Effects of schooling and age on performance in mathematics and science: A between-grade regression discontinuity design applied to Swedish TIMSS 95 data

Christina Cliffordson, University West, Trollhättan and University of Gothenburg. Göteborg, Sweden,
christina.cliffordson@hv.se

Effects of schooling and age on performance in mathematics and science: A between-grade regression discontinuity design applied to Swedish TIMSS 95 data

Christina Cliffordson

University West and University of Gothenburg

christina.cliffordson@hv.se

Abstract

One purpose of the study is to examine the relative effects of schooling and age on performance in mathematics and science by the use of a between-grade regression discontinuity approach applied to the Swedish samples of TIMSS 95. This design relies on the assumptions that there is a sharp age-based decision rule for grade assignment, and that the regression of performance on age is linear. Another main purpose is to investigate the robustness of the results against deviations from these assumptions. The issue has theoretical, methodological and practical implications, especially on the design of international comparative studies of school performance, in which it is necessary to focus on comparability with respect to either age or to the number of years in school. Even though the design has strong advantages, there are some problems as well. The greatest problem is the selection effect caused by the non-sharp decision rule for school start. According to the Swedish TIMSS 95 data the share of under-aged students is between 0.3 % and 0.7% and of over-aged about 3%. As the Swedish data comprises three successive grade cohorts it is possible to bring together students born a particular year in the analysis, and to estimate the selection effect on the within-year regression coefficients. The main pattern of results agrees with that obtained in previous research, showing that the grade effect is about twice as strong as the age effect. The results indicate that the regression of performance on birth date is linear, and that the selection effect on the within-year regression coefficient is generally relatively small.

Keywords: Age effect, between-grade regression discontinuity design, intellectual performance, schooling effect, TIMSS 95,

Introduction

An interesting issue in educational research with considerable theoretical, methodological and practical implications is the relative amount of influence of schooling and chronological age on the development of student school performance. The issue is of great importance in many areas. One example is the design of international comparative studies of school performance, and the interpretation of the results from such studies, in which it is necessary to focus on comparability with respect to either age or to the number of years in school. However, the effects of these two factors are not easily separated because experimental methods cannot be used for ethical and practical reasons. Instead the problem has been studied with correlational and quasi-experimental methods, of which the between-grade regression discontinuity design is held to be the strongest one for disentangling the schooling effect from the effect of age on the devolvement of intellectual performance.

One purpose of the present study is to examine the relative effects of schooling and age on performance in mathematics and science by the use of a between-grade regression discontinuity approach applied to the Swedish samples of TIMSS 95. However, this design relies on the assumptions that there is a sharp age-based decision rule for grade assignment, and that the regression of performance on age is linear, so another main purpose is to investigate the robustness of the results against deviations from these assumptions.

A widely held view claims that schooling can increase intellectual performance, even if reviews of research by, for example, Ceci (1991), Herrnstein and Murray (1994), and Winship and Korenman (1997) have arrived at quite different conclusions about the strength of the effect. Herrnstein and Murray (1994) concluded, on the basis of en empirical study and a review of the literature, that the effect of schooling is rather small. According to the Herrnstein and Murray (1994) empirical study, the net effect of one year of schooling is an increase of 1 IQ point. However, a reanalysis of the data by Winship and Korenman (1997) showed, after correction of some data problems and additional model specifications, an effect of 2.7 IQ per year of schooling. They also reviewed the research on effects of schooling on intelligence with a special emphasis on studies using an analysis of covariance design. The results showed that the estimates varied quite a lot over studies, from 1 to 4 IQ points per year of schooling. Ceci (1991) identified in an influential review of 200 studies, eight different types of designs in this area. He concluded that the results from all the different approaches indicate that schooling exerts an effect on intellectual development, even though the estimates vary over studies between 0.25 to 6 IQ points per year of schooling, and all the studies suffer from different kinds of methodological limitations.

In one of the approaches identified by Ceci (1991) effects of variation in length and/or track of education on intellectual performance have been studied. A classical study in this field was conducted by Härnqvist (1968). He tested a nationally representative sample of 13-year olds with a test battery including a verbal, a spatial and an inductive test. At the age of 18, the male subset of the sample took part in a military enlistment test battery of similar composition. In the analysis, the test results at age 13 were used to control for differences in entry characteristics to the different tracks of education. The study showed that students who had the most academically oriented education gained more in general intelligence compared to those with the least amount of academic education. Several similar studies have been carried out on Swedish data (e.g., Balke-Aurell, 1982; Gustafsson, 2008; Husén & Tuijnman, 1991). There also are other such studies, several of which are Scandinavian (e.g., Lund & Thrane, 1983). All these studies have shown fairly strong schooling effects on intellectual performance, of about 2.0-2.5 IQ points per year of schooling.

It also has been demonstrated (e.g., Balke-Aurell, 1982; Cliffordson & Gustafsson, 2008; Gustafsson, 2008) that it is necessary to consider the fact that education in upper secondary school is not a homogeneous activity, and that the characteristics of different educational tracks are important for effects on different aspects of intellectual performance. Balke-Aurell (1982) concluded that spatial/technical ability factors develop in accordance with verbal and technical types of education, and, to a lesser extent, with type of occupation. Gustafsson (2008) also found differences in effects on intelligence for different tracks of education. The effects were minimal for vocational schooling, and among the academic tracks (Economics, Social Sciences, Natural Sciences and Technology) the lowest effect was observed for the Social Sciences track and the strongest for the Technology track. The mean result for the academic tracks was about 2.5 IQ points per years of study. Furthermore, the results indicated that certain schooling experiences cause improvements both in general cognitive ability, and in specific abilities. The results showed that academic tracks with technical and science orientation cause at least as strong an improvement in visualization ability (*Gv*, fluent thinking with stimuli that are visual in the mind's eye; Carroll, 1993) as in general ability, while for crystallized ability (*Gc*, acculturated knowledge and developed skills; Carroll, 1993) weaker positive effects were obtained for the academic tracks and some of the vocational tracks included in the study, compared to the effects on general cognitive ability.

Cliffordson and Gustafsson (2008) applied a new method, the continuous age/continuous treatment method or the CC method, to a large data set for separating age and schooling effects. The CC method is a similar approach as the regression discontinuity design, but with fewer restrictions. The main finding was that academic schooling has a positive effect on intellectual performance with a size of the estimates, 2.7 IQ points per year of schooling, which agrees with what have been found in previous research using other designs. The strongest effect was found for the Technology track, which was in line with the results by Gustafson (2008). However, the lowest effect was found for the Economics track, which was not the case for that study. There was also a pattern of differential effects on different intellectual tests which matches the curricular emphasis of the different academic tracks. However, for the test which primarily measures fluid ability (G*f*, ability to reason and think in new situations, or decontextualized reasoning ability; Caroll, 1993) and working memory there were positive schooling effects for all the tracks, ranging from 4.2 to 4.8 IQ points per year, which supports previous findings that schooling does effect general intelligence, and not only acquisition of specific information.

The approach which by Ceci (1991) was characterized as the strongest one for disentangling the schooling effect from the effect of age on growth of performance is the between-grade regression discontinuity design (Cahan & Cohen, 1989; Cahan & Davis, 1987). In this design regression techniques are used to estimate the independent effects of age and schooling. The slope of the regression within grades on age estimates the effect of age, and the effect of schooling is represented by the discontinuity between the regression lines for two adjacent grades. That is, the differences between the oldest and the youngest students in each grade estimate the net effect of a one-year difference in age. The difference between the youngest in any given grade level and the oldest ones in the lower adjacent grade level provides the estimate of the effect of one year of schooling. Cahan and Cohen (1989) reported a study with this design, and the results showed that the effect of schooling was about twice as strong as the effect of age. Other studies have replicated the results of the Cahan and Cohen study using similar approaches (e.g., Crone & Whitehurst, 1999; Gustafsson, in press; Stelzl, Merz, Ehlers & Remer, 1995). However, others have replicated the design, but obtained generally weaker and disparate schooling effects (Luyten, 2006; Luyten, Peschar & Coe, 2008).

However, even though this design is strong, there are some problems as well. The greatest problem is the assumption that admission to school is based only on age, according to some arbitrary cut-off date. In reality the admission of some students is delayed or made earlier in time. That is, in any given grade there are students whose age should place them in a higher or lower grade, and there also are missing students who are learning in a higher or lower grade. Furthermore, delayed and accelerated students are not random with respect to age and intellectual development. Early admission is more likely to occur for bright students born just after the cut-off, whereas delay is more likely for intellectually less developed students born just before the cut-off. These selection effects cause the regression on age to be underestimated. In recognition of this problem, Cahan and Cohen (1989) excluded those students who were not normal-aged within their grade level, as well as those whose birthday fell within two months around the cut-off date, that is, the birth dates with the highest proportion of missing students. The regression obtained from these data was then extrapolated to cover the whole range of one year. Luyten (2006) and Luyten, Peschar and Coe (2008) claim, with reference to e.g. Shadish, Cook and Campbell (2002), that if the degree of those students who are not normal-aged within their grade cohort is not excessive, it is still possible to obtain reliable effect estimates. Their way to handle the problem was to just exclude those who were not normal-aged, if the percentage does not exceed 5%. Even though these actions reduce the underestimation of the age effect, it is uncertain to what degree the problem remains.

Another potential problem is the use of regression models which rely on the assumption of linearity of the within-grade regression. If the regression of performance on age is not perfectly linear, the effect of age is underestimated. Both Cahan and Cohen (1989) and Artman, Cahan and Avni-Babad (2006) assert that the deviation from linearity occurs over a span of several years. Even though the within-grade increase in mean raw test scores as a function of age over several years is curvilinear, it can be satisfactorily approximated by short linear segments of two years. However, because the linear regression model is the main tool in the between-grade regression discontinuity design, it is reasonable to investigate the linearity assumption (Baltes & Reinert, 1969).

Method

Participants

The study relies on Swedish TIMSS 95 data, which comprises sample data for 8 855 students distributed amongst three successive grade cohorts, 6^{th}, 7^{th} and 8^{th} (2 831, 4 075 and 1 949 students, respectively) and 270 schools. Age means were 12.9, 13.9 and 14.9, respectively. The share of accelerated students was 0.4, 0.7 and 0.7% and of delayed students 2.9, 3.1 and 3.0%, respectively. Case weights were applied in the analysis. See Table 1 for detailed information based on weights.

Variables

Scores from performance tests in mathematics and science constituted the dependent variables, while age and schooling were the independent variables. Age was measured in terms of month of birth - from the youngest born in December 1983 (=1) to the oldest born in January 1979

(=60). Schooling was measured in terms of grade, indicated with dummy variables.

Method of analysis

As has already been described, the greatest problem of the between-grade regression discontinuity design is the selection effect caused by the non-sharp decision rule for school start. Another potential problem is the assumption of linearity of the within-grade regressions. As the Swedish data comprises sample data for three successive grade cohorts, 6th. 7th and 8th, it is possible to bring together students born a particular year in the analysis, and to estimate the selection effect on the within-year regression coefficients. This was done through estimating several models and specifying different regression equations for accelerated, normal-aged and delayed students. The same techniques to handle the selection problem that were applied by e.g., Cahan and Cohen (1989) and Luyten (2006) were also applied on the current data, and the results from different procedures of analysis were compared.

First, the effects of grade and age on mathematics and science performance, respectively, for normal-aged students from the grade cohorts 6-7 and 7-8, respectively, were estimated. Then, models were estimated which were based only on students born between February-November and March-October, respectively.

The next sequence of analyses was based on one of the three grades at a time, and hence only the within-year effects were estimated. This was done for: a) normal-aged students only; b) normal-aged and accelerated students from the subsequent grade level; c) normal-aged and accelerated students from the subsequent grade level, for whom the scores were corrected by subtracting an approximate estimate for the effect of one more year of schooling; d) normal-aged and delayed students from the lower grade level; e) normal-aged and delayed from the lower grade level, for whom scores were corrected by adding an approximate estimate for the effect of one less year of schooling. For the 6th grade cohort only the accelerated students were available, and for the 8th grade cohort only the delayed student were available. However, for the students born in 1981, normally belonging to the 7th grade cohort, it was possible to also include both the accelerated and the delayed students in one model, and to estimate the within-year effects based on both uncorrected and corrected scores.

The analyses were conducted with Mplus 4.1 (Muthén & Muthén, 2006) under the STREAMS 3.0 (Gustafsson & Stahl, 2005) modeling environment. In order to take account of clustering effects and hierarchical data, the “complex” option offered by the Mplus program was used in the current study. In the “complex” analyses, the standard errors become larger and the t-values become smaller due to loss in information caused by the clustering. The extent of the information loss due to clustering effects is a function of the intra-class correlation and the cluster size (Muthén & Muthén, 2006).

Findings and Discussion

Throughout the presentation of the findings from the analyses regarding performance in mathematics, the corresponding results from analyses based on science performance also will be presented in the text, but for want of space the estimates are not presented in tables and figures.

Descriptive statistics

Table 1 presents means and standard deviations of the mathematics test scores for the three grades. As may be expected, the means are higher and the standard deviations are lower for the accelerators, compared to the normal-aged students´ estimates. For the delayers the means are lower and for two (6th and 8th grades) out of the three grade cohorts the standard deviations are larger (Figure 1). Among the accelerated students 44, 45 and 50% were born in January and 66, 60, and 80% in January or February, respectively, for the 6th, 7th and 8th grade cohorts. Among the delayed students 22, 18 and 27% were born in December and 29, 33 and 38% in December or November, respectively. The remaining delayed students' month of birth were spread over the whole year, which, however, was not the case for the accelerators (Figure 2).

(Take in Table 1 about here)

(Take in Figure 1 and 2 about here)

Corresponding results from the analyses based on science performance showed an equal pattern, but with generally higher means and standard deviations and smaller differences between the means for the accelerated and the delayed students.

The between-grade regression discontinuity models

In order to compare the two commonly used procedures to handle the selection problem that were applied by, e.g. Cahan and Cohen (1989) and Luyten (2006), three groups of students were excluded from the analyses, one at a time. First students were excluded who were under- or over-aged, and then those whose birthday fell within two and four months, respectively, around the cut-off date (Table 2; Figure 3).

The results from the models including the 6th and 7th grade cohorts were in line with the results by, e.g. Cahan and Cohen (1989), that is the effect of schooling was about twice as strong as the effect of age, whereas the results from the models including the 7th and 8th grade cohorts showed varying results over the three datasets, but with generally weaker schooling effects. On comparison, for science performance the schooling effects were stronger, and the age effects were slightly weaker, for grades 6-7. However, for grades 7-8, the schooling effects were weaker, and the age effects were stronger compared to mathematics performance. The difference between the two procedures, applied in order to take care of the selection effect, seems to be very small.

(Take in Table 2 about here)

(Take in Figure 3 about here)

Luyten's (2006) study, which also was based on TIMSS 95 data, but on the lower grades and on data from eight countries, not including Sweden, showed quite disparate results over countries. However, generally the schooling effects on both mathematics and science were stronger than the age effects. For mathematics the schooling effects accounted for between 75 and 38% of the total effects, and for science the schooling effects were generally slightly weaker, but showed a similar pattern over the countries as for mathematics. The results by Gustafsson (in press) based on Swedish PIRLS 2001 data comprising students from grades 3 and 4, who used the Luyten approach to handle the selection problem, showed quite similar coefficients for both schooling and age as was obtained for grades 6-7 in the current study. Luyten, Peschar and Coe (2008) performed a study on PISA 2000 data on reading performance, reading engagement, and reading activities of 15 years old student in England (grades 10-11). They reported significant schooling effects only on reading performance, for which about two third of the total effect was attributed to the effect of schooling. Hence, the results from the current study harmonize with several previously reported findings based on the same approach for various measures of intellectual performance.

It has been observed that students whose admission was delayed are less intellectually developed than the other students in their age group, and that students whose admission was accelerated are more developed. This implies that the remaining students in the appropriate grade are also selected but in the opposite direction. In turn this implies that the within-grade regression coefficients on age should increase if those students who were born in January and December were excluded from the analysis. However, the results do not show this pattern, but when two additional months were excluded from the model the within-grade slopes increased slightly. These results also held true for the science performance data. However, the differences between the results from the three datasets were small and the standard errors rather large, which imply that the differences could be random effects rather than real differences. Thus, the differences in the results between the two procedures, applied in order to take care of the selection effect, seem to be very small.

The within-grade regression models

The selection effect was further investigated through estimating several models and specifying different regression equations for accelerated, normal-aged and delayed students, with and without corrections due to schooling effects. These analyses focused only on the within-grade age effect (Table 3). Generally, for both mathematics and science, the regression coefficients on age increased when the accelerated and delayed students were included in their appropriate grades. However, these estimates are biased due to the influence of one year less or more of schooling. In order to take care of the schooling effects for the included accelerated and delayed students, their scores were corrected by adding and subtracting, respectively, an approximate estimate of the schooling effect. The estimate used, 24 (for both subjects), was the computed mean for the schooling effects from the discontinuity models (see Table 2). As may be expected, the regression age effects generally decreased, but were still slightly higher than the computed effects from the model which included only the normal-aged students.

The results from the model based on the entire age cohort born in 1981, normally belonging to the 7^{th} grade cohort, which included corrected scores for accelerated and delayed students, showed a higher regression coefficient (*b*=1.23) on age compared to the results from the within-grade model comprising only normal-aged students (*b* =0.97) from the same grade cohort. However, the estimate was close to the mean of the coefficients from the discontinuity models comprising normal-aged students from grades 6-7 (*B* =1.12) and grades 7-8 (*b* =1.34), respectively, which was the technique applied by Luyten (2006) (see Table 2). The corresponding mean of the estimates from the discontinuity models based on the procedure used by Cahan and Cohen (1989), when also those whose birthday fell within two months around the cut-off date were excluded, (b=1.15 and 1.05, respectively) was, however, slightly lower.

(Take in Table 3 about here)

With regards to the science performance the pattern of the results was quite similar to those presented for mathematics performance.

Tests of linearity

The linearity assumption was investigated by also applying quadratic- and cubic-function models to data for accelerated, normal-aged and delayed students, with and without corrections due to schooling effects. Table 4 presents tests of significance of the quadratic and cubic terms for the mathematics data.

As is clear from the results, the analyses showed that a linear function provides an adequate description of the relationship between age and performance in mathematics. Similar results were obtained in the corresponding analyses based on the science performance data.

(Take in Table 4 about here)

Conclusion and Implications

The main pattern of results corresponds well with those obtained in a number of previous studies that have investigated various measures/tests of intellectual performance – both those primarily measuring fluid ability (G*f;* e.g., Artman & Cahan, 1993; Cahan & Cohen, 1989; Cahan & Noyman, 2001; Stelzl, Merz, Ehlers & Remer, 1995) as well as those primarily measuring crystallized ability (G*c*; e.g., Cahan & Cohen, 1989; Cahan & Noyman, 2001; Crone & Whitehurst, 1999; Gustafsson, in press; Stelzl, Merz, Ehlers & Remer, 1995) – in revealing that the schooling effect is about twice as strong as the effect of age, even though the effects vary between grades/age.

For both mathematics and science, the schooling effects were weaker for grades 7-8 than for grades 6-7, although the extent of these effects was greater for science than for mathematics. However, the mean difference of performance scores for the 6^{th} and the 7^{th} grade cohorts was greater than the similar difference for the 7^{th} and 8^{th} grade cohorts, and this dissimilarity was greater for science than for mathematics.

A majority of the studies applying a regression discontinuity design in order to measure the relative effects of age and schooling on the development of intellectual performance are based on lower or middle grade students (up to grades 6/7, or up to an age of about 12/13 years) and, generally, they show quite similar results. However, those studies which are based on older students seem to obtain more disparate results. Thus it would appear, based on the research conducted on this issue to date, that the results obtained on the bases of younger students/lower grades could be potentially valuable for the interpretation of the results from international comparative studies, where it is necessary to focus on comparability with respect either to age, or to the number of years in school, but with the restriction of having corresponding ages/grades.

Furthermore, the relative effects may not only vary between grade levels, and it is reasonable to presume that the magnitude of effects may also vary between educational systems. The studies referred to above have been conducted on data from Germany, Israel, the United States and Sweden, all of which can be regarded as “Western-style” schooling systems. Thus, as Artman, Cahan and Avni-Babad (2006) suggest, considering the possibility of significant variability between school systems, the issue of whether the results are generalizable to most educational systems is an empirical one that cannot be answered without additional studies.
Another main purpose was to investigate the robustness of the results against deviation from the assumptions that grade level is solely based on age according to some arbitrary cut-off point, and that the regression of performance on age is linear.

The results indicate that the selection effect caused by the presence of accelerated and delayed students on the estimated within-year regression coefficient is generally relatively small. The difference in the results from those analyses that apply the two most commonly used procedures (e.g., Cahan & Cohen, 1989; Luyten, 2006) designed to avoid bias in the estimation of age and schooling effects caused by the non-sharp decision rule for school start, seems to be negligible. However, the result does not imply that the result would also hold true if the proportion of those students who are not normal-aged within their grade cohort is excessive. In this study, the total percentage of accelerated and delayed students was relatively small, about 3.5 %, and, as Luyten (2006) making reference to e.g. Shadish, Cook and Campbell (2002) has suggested, if the percentage does not exceed 5% it is still possible to obtain reliable effect estimates. The upper limit for the percentage of non normal-aged students is, however, still an open question.

The results from the investigation of the linearity assumption showed that a linear function provides an adequate description of the relationship between age and performance in both mathematics and science. Thus, the assertion by Cahan and Cohen (1989) and Artman, Cahan and Avni-Babad (2006) that, notwithstanding the fact that the within-grade increase in mean scores as a function of age over several years is curvilinear, satisfactory approximations using short linear segments can nevertheless be made, would appear to be reasonable. The linearity assumption was also tested by Luyten (2006) who obtained a similar result. However, as Luyten has stated, it is likely that a curvilinear relationship will be established if the range of the analyses is extended to comprise more than two years.

References

Artman, L., & Cahan, S. (1993). Schooling and the development of transitive inference. *Developmental Psychology, 29* (4), 753-759.

Artman, L., Cahan, S., & Avni-Babad, D. (2006). Age, schooling and conditional reasoning. *Cognitive Development, 21*, 131-145.

Balke-Aurell, G. (1982). *Changes in ability as related to educational and occupational experience*. Göteborg: Göteborg studies in Educational Sciences 40, Acta Universitatis Gothoburgensis.

Baltes, P. B., & Reinert, G. (1969). Cohort effects in cognitive development of children as revealed by cross-sectional sequences. *Developmental Psychology, 1*(2), 169-177.

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

Cahan, S., & Davis, D. (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.

Cahan, S., & Noyman, A. (2001). The Kaufman Ability Battery for Children Mental Processing Scale: A valid measure of "pure" intelligence? *Educational and Psychological Measurement, 61*(5), 827-840.

Carroll, J. B. (1993). *Human cognitive abilities.* Cambridge: Cambridge University Press.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology, 27*(5*)*, 703-722.

Cliffordson, C., & Gustafsson, J.-E. (2008). Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling. *Intelligence, 36*, 143-152.

Crone, D. A., & Whitehurst, G. J. (1999). Age and schooling effects on emergent literacy and early reading skills. *Journal of Educational Psychology, 91*(4*)*, 604-614.

Gustafsson, J.-E. (2008). Schooling and intelligence: Effects of track of study on level and profile of cognitive abilities. In P. Kyllonen, R. Roberts, & L. Stankov (Eds), *Extending Intelligence: Enhancement and New Constructs.* New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.

Gustafsson, J.-E. (in press). Strukturell ekvationsmodellering [Structural equation modeling]. I G. Djurfeldt & M. Barmark (red*). Statistisk verktygslåda II* [Statistical tool-box II]. Lund: Studentlitteratur.

Gustafsson, J.-E., & Stahl, P.-A. (2005). *STREAMS 3.0 User's Guide*. Mölndal, Sweden: MultivariateWare.

Herrnstein, R. J., & Murray, C. (1994). *The Bell Curve. Intelligence and Class Structure in American Life*. New York: The Free Press.

Husén, T., & Tuijnman, A. (1991). The contribution of formal schooling to the increase in intellectual capital. *Educational Researcher, 20*(1), 17-25.

Härnqvist, K. (1968). Relative changes in intelligence from 13 to 18. *Scandinavian Journal of*

Psychology, 9, 50-82.

Muthén, L. K., & Muthén, B. O. (2006). *Mplus User's Guide version 4.1.* Los Angeles, CA: Muthén & Muthén.

Lund, T., & Thrane, V. C. (1983). Schooling and intelligence: A methodological and longitudinal study. *Scandinavian Journal of Psychology, 24*, 161-173.

Luyten, H. (2006). An emperical assessment of the absolute effect of schooling: regression-discontinuity applied to TIMSS-95. *Oxford Review of Education, 32*(3*)*, 397-429.

Luyten, H., Peschar, J., & Coe, R. 2008. Effects of Schooling on Reading Performance, Reading Engagement and Reading Activities of 15-year-olds in England. *American Educational Research Journal 45*(2), 319-342.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston, NY: Houghton Mifflin Company.

Stelzl, I., Merz, F., Remer, H., & Ehlers, Th. (1995). The effect of schooling on the development of fluid and cristallized intelligence: A quasi-experimental study. *Intelligence, 21*, 279–296.

Winship, C., & Korenman, S. (1997). Does staying in school make you smarter? The effect of education on IQ in The Bell Curve. In B. Devlin, S. E. Fienberg, D. P. Resnick, & K. Roeder (Eds.), *Intelligence, Genes, & Success. Scientists Respond to the Bell Curve.* New York: Springer-Verlag, Inc.

Table 1: Means and standard deviations per grade for mathematics performance.

Grade		*Year of birth*	*N*	*%*	*Mean*	*S.D.*
6						
	Accelerated	83	10	0.3	507.52	59.04
	Normal	82	2783	96.8	479.12	75.82
	Delayed	81	84	2.9	422.38	80.97
7						
	Accelerated	82	20	0.7	553.87	66.50
	Normal	81	2816	96.2	520.84	84.72
	Delayed	80	92	3.1	446.79	75.94
8						
	Accelerated	81	20	0.7	647.24	69.39
	Normal	80	2898	96.3	555.86	89.25
	Delayed	79	89	3.0	475.10	103.44

Note. Cases weighted by HOUSE WEIGHT.

Table 2: Unstandardized regression coefficients, standard errors, and t-values from analyses based on normal-aged students, for mathematics.

Models			*Age*				*Schooling*		
Grades	Months of birth	*N*	*B*	*S.E.*	*t-value*	*One year of age*	*b*	*S.E.*	*t-value*
	January- December	5599	1.12	0.32	3.56	13.44	28.43	5.28	5.38
6 -7	February - November	4748	1.15	0.40	2.86	13.81	29.07	6.43	4.53
	March - October	3881	1.43	0.53	2.69	17.10	26.89	7.42	3.63
	January- December	5714	1.34	0.39	3.48	16.07	18.67	6.99	2.67
7 - 8	February - November	4846	1.05	0.52	2.03	12.60	21.94	7.99	2.75
	March - October	3979	1.41	0.67	2.10	16.90	18.07	9.73	1.86

Note. Cases weighted by HOUSE WEIGHT; The lowest grade included in the model is the reference; The cut-off point is the first of January.

Table 3: Unstandardized regression coefficients, standard errors and t-values from analyses based on normal-aged, and normal-aged and accelerated and delayed students, respectively, for mathematics.

Models				*Age*		
Grades	Months of birth	*N*	*b*	*S.E.*	*t-value*	*One year of age*
6	Normal-aged	2783	1.28	0.47	2.74	15.32
6 (7)	Normal-aged and accelerated gr. 7	2803	1.40	0.46	3.02	16.75
6 (7)*	Normal-aged and accelerated gr. 7		1.34	0.46	2.90	16.10
7	Normal-aged	2816	0.97	0.42	2.28	11.60
7 (8)	Normal-aged and accelerated gr.8	2836	1.32	0.45	2.95	15.78
7 (8)*	Normal-aged and accelerated gr.8		1.25	0.44	2.86	15.01
7 (6)	Normal-aged and delayed gr. 6	2900	1.02	0.43	2.40	12.24
7 (6)*	Normal-aged and delayed gr. 6		0.94	0.42	2.23	11.32
(6) 7 (8)	Normal-aged, accelerated gr.8 and delayed gr. 6	2920	1.37	0.45	3.05	16.40
(6) 7 (8)*	Normal-aged and corrected gr.8 and delayed gr. 6		1.23	0.44	2.81	14.72
8	Normal-aged	2898	1.71	0.66	2.59	20.53
8	Normal-aged and delayed gr. 7	2990	1.94	0.63	3.10	23.26
8 (7)*	Normal-aged and delayed gr. 7		1.84	0.63	2.93	22.02

*Note. Cases weighted by HOUSE WEIGHT; *Accelerated and delayed student scores are corrected for the schooling effect with an approximate estimate of 24.*

Table 4: Estimates from quadratic- and cubic-function models, for mathematics.

Models			*Quadratic **2*		*Cubic **3*	
Grades	Month of birth	*N*	*t-value*	*p-value*	*t-value*	*p-value*
6	Normal-aged	2783	0.619	0.536	0.567	0.571
6 (7)	Normal-aged and accelerated gr. 7	2803	0.665	0.506	0.615	0.539
6 (7)*	Normal-aged and accelerated gr. 7		0.587	0.557	0.217	0.828
7	Normal-aged	2816	-0.834	0.404	0.411	0.681
(6) 7 (8)	Normal-aged, accelerated gr.8 and delayed gr. 6	2920	-0.698	0.485	0.307	0.759
(6) 7 (8)*	Normal-aged and corrected gr.8 and delayed gr. 6		-0.673	0.501	0.235	0.815
8	Normal-aged	2898	-1.078	0.281	0.051	0.960
8	Normal-aged and delayed gr. 7	2990	-1.586	0.113	0.142	0.887
8 (7)*	Normal-aged and delayed gr. 7		-1.430	0.153	0.076	0.940

Note. Cases weighted by HOUSE WEIGHT.

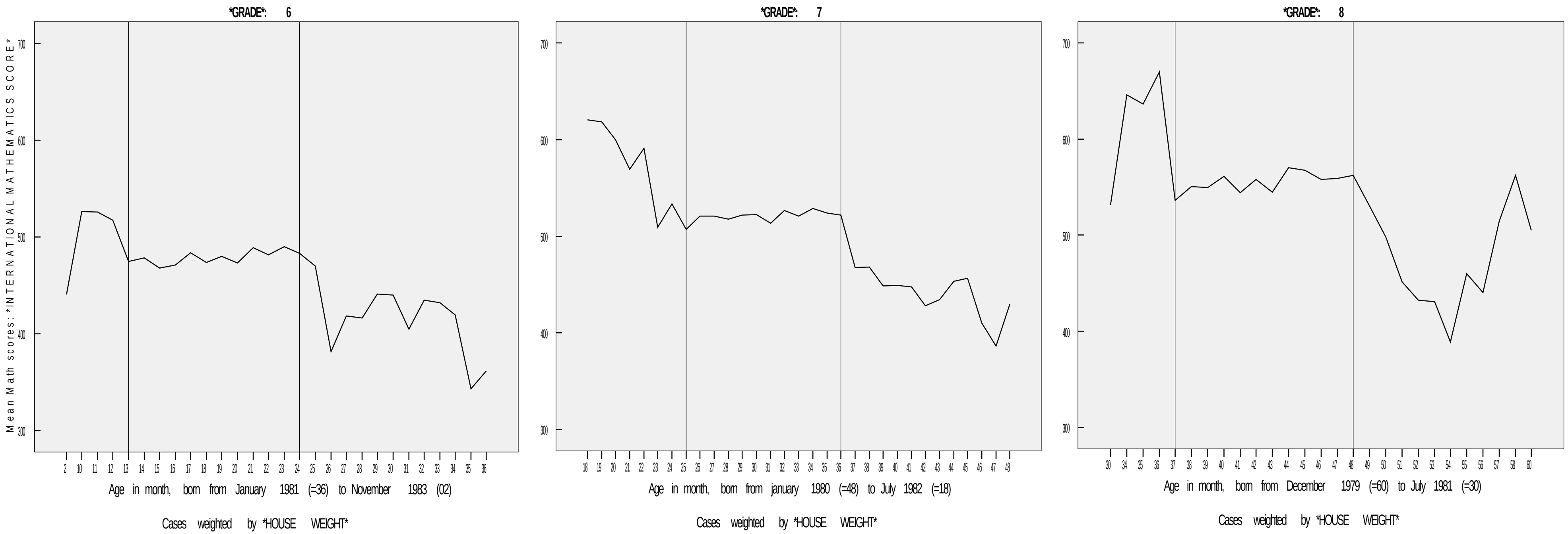

Figure 1: Mean scores for accelerated, normal-aged and delayed students, for mathematics.

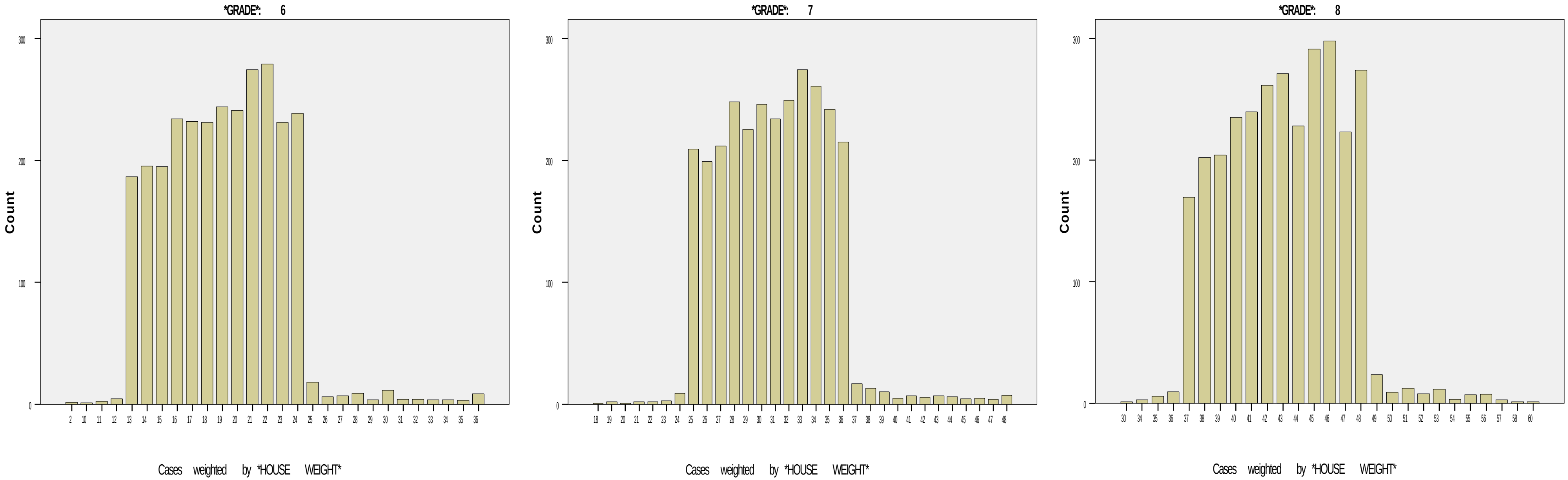

Figure 2: The distribution of the number of students over month of birth, for accelerated, normal-aged and delayed students.

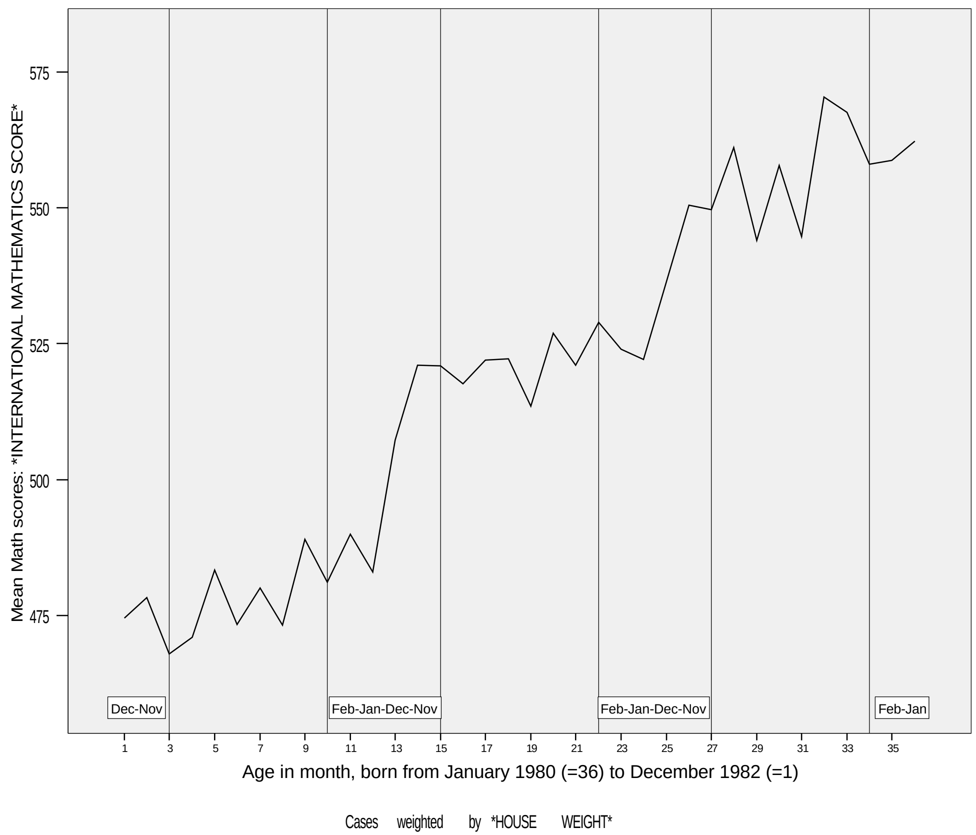

Figure 3: Mean scores for normal-aged students ranging from the youngest in grade 6 (December 1982) to the oldest in grade 8 (January 1980), for mathematics.