00)	(10.00)	(0.50)	
Norway	49.97	49.98	0.52	4,300	49.99	50.00	0.51	5,644
	(9.97)	(9.98)	(0.50)		(10.01)	(10.01)	(0.50)	
Portugal	50.04	50.02	0.51	5,451	50.05	50.05	0.50	6,600
	(9.99)	(10.00)	(0.50)		(9.99)	(9.97)	(0.50)	
Slovak Republic					50.00	50.00	0.49	10,597
					(10.00)	(10.00)	(0.50)	
Spain					50.00	50.00	0.50	7,595
					(10.00)	(10.00)	(0.50)	
Sweden					50.02	50.02	0.51	8,823
					(9.99)	(9.99)	(0.50)	
Turkey					50.00	50.00	0.58	7,839
					(10.00)	(10.00)	(0.49)	

Sample means are appropriately weighted.

Table 2. Quarter of Birth Patterns

	Calendar Quarter of Birth					School Quarter of Birth					
	Q1 (1)	Q2 (2)	Q3 (3)	Q4 (4)	Quarter with most Births (5)	Q1 (6)	Q2 (7)	Q3 (8)	Q4 (9)	Quarter with most Births (10)	Sample Size (11)
Austria	0.250	0.250	0.257	0.244	3	0.262	0.247	0.251	0.241	1	10,573
Belgium	0.244	0.263	0.253	0.241	2	0.241	0.253	0.263	0.244	3	15,650
Canada	0.244	0.258	0.256	0.243	2	0.243	0.256	0.258	0.244	3	40,651
Czech Republic	0.257	0.260	0.255	0.229	2	0.254	0.266	0.242	0.238	2	16,646
Denmark	0.256	0.273	0.244	0.227	2	0.227	0.244	0.273	0.256	3	4,298
England	0.246	0.250	0.254	0.251	3	0.241	0.257	0.245	0.257	2	12,581
Finland	0.251	0.273	0.256	0.220	2	0.220	0.256	0.273	0.251	3	2,920
France	0.237	0.268	0.254	0.241	2	0.241	0.254	0.268	0.237	3	5,610
Greece	0.229	0.266	0.268	0.237	3	0.241	0.250	0.267	0.242	3	13,752
Iceland	0.240	0.261	0.261	0.238	3	0.238	0.261	0.261	0.240	2	7,149
Italy	0.248	0.268	0.246	0.237	2	0.237	0.246	0.268	0.248	3	8,163
Japan	0.239	0.247	0.271	0.244	3	0.239	0.244	0.271	0.247	3	23,425
Korea	0.288	0.220	0.241	0.251	1	0.292	0.250	0.233	0.226	1	17,530
New Zealand	0.244	0.254	0.247	0.255	4	0.256	0.246	0.257	0.241	3	15,392
Norway	0.249	0.283	0.246	0.222	2	0.222	0.246	0.283	0.249	3	10,093
Portugal	0.238	0.257	0.259	0.246	3	0.246	0.259	0.257	0.238	2	12,236
Slovak Republic	0.254	0.255	0.263	0.228	3	0.255	0.259	0.245	0.240	2	10,597
Spain	0.241	0.257	0.254	0.248	2	0.248	0.254	0.257	0.241	3	7,595
Sweden	0.265	0.268	0.252	0.215	2	0.215	0.252	0.268	0.265	3	8,823
Turkey	0.282	0.248	0.226	0.214	1	0.211	0.258	0.248	0.282	4	7,839
Average	0.250	0.259	0.253	0.236		0.241	0.253	0.259	0.246		

All proportions are appropriately weighted.

Table 3. School Quarter of Birth Patterns Across Maternal Education Groups

	Less Educated Mothers					More Educated Mothers					Difference (More-Less)				Z-Statisics			
	Q1 (1)	Q2 (2)	Q3 (3)	Q4 (4)	Sample (5)	Q1 (6)	Q2 (7)	Q3 (8)	Q4 (9)	Sample (10)	Q1 (11)	Q2 (12)	Q3 (13)	Q4 (14)	Q1 (15)	Q2 (16)	Q3 (17)	Q4 (18)
Austria	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Belgium	0.24	0.25	0.27	0.24	6,310	0.25	0.25	0.26	0.24	3,155	0.02	0.00	-0.01	-0.01	1.7	0.0	1.0	0.7
Canada	0.24	0.25	0.26	0.25	12,383	0.25	0.26	0.26	0.23	7,064	0.01	0.01	0.00	**-0.02**	1.5	0.8	0.1	2.4
Czech Republic	0.25	0.26	0.25	0.24	5,895	0.27	0.28	0.22	0.23	2,642	**0.02**	0.01	**-0.03**	-0.01	2.3	1.4	2.9	1.0
Denmark	0.21	0.26	0.29	0.24	1,190	0.22	0.22	0.29	0.27	1,141	0.01	**-0.04**	0.01	0.02	0.6	2.3	0.3	1.4
England	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Finland	0.21	0.26	0.28	0.25	719	0.21	0.27	0.25	0.27	485	0.00	0.01	-0.02	0.02	0.1	0.2	0.9	0.6
France	0.23	0.25	0.28	0.24	2,101	0.23	0.28	0.27	0.22	773	0.00	0.03	-0.01	-0.02	0.0	1.8	0.6	1.3
Greece	0.24	0.23	0.27	0.26	4,473	0.23	0.25	0.26	0.26	1,608	-0.01	0.02	-0.01	-0.01	0.6	1.7	0.6	0.5
Iceland	0.21	0.26	0.26	0.26	1,979	0.22	0.26	0.24	0.28	687	0.01	0.00	-0.02	0.02	0.3	0.0	1.1	0.8
Italy	0.23	0.26	0.26	0.25	4,215	0.24	0.24	0.28	0.24	3,189	0.01	-0.02	0.02	-0.01	0.6	1.8	1.9	0.8
Japan	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Korea	0.28	0.24	0.24	0.24	3,900	0.30	0.25	0.23	0.22	6,681	**0.02**	0.01	-0.02	-0.01	2.5	0.8	1.9	1.6
New Zealand	0.26	0.24	0.25	0.25	4,963	0.25	0.26	0.25	0.24	2,285	-0.01	0.01	0.00	-0.01	0.8	1.3	0.1	0.7
Norway	0.21	0.25	0.28	0.26	2,112	0.22	0.24	0.28	0.27	1,270	0.01	-0.02	0.00	0.01	0.5	1.2	0.1	0.8
Portugal	0.25	0.26	0.25	0.24	3,763	0.23	0.25	0.28	0.23	2,032	-0.02	-0.01	**0.03**	0.00	1.6	0.6	2.2	0.2
Slovak Republic	0.25	0.26	0.24	0.25	6,339	0.26	0.26	0.24	0.23	2,861	0.01	0.01	0.00	-0.02	1.2	0.7	0.1	1.8
Spain	0.24	0.25	0.26	0.25	4,267	0.27	0.25	0.26	0.22	1,815	0.02	0.00	0.00	**-0.03**	1.8	0.4	0.3	2.6
Sweden	0.22	0.24	0.27	0.27	3,365	0.22	0.27	0.27	0.25	1,796	0.00	**0.03**	0.00	-0.02	0.0	2.1	0.1	1.9
Turkey	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

All proportions are appropriately weighted. Bold fractions are statistically different at the 5% level or better

"More" educated is approximately in the top 30-40 percent of the national education distribution (see Appendix Table 1 detailed education definitions).

Table 4. Births by Maternal Education and School Start Date

	H.S. Drop Mean	H.S. Graduate	Some College	College Graduate	H.S. Drop Mean	H.S. Graduate	Some College	College Graduate	H.S. Drop Mean	H.S. Graduate	Some College	College Graduate	Sample Size
	Oldest Birth Quarter				Summer Birth (June-Aug)				Winter Birth (Dec-Feb)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Overall	0.2527	**-0.0015**	**-0.0038**	**-0.0087**	0.2563	**0.0055**	**0.0060**	**0.0054**	0.2489	**-0.0045**	**-0.0076**	**-0.0135**	9,525,687
		(0.0004)	(0.0005)	(0.0005)		(0.0004)	(0.0005)	(0.0005)		(0.0004)	(0.0005)	(0.0005)	
By Start Date													
January	0.2472	-0.0032	**-0.0085**	**-0.0091**	0.2531	**0.0050**	**0.0103**	**0.0058**	0.2516	**-0.0050**	**-0.0100**	**-0.0143**	438,055
		(0.0023)	(0.0026)	(0.0026)		(0.0023)	(0.0026)	(0.0027)		(0.0023)	(0.0026)	(0.0026)	
June	0.2562	**0.0085**	**0.0090**	**0.0067**	0.2562	**0.0085**	**0.0090**	**0.0067**	0.2530	**-0.0144**	**-0.0163**	**-0.0221**	250,542
		(0.0026)	(0.0030)	(0.0034)		(0.0026)	(0.0030)	(0.0034)		(0.0025)	(0.0029)	(0.0033)	
August	0.2582	0.0042	-0.0008	**-0.0069**	0.2602	0.0006	0.0028	0.0018	0.2495	**-0.0068**	**-0.0070**	**-0.0191**	221,335
		(0.0028)	(0.0032)	(0.0036)		(0.0028)	(0.0032)	(0.0036)		(0.0028)	(0.0031)	(0.0035)	
September	0.2563	-0.0003	**-0.0022**	**-0.0072**	0.2569	**0.0052**	**0.0056**	**0.0051**	0.2490	**-0.0041**	**-0.0072**	**-0.0136**	4,552,343
		(0.0006)	(0.0007)	(0.0008)		(0.0006)	(0.0007)	(0.0008)		(0.0006)	(0.0007)	(0.0008)	
October	0.2505	**-0.0026**	**-0.0044**	**-0.0098**	0.2563	**0.0077**	**0.0077**	**0.0057**	0.2500	**-0.0047**	**-0.0087**	**-0.0132**	1,390,886
		(0.0011)	(0.0013)	(0.0014)		(0.0011)	(0.0013)	(0.0014)		(0.0011)	(0.0013)	(0.0014)	
December	0.2478	**-0.0036**	**-0.0065**	**-0.0117**	0.2554	**0.0053**	**0.0053**	**0.0062**	0.2478	**-0.0036**	**-0.0065**	**-0.0117**	2,672,526
		(0.0008)	(0.0009)	(0.0010)		(0.0008)	(0.0009)	(0.0010)		(0.0008)	(0.0009)	(0.0010)	

All estimates are from linear probability models. U.S. Natality data for 1997-99. All models also include continuous controls for maternal age and birth order, and indicator variables for black, Hispanic, other race, immigrant mother, child birth order, survey year 1998, survey year 1999, and a constant.

Table 5. The Impact of Relative Age on Test Scores at the Fourth Grade Level

	Math						Science					
	OLS (1)	RF (2)	FS (3)	Linear IV (4)	Quarter IV (5)	OID (6)	OLS (7)	RF (8)	FS (9)	Linear IV (10)	Quarter IV (11)	OID (12)
Austria	**-0.326**	**0.102**	**0.552**	**0.186**	**0.181**	0.2	**-0.302**	**0.134**	**0.552**	**0.242**	**0.213**	2.9
	(0.027)	(0.051)	(0.031)	(0.096)	(0.092)		(0.027)	(0.049)	(0.031)	(0.093)	(0.091)	
Canada	-0.013	**0.145**	**0.601**	**0.242**	**0.254**	0.3	**-0.045**	**0.140**	**0.601**	**0.233**	**0.210**	0.1
	(0.021)	(0.040)	(0.026)	(0.067)	(0.067)		(0.020)	(0.040)	(0.026)	(0.067)	(0.065)	
Czech Republic	**-0.233**	**0.141**	**0.494**	**0.285**	**0.285**	0.1	**-0.161**	**0.210**	**0.494**	**0.426**	**0.427**	0.5
	(0.024)	(0.035)	(0.020)	(0.073)	(0.073)		(0.025)	(0.035)	(0.020)	(0.075)	(0.074)	
England	**0.318**	**0.364**	**0.956**	**0.380**	**0.373**	0.1	**0.308**	**0.325**	**0.956**	**0.340**	**0.337**	0.3
	(0.035)	(0.038)	(0.008)	(0.040)	(0.041)		(0.035)	(0.038)	(0.008)	(0.040)	(0.042)	
Greece	**0.088**	**0.246**	**0.883**	**0.278**	**0.299**	2.0	**0.070**	**0.294**	**0.883**	**0.333**	**0.360**	0.2
	(0.033)	(0.043)	(0.018)	(0.049)	(0.050)		(0.033)	(0.045)	(0.018)	(0.051)	(0.053)	
Iceland	**0.224**	**0.303**	**0.985**	**0.308**	**0.284**	0.3	**0.219**	**0.294**	**0.985**	**0.299**	**0.281**	1.3
	(0.047)	(0.050)	(0.008)	(0.051)	(0.054)		(0.047)	(0.051)	(0.008)	(0.052)	(0.054)	
Japan	**0.243**	**0.280**	**0.962**	**0.291**	**0.292**	1.5	**0.306**	**0.353**	**0.962**	**0.366**	**0.370**	0.2
	(0.031)	(0.031)	(0.005)	(0.032)	(0.033)		(0.032)	(0.031)	(0.005)	(0.032)	(0.033)	
Korea	0.021	**0.160**	**0.724**	**0.220**	**0.228**	0.0	**0.051**	**0.207**	**0.724**	**0.286**	**0.273**	0.7
	(0.022)	(0.037)	(0.022)	(0.052)	(0.052)		(0.022)	(0.038)	(0.022)	(0.053)	(0.053)	
New Zealand	0.042	**0.235**	**0.467**	**0.503**	**0.489**	1.3	0.032	**0.210**	**0.467**	**0.449**	**0.447**	0.4
	(0.036)	(0.043)	(0.023)	(0.095)	(0.089)		(0.038)	(0.044)	(0.023)	(0.097)	(0.091)	
Norway	**0.178**	**0.260**	**0.935**	**0.278**	**0.269**	1.2	**0.171**	**0.275**	**0.935**	**0.294**	**0.281**	0.7
	(0.040)	(0.043)	(0.011)	(0.046)	(0.047)		(0.042)	(0.044)	(0.011)	(0.047)	(0.049)	
Portugal	**-0.162**	**0.230**	**0.738**	**0.312**	**0.287**	0.3	**-0.138**	**0.225**	**0.738**	**0.304**	**0.291**	0.4
	(0.013)	(0.041)	(0.044)	(0.062)	(0.061)		(0.014)	(0.041)	(0.044)	(0.062)	(0.061)	

All models are appropriately weighted and heteroskastic-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. F-statistics for first stage range from 100-35,344.

RF stands for reduced form and reports the coefficient estimate for assigned relative age from equation (11).

FS stands for first stage and reports the coefficient estimates for assigned age from equation (10).

Table 6. Failure and Acceleration

	Fourth Grade				Eighth Grade			
	Failure		Acceleration		Failure		Acceleration	
	Mean (1)	Rel. Age (2)	Mean (3)	Rel. Age (4)	Mean (5)	Rel. Age (6)	Mean (7)	Rel. Age (8)
Austria	0.179	**-0.0247**	0.026	**0.0037**	0.181	**-0.0212**	0.018	**0.0002**
		(0.0018)		(0.0010)		(0.0017)		(0.0001)
Belgium					0.243	**-0.0144**	0.013	**0.0025**
						(0.0014)		(0.0002)
Canada	0.150	**-0.0265**	0.027	**0.0018**	0.195	**-0.0298**	0.015	**0.0031**
		(0.0014)		(0.0006)		(0.0012)		(0.0002)
Czech Republic	0.182	**-0.0379**	0.006	**0.0006**	0.166	**-0.0281**	0.013	**-0.0013**
		(0.0014)		(0.0003)		(0.0013)		(0.0002)
Denmark					0.094	**-0.0203**	0.043	**0.0090**
						(0.0012)		(0.0008)
England	0.011	**-0.0008**	0.009	**0.0015**	0.011	**-0.0014**	0.010	**0.0016**
		(0.0002)		(0.0003)		(0.0004)		(0.0003)
Finland					0.041	**-0.0040**	0.008	**0.0016**
						(0.0012)		(0.0004)
France					0.370	**-0.0151**	0.037	**0.0071**
						(0.0020)		(0.0006)
Greece	0.043	**-0.0060**	0.009	**0.0017**	0.126	**-0.0112**	0.011	**0.0019**
		(0.0009)		(0.0003)		(0.0011)		(0.0003)
Iceland	0.008	-0.0002	0.005	**0.0009**	0.006	**-0.0008**	0.010	**0.0020**
		(0.0004)		(0.0002)		(0.0004)		(0.0005)
Italy					0.116	**-0.0042**	0.050	**0.0101**
						(0.0011)		(0.0006)
Japan	0.005	**-0.0005**	0.006	**0.0011**	0.003	**-0.0003**	0.003	**0.0005**
		(0.0002)		(0.0002)		(0.0001)		(0.0001)
Korea	0.120	**-0.0152**	0.029	**0.0037**	0.035	**-0.0068**	0.018	**0.0031**
		(0.0013)		(0.0006)		(0.0005)		(0.0003)
New Zealand	0.072	**-0.0119**	0.050	**0.0079**	0.091	**-0.0185**	0.059	**0.0121**
		(0.0017)		(0.0017)		(0.0011)		(0.0007)
Norway	0.010	**-0.0020**	0.009	**0.0012**	0.021	**-0.0044**	0.019	**0.0036**
		(0.0003)		(0.0005)		(0.0004)		(0.0006)
Portugal	0.217	**-0.0146**	0.004	0.0002	0.378	**-0.0171**	0.005	**0.0010**
		(0.0018)		(0.0003)		(0.0019)		(0.0002)
Slovak Republic					0.061	**-0.0108**	0.009	0.0003
						(0.0007)		(0.0002)
Spain					0.275	**-0.0125**	0.003	0.0002
						(0.0016)		(0.0003)
Sweden					0.031	**-0.0048**	0.007	**0.0010**
						(0.0007)		(0.0003)
Turkey					0.288	**-0.0189**	0.063	**0.0107**
						(0.0017)		(0.0009)

All models are appropriately weighted and heteroskastic-consistent standard errors are in parentheses.
Bold coefficents are significant at the 5 percent level or better.

Table 7. The Impact of Relative Age on Test Scores at the Eighth Grade Level

	Math						Science					
	OLS (1)	RF (2)	FS (3)	Linear IV (4)	Quarter IV (5)	OID (6)	OLS (7)	RF (8)	FS (9)	Linear IV (10)	Quarter IV (11)	OID (12)
Austria	**-0.348**	**0.103**	**0.616**	**0.167**	**0.149**	0.4	**-0.321**	**0.116**	**0.616**	**0.189**	**0.172**	3.0
	(0.024)	(0.043)	(0.026)	(0.072)	(0.071)		(0.025)	(0.043)	(0.026)	(0.071)	(0.071)	
Belgium	**-0.471**	**0.098**	**0.719**	**0.137**	**0.118**	2.9	**-0.320**	**0.121**	**0.719**	**0.168**	**0.153**	0.2
	(0.014)	(0.035)	(0.023)	(0.051)	(0.051)		(0.016)	(0.034)	(0.023)	(0.049)	(0.048)	
Canada	**-0.125**	**0.094**	**0.507**	**0.186**	**0.141**	3.9	**-0.212**	**0.110**	**0.507**	**0.218**	**0.181**	4.0
	(0.014)	(0.028)	(0.021)	(0.056)	(0.054)		(0.016)	(0.028)	(0.021)	(0.057)	(0.056)	
Czech Republic	**-0.391**	0.066	**0.563**	0.117	**0.139**	5.3	**-0.294**	**0.096**	**0.563**	**0.171**	**0.203**	7.6
	(0.022)	(0.036)	(0.018)	(0.065)	(0.063)		(0.022)	(0.035)	(0.018)	(0.064)	(0.063)	
Denmark	**-0.243**	0.030	**0.557**	0.054	0.049	0.4	**-0.204**	0.040	**0.557**	0.071	0.082	0.9
	(0.037)	(0.045)	(0.026)	(0.081)	(0.080)		(0.037)	(0.045)	(0.026)	(0.082)	(0.081)	
England	**0.065**	**0.186**	**0.950**	**0.196**	**0.181**	1.1	**0.088**	**0.179**	**0.950**	**0.189**	**0.183**	0.1
	(0.035)	(0.036)	(0.010)	(0.039)	(0.039)		(0.034)	(0.036)	(0.010)	(0.038)	(0.039)	
Finland	**-0.218**	0.067	**0.916**	0.073	0.044	0.1	**-0.173**	**0.119**	**0.916**	**0.130**	0.106	4.4
	(0.048)	(0.061)	(0.020)	(0.066)	(0.067)		(0.051)	(0.059)	(0.020)	(0.065)	(0.066)	
France	**-0.371**	**0.085**	**0.595**	**0.142**	0.117	5.7	**-0.266**	**0.100**	**0.595**	**0.168**	**0.135**	2.7
	(0.015)	(0.040)	(0.040)	(0.072)	(0.070)		(0.015)	(0.040)	(0.040)	(0.071)	(0.070)	
Greece	**-0.208**	**0.187**	**0.790**	**0.236**	**0.249**	0.3	**-0.179**	**0.184**	**0.790**	**0.233**	**0.240**	1.7
	(0.017)	(0.033)	(0.020)	(0.043)	(0.045)		(0.018)	(0.033)	(0.020)	(0.043)	(0.045)	
Iceland	-0.021	**0.128**	**0.958**	**0.133**	**0.146**	0.0	0.055	**0.210**	**0.958**	**0.219**	**0.230**	1.1
	(0.052)	(0.057)	(0.016)	(0.060)	(0.062)		(0.052)	(0.058)	(0.016)	(0.061)	(0.063)	
Italy	**-0.249**	**0.087**	**0.799**	**0.109**	**0.088**	3.4	**-0.223**	**0.105**	**0.799**	**0.132**	**0.122**	6.7
	(0.017)	(0.034)	(0.021)	(0.043)	(0.044)		(0.019)	(0.034)	(0.021)	(0.043)	(0.044)	
Japan	**0.234**	**0.242**	**0.983**	**0.247**	**0.240**	2.7	**0.231**	**0.243**	**0.983**	**0.248**	**0.245**	1.2
	(0.024)	(0.024)	(0.003)	(0.025)	(0.025)		(0.024)	(0.024)	(0.003)	(0.025)	(0.025)	
Korea	**-0.044**	**0.077**	**0.852**	**0.090**	**0.081**	6.8	0.038	**0.152**	**0.852**	**0.179**	**0.176**	2.8
	(0.029)	(0.030)	(0.011)	(0.036)	(0.036)		(0.028)	(0.030)	(0.011)	(0.035)	(0.036)	
New Zealand	**-0.247**	**0.141**	**0.390**	**0.361**	**0.279**	2.3	**-0.254**	**0.152**	**0.390**	**0.389**	**0.334**	2.8
	(0.023)	(0.029)	(0.016)	(0.078)	(0.070)		(0.023)	(0.029)	(0.016)	(0.079)	(0.070)	
Norway	0.026	**0.239**	**0.846**	**0.283**	**0.264**	2.5	**0.048**	**0.267**	**0.846**	**0.316**	**0.318**	3.2
	(0.038)	(0.043)	(0.014)	(0.051)	(0.052)		(0.039)	(0.042)	(0.014)	(0.050)	(0.052)	
Portugal	**-0.264**	**0.174**	**0.568**	**0.305**	**0.247**	8.2	**-0.219**	**0.162**	**0.568**	**0.284**	**0.258**	3.2
	(0.009)	(0.036)	(0.044)	(0.075)	(0.071)		(0.010)	(0.162)	(0.044)	(0.072)	(0.069)	
Slovak Republic	**-0.258**	**0.165**	**0.803**	**0.206**	**0.203**	1.7	**-0.232**	**0.134**	**0.803**	**0.167**	**0.169**	0.7
	(0.023)	(0.030)	(0.012)	(0.037)	(0.038)		(0.024)	(0.030)	(0.012)	(0.037)	(0.038)	
Spain	**-0.325**	**0.155**	**0.765**	**0.203**	**0.194**	5.2	**-0.271**	**0.144**	**0.765**	**0.189**	**0.197**	1.2
	(0.012)	(0.035)	(0.028)	(0.048)	(0.049)		(0.014)	(0.034)	(0.028)	(0.047)	(0.049)	
Sweden	**-0.100**	**0.172**	**0.917**	**0.188**	**0.182**	0.1	**-0.091**	**0.184**	**0.917**	**0.201**	**0.194**	0.5
	(0.030)	(0.033)	(0.010)	(0.036)	(0.037)		(0.031)	(0.033)	(0.010)	(0.036)	(0.037)	
Turkey	**-0.145**	**0.062**	**0.574**	0.107	0.108	0.1	**-0.095**	**0.105**	**0.574**	**0.183**	**0.180**	0.6
	(0.013)	(0.038)	(0.033)	(0.067)	(0.069)		(0.013)	(0.038)	(0.033)	(0.068)	(0.069)	

All models are appropriately weighted and heteroskastic-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. F-statistics for first stage range from 63-121,801.

Table 8. Relative Age Effects by Gender

	Fourth Grade Boys				Fourth Grade Girls				Eighth Grade Boys				Eighth Grade Girls			
	Math		Science		Math		Science		Math		Science		Math		Science	
	RF	IV	RF	IV	RF	IV	RF	IV	RF	IV	RF	IV	RF	IV	RF	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Austria	**0.188**	**0.340**	**0.220**	**0.398**	0.019	0.034	0.051	0.092	**0.164**	**0.251**	**0.181**	**0.277**	0.034	0.059	0.043	0.075
	(0.058)	(0.113)	(0.061)	(0.118)	(0.083)	(0.151)	(0.075)	(0.140)	(0.058)	(0.093)	(0.059)	(0.095)	(0.064)	(0.112)	(0.061)	(0.108)
Belgium									**0.143**	**0.197**	**0.167**	**0.231**	0.053	0.074	0.074	0.103
									(0.049)	(0.071)	(0.046)	(0.066)	(0.051)	(0.073)	(0.050)	(0.072)
Canada	**0.288**	**0.471**	**0.259**	**0.424**	0.011	0.019	0.028	0.047	**0.155**	**0.317**	**0.159**	**0.324**	0.034	0.065	0.062	0.119
	(0.056)	(0.094)	(0.058)	(0.097)	(0.057)	(0.096)	(0.055)	(0.092)	(0.038)	(0.081)	(0.038)	(0.080)	(0.040)	(0.077)	(0.042)	(0.081)
Czech Republic	**0.140**	**0.236**	**0.230**	**0.389**	**0.142**	**0.366**	**0.189**	**0.486**	0.063	0.098	0.078	0.122	0.068	0.142	0.115	**0.238**
	(0.049)	(0.085)	(0.049)	(0.086)	(0.050)	(0.134)	(0.052)	(0.140)	(0.051)	(0.080)	(0.049)	(0.078)	(0.051)	(0.107)	(0.050)	(0.107)
Denmark									0.057	0.100	0.063	0.110	0.004	0.007	0.018	0.033
									(0.062)	(0.109)	(0.062)	(0.109)	(0.065)	(0.119)	(0.067)	(0.123)
England	**0.401**	**0.425**	**0.344**	**0.365**	**0.324**	**0.335**	**0.305**	**0.315**	**0.170**	**0.176**	**0.173**	**0.180**	**0.200**	**0.214**	**0.182**	**0.194**
	(0.051)	(0.054)	(0.052)	(0.055)	(0.056)	(0.058)	(0.057)	(0.059)	(0.050)	(0.052)	(0.050)	(0.052)	(0.052)	(0.056)	(0.052)	(0.056)
Finland									0.161	0.171	0.102	0.109	-0.026	-0.030	0.136	0.152
									(0.084)	(0.090)	(0.081)	(0.087)	(0.087)	(0.098)	(0.086)	(0.097)
France									0.063	0.101	0.071	0.115	0.107	0.187	**0.130**	**0.228**
									(0.056)	(0.095)	(0.055)	(0.092)	(0.058)	(0.111)	(0.058)	(0.111)
Greece	**0.350**	**0.400**	**0.415**	**0.475**	**0.138**	**0.155**	**0.170**	**0.190**	**0.128**	**0.156**	**0.142**	**0.173**	**0.242**	**0.318**	**0.224**	**0.294**
	(0.061)	(0.071)	(0.060)	(0.071)	(0.060)	(0.068)	(0.065)	(0.074)	(0.046)	(0.057)	(0.047)	(0.058)	(0.047)	(0.066)	(0.046)	(0.064)
Iceland	**0.292**	**0.300**	**0.321**	**0.330**	**0.316**	**0.317**	**0.265**	**0.266**	0.073	0.076	0.102	0.107	**0.184**	**0.191**	**0.315**	**0.328**
	(0.071)	(0.073)	(0.071)	(0.073)	(0.072)	(0.073)	(0.072)	(0.073)	(0.078)	(0.082)	(0.077)	(0.081)	(0.084)	(0.088)	(0.086)	(0.091)
Italy									0.019	0.024	0.021	0.026	**0.159**	**0.203**	**0.194**	**0.247**
									(0.047)	(0.058)	(0.046)	(0.057)	(0.050)	(0.067)	(0.050)	(0.068)
Japan	**0.263**	**0.267**	**0.297**	**0.303**	**0.297**	**0.315**	**0.404**	**0.429**	**0.221**	**0.224**	**0.198**	**0.200**	**0.264**	**0.269**	**0.288**	**0.294**
	(0.042)	(0.043)	(0.042)	(0.042)	(0.045)	(0.048)	(0.045)	(0.048)	(0.034)	(0.034)	(0.033)	(0.034)	(0.035)	(0.036)	(0.035)	(0.036)
Korea	**0.155**	**0.207**	**0.190**	**0.253**	**0.163**	**0.233**	**0.223**	**0.319**	0.038	0.045	**0.122**	**0.143**	**0.108**	**0.127**	**0.177**	**0.207**
	(0.053)	(0.071)	(0.053)	(0.072)	(0.053)	(0.076)	(0.053)	(0.078)	(0.044)	(0.051)	(0.042)	(0.049)	(0.042)	(0.050)	(0.042)	(0.050)
New Zealand	**0.212**	**0.445**	**0.169**	**0.355**	**0.260**	**0.569**	**0.253**	**0.553**	**0.169**	**0.407**	**0.154**	**0.372**	**0.115**	**0.313**	**0.150**	**0.410**
	(0.058)	(0.126)	(0.057)	(0.123)	(0.063)	(0.144)	(0.067)	(0.153)	(0.040)	(0.102)	(0.040)	(0.102)	(0.042)	(0.119)	(0.042)	(0.121)
Norway	**0.235**	**0.259**	**0.253**	**0.278**	**0.277**	**0.289**	**0.294**	**0.307**	**0.207**	**0.243**	**0.173**	**0.203**	**0.271**	**0.323**	**0.359**	**0.428**
	(0.062)	(0.069)	(0.066)	(0.073)	(0.059)	(0.061)	(0.059)	(0.062)	(0.063)	(0.074)	(0.056)	(0.066)	(0.059)	(0.072)	(0.062)	(0.075)
Portugal	**0.259**	**0.333**	**0.262**	**0.337**	**0.201**	**0.286**	**0.185**	**0.264**	**0.246**	**0.453**	**0.173**	**0.317**	0.099	**0.167**	**0.149**	**0.252**
	(0.058)	(0.086)	(0.058)	(0.085)	(0.058)	(0.090)	(0.058)	(0.090)	(0.051)	(0.118)	(0.049)	(0.106)	(0.052)	(0.095)	(0.052)	(0.098)
Slovak Republic									**0.146**	**0.173**	**0.130**	**0.154**	**0.186**	**0.244**	**0.139**	**0.182**
									(0.040)	(0.049)	(0.040)	(0.048)	(0.043)	(0.058)	(0.043)	(0.058)
Spain									**0.168**	**0.216**	**0.142**	**0.183**	**0.140**	**0.187**	**0.146**	**0.194**
									(0.049)	(0.066)	(0.048)	(0.065)	(0.049)	(0.070)	(0.049)	(0.069)
Sweden									**0.172**	**0.184**	**0.209**	**0.224**	**0.172**	**0.190**	**0.163**	**0.180**
									(0.046)	(0.050)	(0.046)	(0.050)	(0.047)	(0.052)	(0.048)	(0.053)
Turkey									0.092	0.159	**0.160**	**0.276**	0.041	0.071	0.067	0.118
									(0.059)	(0.104)	(0.055)	(0.099)	(0.050)	(0.088)	(0.052)	(0.092)

All models are appropriately weighted and heteroskastic-consistent standard errors are in parentheses. Bold coefficents are significant at the 5 percent level or better.

Table 9. The Impact of Relative Age with Expanded Controls

	Fourth Grade					Eighth Grade				
	Math		Science		Sample	Math		Science		Sample
	RF (1)	IV (2)	RF (3)	IV (4)	Size (5)	RF (6)	IV (7)	RF (8)	IV (9)	Size (10)
Austria	0.097	0.176	**0.139**	**0.252**	4,431	**0.114**	**0.182**	**0.129**	**0.206**	5,192
	(0.052)	(0.098)	(0.050)	(0.097)		(0.041)	(0.066)	(0.040)	(0.065)	
Belgium						**0.090**	**0.12401**	**0.117**	**0.161**	14,818
						(0.032)	(0.046)	(0.031)	(0.044)	
Canada	**0.121**	**0.195**	**0.122**	**0.198**	13,793	**0.078**	**0.169**	**0.068**	**0.146**	15,244
	(0.038)	(0.062)	(0.039)	(0.063)		(0.034)	(0.074)	(0.033)	(0.072)	
Czech Republic	**0.120**	**0.241**	**0.198**	**0.397**	6,112	**0.085**	**0.151**	**0.103**	**0.183**	9,551
	(0.034)	(0.070)	(0.034)	(0.071)		(0.034)	(0.060)	(0.033)	(0.060)	
Denmark						0.019	0.034	0.057	0.103	3,763
						(0.045)	(0.080)	(0.044)	(0.080)	
England	**0.314**	**0.329**	**0.263**	**0.071**	4,632					
	(0.041)	(0.043)	(0.040)	(0.042)						
Finland						0.069	0.075	**0.124**	**0.135**	2,781
						(0.058)	(0.064)	(0.057)	(0.062)	
France										
Greece	**0.167**	**0.187**	**0.232**	**0.260**	4,939	**0.177**	**0.221**	**0.181**	**0.227**	7,067
	(0.044)	(0.049)	(0.044)	(0.049)		(0.032)	(0.042)	(0.032)	(0.041)	
Iceland	**0.257**	**0.262**	**0.265**	**0.271**	2,968	0.101	**0.105**	**0.209**	**0.218**	3,496
	(0.053)	(0.054)	(0.054)	(0.055)		(0.054)	(0.056)	(0.055)	(0.058)	
Italy						**0.117**	**0.149**	**0.156**	**0.200**	4,635
						(0.041)	(0.054)	(0.041)	(0.054)	
Japan										
Korea	**0.145**	**0.199**	**0.196**	**0.269**	5,239	**0.082**	**0.100**	**0.153**	**0.184**	11,625
	(0.036)	(0.050)	(0.037)	(0.051)		(0.028)	(0.033)	(0.028)	(0.034)	
New Zealand	**0.191**	**0.398**	**0.153**	**0.319**	4,236	**0.148**	**0.371**	**0.157**	**0.394**	6,303
	(0.042)	(0.090)	(0.041)	(0.088)		(0.034)	(0.091)	(0.034)	(0.090)	
Norway	**0.254**	**0.273**	**0.272**	**0.292**	3,915	**0.186**	**0.221**	**0.218**	**0.258**	5,415
	(0.043)	(0.046)	(0.044)	(0.047)		(0.042)	(0.050)	(0.039)	(0.047)	
Portugal	**0.213**	**0.288**	**0.215**	**0.290**	4,829	**0.152**	**0.253**	**0.150**	**0.249**	6,334
	(0.040)	(0.059)	(0.040)	(0.058)		(0.035)	(0.065)	(0.035)	(0.064)	
Slovak Republic						**0.173**	**0.214**	**0.148**	**0.183**	10,005
						(0.028)	(0.036)	(0.028)	(0.036)	
Spain						**0.170**	**0.222**	**0.143**	**0.187**	7,097
						(0.034)	(0.047)	(0.033)	(0.045)	
Sweden						**0.158**	**0.172**	**0.162**	**0.175**	8,371
						(0.031)	(0.034)	(0.031)	(0.034)	
Turkey						**0.105**	**0.179**	**0.142**	**0.241**	6,759
						(0.039)	(0.069)	(0.039)	(0.069)	

All models are appropriately weighted and heteroskastic-consistent standard errors are in parentheses.
Bold coefficents are significant at the 5 percent level or better.

Table 10. The Impact of Relative Age on College Completion in New Zealand

Relative Quarter of Birth	% of Population with BA+ (1)	Z-Statistic (2)
Quarter 1 (Youngest)	0.146	
Quarter 2	0.149	3.8
Quarter 3	0.156	6.1
Quarter 4	0.152	2.1

The New Zealand data is from a special tablulation of the 2001Census provided by Statistics New Zealand for individuals born New Zealand from 1967-1976. The Z-statistics are for the null hypothesis that the relevant relevant quarter is the same as the youngest quarter.

Appendix Table 1. Percentage of Mothers in Each Education Category

	1	2	3	4	5	6	N
Austria	--	21.9	73.1	--	--	5.0	4,187
Belgium	9.3	18.7	31.3	6.7	13.6	20.4	9,465
Canada	5.5	12.7	22.8	12.0	11.9	35.2	19,447
Czech Republic	5.7	28.2	39.4	8.7	2.9	15.3	8,537
Denmark	--	17.7	33.6	--	39.3	9.4	2,323
England	--	--	--	--	--	--	--
Finland	25.5	8.4	25.2	25.2	5.3	10.5	1,204
France	12.1	28.9	32.5	--	13.7	12.8	2,874
Greece	33.3	18.2	21.6	9.8	3.5	13.5	6,081
Iceland	17.9	14.8	15.9	21.9	6.3	23.3	2,666
Italy	23.3	34.4	30.5	1.8	4.3	5.7	7,404
Japan	--	--	--	--	--	--	--
Korea	18.6	21.0	41.7	5.6	1.1	11.9	10,581
New Zealand	1.8	22.0	31.4	13.6	6.1	25.2	7,248
Norway		20.8	20.2	22.4	9.8	26.8	3,382
Portugal	66.6	16.1	7.1	2.5	1.7	6.0	5,665
Slovak Republic	7.8	15.5	49.4	7.4	3.5	16.5	9,200
Spain	58.1	12.9	8.4	5.3	3.3	12.0	6,082
Sweden		20.2	45.5		10.6	23.8	5,161
Turkey	85.4	2.8	8.4	0.8	0.3	2.3	7,225

Sample weights used. Shaded entries are included in the "more" educated category.

Education Categories:
1. Primary education only
2. Some secondary education
3. High school graduate
4. Vocational training
5. Some university
6. Undergraduate degree