Summer learning and the effect of schooling: Evidence from Sweden

Lindahl, M.

Summer Learning and the Effect of Schooling: Evidence from Sweden

Mikael Lindahl

University of Amsterdam and IZA, Bonn

Current draft: January 24, 2001

(First draft: September 21, 1999)

ABSTRACT

Using that schools are in session during the school year and out of session during the summer, it is possible to isolate the effect of schooling on learning. This natural experiment situation can also be used to see whether schooling compensates for disadvantageous social backgrounds. Using a new sample of Swedish sixth grade pupils, results are that math skills are lost when pupils are not in school, pupils with non-Swedish parents learn relatively more during the school year, and learning is unrelated to pupils' parents' socioeconomic level during both the summer and the school year.

*Most of this study was conducted when I was affiliated with Swedish Institute for Social Research at Stockholm University. I thank Mahmood Arai, Anders Björklund, Per-Anders Edin, Fredrik Heyman, Martin Hörnqvist, Alan Krueger, Roope Uusitalo, Björn Öckert and seminar particpants at SOFI and FIEF for valuable comments and discussions. I also thank Peter Björklund and Ossian Wennström for excellent help with the data collection, and Judy Petersen for improving the English. I am most grateful to the pupils and teachers who participated in this study. I also thank HSFR and Handelsbanken for financial support.
**Address: Department of General Economics/NWO 'Scholar', Faculty of Economics and Econometrics, University of Amsterdam, Roeterstraat 11, NL-1018 WB, Amsterdam, The Netherlands. E-mail: mlindahl@fee.uva.nl

I. Introduction

A large literature on what makes school works has emerged from economists, especially in recent years. The aim has often been to estimating the effect of different school characteristics on pupils' scholastic achievement. The results from this literature are still somewhat mixed.[1] However, by only including observable school characteristics, this literature might miss important points regarding the total effect of schooling on pupils' achievement. In this paper I instead try to estimate the total effect of schooling, following a path from the educational sociology literature. This literature compares learning during the school year, when schools are open, with learning during the summer, when schools are closed. The thought is that learning during the summer vacation can be viewed as a reasonable counter factual to learning during the school year. First, by comparing the mean test score change during the school year, with the mean test score change during the summer vacation (adjusting for different period lengths), the causal effect of formal schooling on learning can be estimated. Hence, we can learn how much skill that are added to the average pupils knowledge level by schooling, compared with if this pupil did not go to school at all. Second, by comparing test score changes during the school year and summer for particular groups, it can be estimated which pupils who gain the most from schooling. By using information on pupils' demographics and social background, we can therefore learn whether schooling compensates for "disadvantage" backgrounds, i.e. whether schooling mitigates the strong connection between family background and scholastic achievement.

The US literature on summer learning is extensive. Cooper et al. (1996) survey the literature on summer learning before 1995. They also conduct a meta-analysis of

[1] See for instance Burtless (1996) and the papers in the symposium on school quality and educational outcomes in Review of Economics and Statistics (Moffitt 1996).

39 studies on summer learning. They find that achievement test scores decline when schools are not in session, and this negative, summer-learning effect increases with grade level. The achievement test- score decline is more pronounced for math than for reading, but it is not systematically related to racial background or gender. They also found that the achievement test-score decline is larger for pupils from low-income families in reading but not in math.[2] The literature studying both summer and school year learning are less extensive. A classical reference is Heyns (1978) who, using a reading test, found that pupils in Atlanta schools gained more during the school year compared to the summer and that learning during summer was negatively related to parental income. Entwisle, Alexander and Olson (1997) followed 790 Baltimore pupils from first to sixth grade. They found that school-year learning was unrelated to socioeconomic level but that for math and reading, summer learning was negatively associated with a pupil's socioeconomic level.[3] To my knowledge, studies of seasonal learning have not been done outside the US and Canada.

It is not clear whether or not the results from studies on seasonal learning carry over to conclusions about policy issues such as an extension of the school year or the introduction of summer programs targeted to specific groups. For this we must know what effect a marginal change in the school year length has on pupils' learning. Card & Krueger (1992) founds that term length has a significant positive effect on subsequent earnings and education, if other school characteristics are not controlled for. However, if pupil-teacher ratio and teacher wages are controlled for, the term-

[2] One should note however that few of these studies contained socioeconomic-level measures, such as parental income or education. And among those that did, the income measure was often whether or not the pupil was on a meal plan.

[3] Another early reference is Murnane (1975), who found that inner-city children in US gained reading skills during the school-year, whereas the development of reading skills where non-positive during the summer. O'Brien (1998) found that pupils who were non-white and from low-income families, lost math and reading skills during the summer, but that they gained during the school year. Also, see other studies by Entwisle and Alexander (1992, 1994), which analyze the influence of race on math and reading skills, using the seasonal approach.

length effect becomes insignificant. Frazier and Morrison (1999) compared pupils with longer and shorter school years and found that pupils having a longer school year learned relatively more. Pischke (1999) use the feature of a school reform in West-Germany in 1966-67. Before the reform, the school-year lengths varied across states, whereas the reform unified the number of weeks in schools across states. Results are that a shorter school year did lead to increased grade-repetition, but that no effect was found on later education or earnings. Regarding studies of the effect of summer programs on pupils' learning, in general only small effects have been found.[4]

This paper uses data from a new sample of pupils from schools in Stockholm, Sweden's capital. The sample contains information on scores from an identical math test for the same pupils in the spring of the fifth grade and in the fall and spring of the sixth grade, making it possible to compare test-score changes when school is in and out of session. The sample also contains socioeconomic measures from register data and data on pupils' demographics from a questionnaire answered by teachers. Hence, we can also investigate whether or not test-score changes vary by social backgrounds, during the school year and during the summer.

It should be of great interest to contrast the results from existing US studies with results from this study. One reason is that Swedish school policy has deliberately used schools as an instrument to equalize opportunities of individuals, probably to a larger extent than in the US.[5] Another reason is the smaller degree of income inequality in Sweden compared to the US, which points to more homogenous social background among pupils.[6] The conclusions from the US studies are that schools seem to have an equalizing role compared to the home and that the social

---

[4] See Heyns (1987) for a review of the literature on the effects of summer programs, and for a comparison with the literature on summer learning.
[5] See Erikson and Jonsson (1996).
[6] See Björklund (1998).

backgrounds of children affect the learning rate when children are out of school. Due to the different nature of the societies, conclusions from the US studies might not be transferable to Sweden.

The purpose of this paper is to answer the following questions. First, how much do pupils learn during the school year compared to the summer break? Second, is school year and summer learning correlated with socioeconomic background and demographic factors? Third, what is the causal effect on learning, of a change in the length of the school year? I also use the results to generate an estimate of the average rate of return to years of schooling, on earnings.

Section II presents the data collection procedure and descriptive statistics for the sample, and discusses some issues related to the data. Section III compares how much is learned when school is in and out of session. In section IV, learning during the school year and during the summer are regressed on covariates. Section V relates scores on a math test for earlier cohorts of pupils in Sweden to their later earnings. I use these results in an attempt to roughly calculate the effect of a longer school year on adult earnings. Section VI concludes.

II. Data

A. Data collection

The Swedish National Agency for Education distributes a math test in the early spring semester of the fifth grade to all schools in Sweden. This study's original sampling design was to randomly choose 16 schools among all the schools in the Stockholm municipality that had fifth grade classes (in the 1997/98 school year) and to add 4 schools, which were randomly chosen from areas in Stockholm with the highest socioeconomic level (2 schools) and the lowest socioeconomic level (2 schools). Math skills for all pupils in the classes, in all 20 schools, were to be tested in the fall and spring of sixth grade, using the same test that they took in the fifth grade.

But because some principals and teachers chose not to participate in this study, the sample was drawn differently. Stockholm schools were contacted at the start of the fall, sixth-grade semester in 1998. Of the first 22 randomly chosen schools (based on all schools in Stockholm that were in the official phone book and that had fifth grade classes), 10 schools that had sixth graders (in the 1998/99 school year) chose to participate. Two more schools, from parts of Stockholm similar to the second and third of the randomly sorted schools that did not participate, acted as replacements for these schools. In addition, two schools from the Stockholm area with the lowest socioeconomic level and two schools from the Stockholm area with the highest socioeconomic level were added. Within the participating schools, most (but not all) classes participated. The schools and classes that declined to participate stated that they were not able to hand over the fifth-grade tests because they felt they did not have the time or that they did not want to make the pupils "suffer" once again by making them take this test. The sample used in this study consists of 33 classes from 16 schools.[7]

The math test consisted of selected parts of the test that the Swedish National Agency for Education distributed in the spring of 1998. I distributed the math test to the pupils at the start and end of the sixth grade. The national math test in the spring of 1998 was administered without my participation. It had 10 separate parts of which six were to be done individually, two were to be done in pairs, and two were to be

[7] This is true for fifth grade. Because pupils at one of the included schools were transferred between fifth and sixth grade to another school in the sample, 32 school classes in 15 schools remained when the testing in the sixth grade was done. Also, there are a larger number of school classes, if a class is counted as the group of pupils that are taught in math. See Lindahl (2000).

done by groups of pupils. Of the six individual tests, I picked four (minus some unclear questions) to be repeated by the pupils at the start and end of the sixth grade. The four test parts were estimated to take an equal length of time to complete (30 minutes). These test parts included math questions on the four fundamental rules of arithmetic (part b), problems formulated in short text and on means (part c), problems based on newspaper ads and on area (part d), and problems based on a story the pupils had to read (part e). The completed tests were handed over to me for grading (with help from an assistant).[8] The maximum score on each test were 18 points. I was present in most schools during some parts of the test in the fall and spring of sixth grade.

The test taken by the pupils at the start and end of the sixth grade were to be done under as similar conditions as possible as the fifth grade test. The conditions, regarding time allowed and help given, varied among pupils within and among classes. I instructed teachers that the sixth grade tests should be conducted under the same conditions as the fifth grade test. Because I analyze changes in test scores between test occasions, different test conditions among schools should not be a serious problem. A total of 556 pupils took at least one part of the test on all three occasions and did the test under similar conditions on all three occasions.[9]

Among the 556 pupils, only 256 had done all parts of the test on all test occasions. To be able to use as large sample as possible, I deal with this by imputing scores for the missing test parts on each occasion. In other words, I regress the score on one test part on the scores on the other test parts on one occasion and predict

[8] Because I graded most of the test answers, there could be bias in the point settings. I compared points (based on corrected answers for 50 tests) that I graded with the same tests that an assistant graded. The correlation for all parts of the test was more than 0.92.

[9] A total of 569 pupils took at least one part of the test in the spring of fifth grade and spring and fall of sixth grade. For 11 pupils, conditions during test taking were stated by the teachers to be clearly

scores for the test parts with missing values. Because the same test was used on all three occasions, I can use these raw scores to calculate the summer and the school-year test-score changes.

Data on school, class, and teacher characteristics were taken from answers to a questionnaire, which I gave to the teachers at the time of the fall sixth grade test.[10] Teachers were asked to provide information on class sizes, their own education and experience and their pupils' characteristics (gender and number of pupils in the class that had no Swedish parent). Besides the questionnaire, I asked the teachers to list the pupils with no Swedish parent, to elaborate on unclear answers and to state whether or not mathematics was taught in classes other than the regular classes.[11] If so, the teacher would state which pupils belonged to which math class. Based on this, actual class sizes in mathematics were calculated.[12]

To get information on the pupil's social background, the addresses of the pupils (from the class lists) were matched with block data on the education and incomes of the parents (from registers).[13] For the purpose of this project, Statistics Sweden calculated block-means (for income) and block-fractions (for educational levels). Education was given in seven categories from which average years of schooling within each block were calculated.[14] The income variables were block-means and block-medians of family income (income from work and capital), block-

different between the first two test occasions and for two pupils, between the last two test occasions. These observations were deleted from the estimations.

[10] Appendix 1 contains the translated questionnaire.

[11] These questions were asked at the fall or spring sixth grade test occasion.

[12] This information is used in Lindahl (2000).

[13] Because the actual parents were not known, Statistics Sweden restricted the calculation on block-means and block-fractions to be for individuals of ages 28-54 and with children ages 10-12. Statistics Sweden required, for safety reasons, that there be at least three observations within each block. So if this was not fulfilled, I restricted the individuals to ages 28-54 with children.

[14] This was done by regressing years of schooling (from a survey questionnaire) on dummies for educational level achieved (from registers) using a sample with the same age restrictions from the 1991 Swedish Level of Living Survey. The estimates from this regression were then used to calculate mean years of schooling within each block.

means and block-medians of family disposable income and block standard deviations for these income variables. In this study, I use mean family income.[15] For the 8 pupils with missing address information, their class' average years of schooling and family income were assigned. This was also done in the additional five cases for which family income data was missing.

B. *Description of data*

Table 1 presents descriptive statistics for the sample used in this study. Test score is the sum of the scores on the four test parts. The minimum possible score is 0, and the maximum possible score is 72. The test seems to allow for improvements and reductions of the scores from one test occasion to another. This suggests that ceiling and floor effects are small. The test scores increased with test occasions. Correlations between the test scores at the different test occasions are between 0.73-0.78 (Table 2).

The demographic and family background variables are *girl*, *non-Swedish parents*, *parents' years of education*, and the *logarithm of family income*. *Non-Swedish parents* is an indicator that equals one if the pupil's parents do not have Swedish as their native language, and zero otherwise. The *socioeconomic index* variable is created by first standardizing parents' years of education and the logarithm of family income, and then taking the average of these standardized variables. Hence, parents' years of education and the logarithm of family income are weighted equally. Table 3 shows correlations among these variables. Among reasons to use the *socioeconomic index* variable in the analysis (instead of parents' years of education and the logarithm of family income) are that the last two are highly correlated and that we are not primarily interested in separate effects of these variables on pupils'

[15] I get very similar results if other income measures are used instead.

learning. When these variables are correlated with class size (see Lindahl, 2000) it is evident that pupils with non-Swedish parents and with parents with low education and income are more likely to be found in smaller classes. This confirms the fact that school resources in the Municipality of Stockholm are re-distributed with the aim to help disadvantaged neighborhoods.

A potential problem is within-class attrition of pupils. A total of 701 pupils belonged to the sampled classes in fifth and sixth grade and should have been tested on all three occasions. To get a picture of the seriousness of this within-class attrition, Table 4 presents family background characteristics for the 145 pupils that should have, but were not, available for testing on all occasions.[16] A comparison with Table 1 shows that sample attrition is more common among pupils with parents who are non-Swedish and who have lower-than-average education and income. Because some of these pupils took the test on two (but not all three) test occasions, we can check the sensitivity of attrition on the results by adding these pupils in the estimations. Adding the 35 pupils in the summer regressions and the 60 pupils in the school year regressions, for which test scores are available, does not statistically significantly alter the results.

C. Data issues

Potential problems with the data used in this study are that using the same test twice could lead to improvements in test scores without improvements in math skills. Another potential problem is that the tests were taken quite far from the end of the spring of fifth grade and the fall of sixth grade. For both reasons the absolute test-

[16] This included 13 pupils with different conditions between the test occasions.

score change during the summer could be overestimated. Potentially, estimates of other parameters could also be biased. In this section I present new ways to test for whether or not these problems exist.

1. Consequences of using the same test several times

If taking the same test more than once leads to an improvement in test scores without improving math skills (because pupils' learn to answer particular questions), this could lead to biased estimates. This could be particularly prevalent here because the same questions are repeated. Observe that the opposite could also be true, that is, taking the same test twice could make the pupils bored or tired, which could lead to underestimates of the true test-score change during summer.

Note that some pupils did not take all test parts on all occasions. Using this feature of the sample makes it possible to test whether or not re-test bias exists. By imputing scores on missing test parts in the spring of the fifth grade (using all existing test parts on this occasion), it is possible to compare score changes on the test parts taken on the first two test occasions, with the score changes on the test parts taken for the first time on the second test occasion. The scores in the test parts, where the fall of the sixth grade is missing, are not used in these estimations.

We can also test for re-test bias by only looking at within-pupils variations in test scores. To be included in the sample, the following conditions must then be met. First, the pupil must have taken at least one of the same test parts in spring of the fifth grade and fall of the sixth grade. Second, in the spring of the fifth grade, the pupil did not take at least one of the other test parts that were taken in the fall of the sixth grade. These conditions were met for 142 pupils.

The following equation is estimated:

(1) $$\Delta T_{sum,i,k} = a + bD_{i,k} + v_i + e_{i,k},$$

where $\Delta T_{sum,i,k}$ is the change in raw test scores between spring in the fifth grade and fall in the sixth grade for pupil i at test part k, *with* k=1,2,3,4; $D_{i,k}$ is a dummy that takes on the value one for pupil i if test part k was not taken in spring of the fifth grade, and zero otherwise; v_i is an unobservable individual-specific, test-score change effect and $e_{i,k}$ is a random error term.[17]

Columns 1 and 2 of Table 5 present estimates from a regression of equation (1) without individual fixed effects. The estimate in column 1 is negative and statistically significant. So pupils who took the fall of the sixth grade test for the first time score on average 0.94 raw points (or 0.24 standard deviations) lower on each test part. The estimates in column 2 reveal no statistically significant difference in the re-test bias among the four test parts. Columns 3 and 4 of Table 5 present estimates from equation (1), for the 142 pupils who did not take at least one test part in the spring of the fifth grade. On average, about one raw-point, test-score change for each test part can be attributed to the fact that the same test is used in spring of the fifth grade and in fall of the sixth grade. Because the tests are of different kinds, $D_{i,k}$ in equation (1) were interacted with dummies for different test parts. Even though the estimations in Columns 3 and 4 are done for a small sub-sample, the results from Columns 1 and 2 are basically confirmed. .

What then are the consequences of these results? If the re-test bias is the same between the spring fifth-grade scores and fall sixth-grade scores as between the fall

[17] The sample contains too few observations where the first test was taken in the spring of the sixth grade. So it is not possible to test for re-test bias in the estimate of the average school-year test-score change.

sixth-grade scores and the spring sixth-grade scores, then an estimate of the average difference between absolute learning during the school year and the summer will be consistent, since the biases will cancel out. If the re-test bias is independent of pupils' characteristics, a regression of summer and school-year learning on pupils' characteristics will also lead to consistent estimates. What will not lead to consistent estimates, without further adjustments, is an estimate of the average absolute learning during the school year and summer, separately. Also note that if an individual took a test part for the first time in the fall or spring of sixth grade, no re-test adjustment for these scores should be made. But for most individuals, re-test corrections will be necessary. One way to correct for this is to use the estimate of the re-test effects on the scores, from Table 5, and simply subtract this number (assumed constant among pupils) from the test score on each test part. When the re-test corrections are done in the rest of the paper, separate adjustment for each test part, based on the estimates in column 2 of Table 5, are used because these estimates are produced using the entire sample.[18]

2. Is learning linear?

Ideally, the tests should have been taken at the exact end and start of the semesters. This was not the case here and is hardly possible to administer in practice. There were, on average, 26.9 weeks between the spring of the fifth-grade and the fall of the sixth-grade test occasions, and 27.1 weeks between the fall and spring of the sixth grade test occasions. The pre-summer fifth-grade test was given in the February-June period.[19] The post-summer sixth-grade test was given from the last week of

---

[18] Very similar estimates are produced if the estimates from the other columns in Table 5 are used instead.

[19] The question to the teachers regarding when in the spring of the fifth grade the test were administered, concerned all test parts included in the National Agency of Education test.

September to the first week of November. And the pre-summer sixth-grade test was given within the last three weeks of the school year. On average, the pre-summer fifth-grade test was taken 9.7 weeks before the summer break and the post-summer sixth-grade test was taken 7.2 weeks after the summer break. In comparison with previous studies, our test dates could be somewhat further from the start and end dates of the summer.[20]

If learning is linear during the school year, we can simply predict scores at the end and start of the sixth grade. Further, if we assume that the individual learning rate is the same in the fifth and the sixth grade, conditional on observed school – and class variables, we can also predict scores at the end of the fifth grade. In figure 1 we illustrate how this would be done for the mean pupil, using actual data. By plotting the actual mean scores, we see that the scores increase between spring of the fifth grade and fall of the sixth grade, i.e. over summer. However, since there is some time between the test dates and the summer vacation, we see that if we assume linear learning, the mean test scores actually decreases over the summer. In previous studies of summer learning mean actual scores are simply used to determine whether or not test scores decrease over summer, although some authors are aware of the problem with non-ideal test dates.[21]

Since the assumption that pupils learn linearly are not obvious and to my knowledge has never been tested, we proceed by presenting a new way to test for this. This is done by using the fact that there is some variation in the time at which the fall sixth-grade tests were taken. The weeks between the fall and spring of the sixth grade test occasions vary between 24 and 31.5 weeks. The following equation is estimated:

---

[20] Entwisle, Alexander, & Olson (1997) state that the tests were administered in May and October. So between 17-26 weeks passed between the test occasions.

(2) $w_i = p_0 + p_1 t_i + u_i$,

where w_i is the weekly (absolute) learning rate during the sixth grade school year for pupil i; t_i is the weeks between the fall and spring sixth-grade test occasions for pupil i; and u_i is a random error term. The null hypothesis of linear learning is tested by testing p_1=0.

Table 6 presents the estimates of p_0 and p_1. Column 1 contains an estimate of only the constant. The weekly learning rate is 0.25 raw points on the sum of the four test parts. Columns 2 and 3 show that the weekly learning rate is negatively associated with the number of weeks between test occasions, that is, the closer between test taking occasions, the higher the weekly learning rate. This indicates that non-linear learning effects are present. Column 5 shows that when the scores are adjusted for re-test effects, the effect is still present. In column 6, where re-test adjustment is made and additional controls are included, the non-linear learning term is still negative but only weakly significant (p-value=0.114).[22] However, one must bear in mind that these estimates are unadjusted for the possibility that the pupils remembers the test better the closer they were taken during the sixth grade. Still, based on these results we will make several attempts to check the sensitivity of the results reported in the next two sections.

III. Absolute learning during the summer and during the school year

Table 7 presents levels and changes in the mean test scores. Column 1 of Table 7 reports mean raw scores on the three test occasions. The summer gain is 1.80

---

[21] See Heyns (1987) and Cooper et al. (1996).

[22] A straightforward calculation, from column 2, reveals that the weekly learning rate changes from positive to negative somewhere from 21-22 weeks between test takings. When we adjust for re-test

raw points (0.18 standard deviations), and the school year gain is 6.62 raw points (0.77 standard deviations). Thus, clearly less improvement is seen between the dates including the summer vacation. It also looks as though pupils actually gain over the summer. This might be because the tests were taken several weeks before and after summer break or because there are re-test effects from taking the same test twice.

A correction for re-test bias in the summer change estimate can be made if we assume that the re-test bias is independent of pupil characteristics. The estimates of re-test bias for the separate test parts in column 2 of Table 5 are used. For example, 0.58 points are subtracted from the point score on the fall sixth-grade *test part c* for those who took *test part c* in spring of the fifth grade. Assuming that the re-test bias is the same between the second and third test occasions, as between the first and the second test occasions, corrections for re-test bias on the spring sixth-grade test can also be done. Following the previous example, 1.16 points are subtracted from *test part c* on the spring sixth-grade test for those pupils who took this test part in the spring of fifth grade and in the fall of sixth grade.[23] Column 2 of Table 7 presents the resulting estimates. The summer change is now negative. The loss is 1.13 raw points (0.12 standard deviations). The school-year gain is 3.30 raw points (0.39 standard deviations).

As mentioned in Section II.C, the tests were in reality not taken at the end of fifth and at the start of sixth grade. If learning is linear during the school year, we can predict scores at the start and end of the sixth grade, because the tests were taken twice in the sixth grade. Also, we can predict test scores at the end of fifth grade, if

---

bias, the sign changes somewhere between 17-18 weeks of test takings. This does not necessarily mean that the weekly learning rate becomes negative after the same weeks of schooling.

[23] Assuming the same re-test bias between the second and third test occasions as between the first and the second test occasions, gives a lower bound estimate on the school's effect on learning, if the re-test bias between the first and the second test occasions, is the largest.

we assume that the learning pace is the same in fifth and sixth grade (conditional on changing school characteristics). Appendix 2 outlines the method for doing this and Figure 1 illustrates it using actual data. Column 3 of Table 7 presents levels and changes in predicted raw scores, assuming linear learning during the school year and during the summer. Skills then seem to be lost when school is out of session. The 10 weeks of summer break yield a loss of 2.18 raw points (0.16 standard deviations) on the test. Because about four times more points are gained over the 38-week-long school year, the first 10 weeks of the school year are needed to get pupils back to the same achievement level that they had at the end of the previous school year. The last row in column 3 presents an estimate of the causal effect of spending 10 weeks in school, assuming linear learning. This effect is estimated to be 4.64 raw points (0.29 standard deviations).

If learning increases, but at a decreasing rate during the school year, it might be more accurate to compare the raw test score in spring of the fifth grade with the predicted raw scores in fall of the sixth grade, assuming linear learning. The estimated summer loss will then be a lower bound of the true loss, since we assumed that nothing is learnt during the last 7.2 weeks of the fifth grade school-year, and that learning is linearly increasing (but not at an decreasing rate) during the first 9.7 weeks of the sixth grade school-year. Column 4 of Table 7 shows the estimates. The summer test-score change is a loss of 0.65 raw points. This difference is not statistically significant (p-value is 0.19). But if we also adjust for re-test bias, the summer loss increases to 2.28 raw points (0.2 standard deviations), which is statistically significant.

If the raw scores on the *test part e* (the combined math and reading test) are computed, it shows that scores increase between the spring of the fifth grade and fall of the sixth grade. This difference is weakly statistically significant (p-value is 0.12), but includes, on average, seven weeks of the sixth-grade school year. If a correction is made for re-test effects, test scores decrease during summer and this decrease is statistically significant.

This analysis suggests that, on average, sixth-grade pupils in Stockholm lose math skills when school is out of session. Less surprising is that the school environment, on average, is better at producing knowledge than the home environment, that is, the effect of schooling on learning is positive.

IV. Family background and learning during the summer and during the school year

By regressing summer and school-year test-score changes on pupils' social backgrounds, we can learn whether or not the school acts as an equalizing force with regard to test scores. If summer learning is positively related to socioeconomic level, and school-year learning is unrelated to socioeconomic level, which results from US studies suggest, this means that schooling acts as an equalizing factor.

It is well known that pupils with low socioeconomic backgrounds have a lower achievement level than pupils with high socioeconomic backgrounds. To see whether or not this is the case in this data set, I regress the level of test scores on girl, non-Swedish parent, and the socioeconomic variables. Table 8 reports the estimates. Pupils score higher if they have Swedish parents and if their parents have higher education and higher income. A pupil, who has parents that are Swedish and are in the 90^{th} percentile of the socioeconomic index, scored more than 12.6 raw points (almost

1 standard deviation) higher in the spring fifth grade test, compared to a pupil who has non-Swedish parents and are in the 10th percentile of the socioeconomic index.[24]

A. Methodological issues

How learning, when school is in session and when it is not, is related to pupils' social backgrounds can be estimated as:

$$(3) \qquad \Delta T_{sy,i} = a_0 + a_1 F_i + e_{sy,i}$$

$$(4) \qquad \Delta T_{sum,i} = b_0 + b_1 F_i + e_{sum,i},$$

where $\Delta T_{sy,i}$ and $\Delta T_{sum,i}$ are the changes in test scores, that is, learning of pupil i during the school year (*sy*) and during the summer (*sum*); F_i is a vector of demographic and family background variables for pupil *i*. These are a dummy for girl, a dummy for non-Swedish parents, and the value of the socioeconomic index; $e_{sum,i}$ and $e_{sy,i}$ are random error terms that are assumed to be uncorrelated with F_i. Learning over the school year in equation (3) might be affected by school characteristics, such as class size and teacher variables. Because these school characteristics probably are a function of family background, no school controls are included in the estimations. Also, note that the functional forms of equations' (3) and (4) are not clear from theory.

A potentially important issue is whether or not lagged test scores should be included as an explanatory variable. If so, instead of estimating equations' (3) and (4), we should estimate:

---

[24] Parents' education, Log(family income) and socioeconomic level in the 90th percentile are 14.79, 13.30 and 1.21, and in the 10th percentile are 9.80, 11.96 and –1.19, respectively.

(5) $$\Delta T_{sy,i} = a_0 + a_1 F_i + a_2 T_{bsy,i} + e_{sy,i}$$

(6) $$\Delta T_{sum,i} = b_0 + b_1 F_i + b_2 T_{bsum,i} + e_{sum,i},$$

where $T_{bsy,i}$ and $T_{bsum,i}$ are the start of the school year and start of the summer test scores.

An argument in favor of estimating equations' (5) and (6) is that ceiling and floor effects are probably present. Also, regression to the mean might be present, that is, it is easier for weak pupils to improve on tests, due to their low starting knowledge. Another issue concerns the type of the test. We cannot know a priori that the test design is such that an absolute improvement in scores is translated into a comparable absolute improvement in mathematical knowledge, which we want to capture, in all parts of the test-score distribution. Not including lagged test scores, is the same as assuming previous test scores having no effect on learning. An argument against including lagged test scores is that this is an endogenous variable with respect to family background characteristics. So if family background affects learning with a lag, this effect is then not captured. Also, if equations' (3) and (4) are estimated no adjustment for measurement error in the test scores is necessary. But if equations' (5) and (6) are estimated, badly measured test scores bias all estimates.[25]

In order to correct the estimates for measurement error in the test scores, we need to know the reliability of the test used in this study.[26] We can use that the four

[25] Normally classical measurement error in an explanatory variable will bias the effect of this variable towards zero. However, since in equations' (5) and (6) the lagged test score variable is included also as a part of the outcome variable this is no longer true. Instead the effect of lagged test scores on the change in test scores will be biased away from zero.

[26] Note that because the scores on the test parts are averaged, measurement errors due to "bad test days" are likely to be diminished. But there could still be measurement errors due to the failure of these particular test parts to capture true math skills.

parts of the test were of different kinds and in the spring of the fifth grade, they were all taken on different dates. So deviations from the true results on these parts should be independent. Using the estimates from the correlation matrix in Table 9, the reliability ratio is calculated to 0.7861.[27] The reliability ratio for fall and spring of the sixth grade tests are slightly higher (0.81 and 0.83). But because some of these test parts were taken on the same day, I use 0.7861 as an estimate of the reliability ratio of the test on all three occasions to correct the estimates in the regressions.[28] Another way to correct for this bias is to use the spring fifth-grade scores as an instrument for the fall sixth-grade scores, in school-year learning regressions. If there is no serial correlation in test scores, this will also correct for this type of bias. It is however not possible to use this last method in the summer-learning regressions, because we have no information on test scores before the spring of the fifth grade.

The choice of scaling of the test scores is not obvious. Whether or not the scores should be transformed from raw scores to, for example, percentile ranks, depends on how gains in different parts of the distribution should be weighted. Using percentile ranks implies lower weights to test-score gains in the upper and lower part of the distribution of test scores (because the distribution of test-score changes in this sample have relatively low tails). Disadvantages in using percentile scores are that ceiling and floor effects might become more severe.[29] I check the sensitivity to this choice by also reporting estimates using percentile ranks, when the result differs.[30]

---

[27] This is an estimate of the *alpha reliability ratio* (see Cronbach, 1951), which calculates the reliability ratio of the sum (or average) of several variables. So 0.7861 is an estimate of the reliability ratio for the sum of the raw scores on these four test parts. Note that the reliability ratio of a sum and an average of variables are the same. For a sensitivity analysis of the *alpha reliability ratio* to different assumptions, see Reuterberg & Gustafsson (1992).

[28] The reliability ratio for the summer change in test scores is 0.3461.

[29] This means that pupils at the lowest end of the distribution cannot lose, and pupils at the highest end of the distribution cannot gain in test scores between two points in time.

[30] Standardizing the test scores to a variable with a mean of zero and a standard deviation of one, produces exactly the same p-values in the regressions. But if changes of two standardized variables are

B. Results

Panel A of Table 10 presents results where summer learning is related to demographic and family background variables. No statistically significant effects from gender, parents' nationality, or socioeconomic index are found. The quadratic effect in column 5 can be rejected.[31] In column 6 an interaction term between non-Swedish parents and the socioeconomic index is added. It seems that pupils with parents who are non-Swedish and have a low value on the socioeconomic index lose the most during summer. Similar results are obtained if school indicators are controlled for.

In Panel B of Table 10, learning during the school year is regressed on family background variables. Having non-Swedish parents is positively related to school-year learning, even if socioeconomic level is controlled for. If a quadratic in socioeconomic level is added, this term is positive and statistically significant. This also makes the estimate on the dummy for non-Swedish parent decrease somewhat. If school dummies are added, similar results as before are produced, although the estimate on the socioeconomic index becomes more negative (p-value=0.14). If school characteristics, such as class size and teacher experience are added, the estimate for non-Swedish parent decreases to almost half of the estimate in column 3 in Panel A of Table 10 and becomes statistically insignificant. This suggests that the effect of having non-Swedish parents on school-year learning partly works through school characteristics.

---

used as dependent variable, this is not the case (because the difference between two standardized variables will produce a standard deviation that differs from one).

[31] Including the logarithm of family income gives a statistically insignificant linear effect but a statistically significant quadratic effect, which decreases with family income.

Table 11 show that if start-of-period test scores are controlled for, in column 2, the results change quite dramatically. In Panel A of Table 11, pupils with parents who are Swedish or have high socioeconomic level are estimated to gain more on the math test when school is out of session. In Panel B of Table 11, having parents with high socioeconomic level suggests a higher school-year gain, whereas non-Swedish parent is insignificant. In Table 11, the estimate on previous test scores is highly negative, indicating strong regression to the mean. One reason to be suspicious about this result is that the measurement error in start-of-period test scores will bias the estimate toward regression to the mean and as is clear from Tables 11, toward socioeconomic level effects on learning.

Table 11 also shows results from estimations in which seasonal learning is related to family background variables, with adjustment for measurement error in start-of-period scores. The regressions in Panel A of Table 11 impose the assumed reliability ratio of 0.7861, for the spring of the fifth-grade test score. None of the variables are statistically significant in column 3. Panel B of Table 11 presents estimates where several attempts are made to control for measurement error in test scores. The general picture is that pupils with non-Swedish parents learn more during the school year, but that parents' socioeconomic level does not have any (at least linear) effect on school-year learning.

The results in this section are not altered if re-test adjustments are made, if linear learning is assumed, or if corrections based on the non-linear learning assumption are made. It is worth pointing out that if the raw test scores are scaled in percentile ranks, the result that pupils with non-Swedish parents learn relatively more during the school year still holds, but becomes a bit more uncertain.[32]

[32] Using the sum of the raw scores, as in Table 10, the p-values from a test of no effect of non-Swedish parents are 0.01 without controls and 0.04 with controls. Using the percentile rank of the sum of the

One of the four test parts used in this paper was a test where the pupils were asked to read a story, and then answer math questions based on information given in the story (test part e). In order to investigate whether the results above holds for reading, I estimate seasonal regressions based on the results on this combined math-reading test. The results, presented in Table 12, are that pupils that have non-Swedish parents do lose skills during the summer but make equivalent gains during the school year. This is not surprising since pupils with non-Swedish parents are more likely to spend the summer talking their native language, and hence lose reading skills in Swedish.

When the test was taken in the spring of the fifth grade, only the teacher of the class monitored the pupils when they took the test. During the fall of the sixth grade test, the teacher and I, monitored the pupils when they took half the test. During the other half of the test (taken on a different day), only the teacher was present. Hence, we can test for whether or not teachers try to affect the results of their classes, and if this varies systematically with the characteristics of the pupils.[33] Doing this, I find the test-score changes being positively related to parents' socioeconomic level, when I do the analysis for the parts that I did not monitor. This could suggest that teachers in classes that consist of pupils with disadvantaged socioeconomic backgrounds try to raise the test scores of those pupils, perhaps through extra help.

The conclusion from the exercise in this section is that no consistent effect from parents' socioeconomic level (that is education and family income) on seasonal learning is found. This is contrary to results for the US. In Sweden, school-year learning is higher for pupils with non-Swedish parents. The difference between

---

raw scores, the p-values from a test of no effect of non-Swedish parents are 0.11 without controls and 0.08 with controls. Using the sum of the percentile ranks of the separate raw test part scores, the p-values from a test of no effect of non-Swedish parents are 0.09 without controls and 0.18 with controls.

[33] This can be done for 400 pupils only.

school-year and summer learning is weakly statistically significantly related to the dummy for non-Swedish parents (p-value=0.054 with no controls, and p-value=0.162 with controls). Using the combined math-reading test however, we do find that pupils' with non-Swedish parents lose relatively more skills during the summer, but make equivalent gains during the school year. This mimics the results for blacks and low-income families in the US.

V. Test scores, the length of the school year and subsequent earnings

In this section I estimate the effect of math test scores on subsequent earnings for an earlier cohort of sixth-grade pupils in Sweden. This is done in order to understand the magnitude of the estimates in the previous sections, to be able to calculate a lower-bound estimate of the effect of an increase in the length of the school year on subsequent earnings, and to learn about the importance of math test scores for earnings in Sweden in general.

A. The impact of math skills on subsequent earnings using another data set

Several studies in Sweden have included test scores in earnings regressions.[34] Kjellström (1999) and Meghir & Palme (1999) estimate the rate of return to education, in part by controlling for test scores. Here, I use the same data source as they do.[35] However, since the purpose is to estimate the total return to test scores, instead of the return to test scores controlling for subsequent education, I do the analysis without controlling for education and other variables, that in part are determined by test scores.

[34] It is not possible to relate the scores of the test, used in this paper, to subsequent earnings, since this type of test just recently were starting to be given to pupils in Sweden.
[35] See Härnqvist et al. (1994) (in Swedish) for a description of the data and for a review of studies that have used this data. See Kjellström (1999) for a description of the data in English.

Table 13 presents descriptive statistics. The individuals included in the Gothenburg data set were born in 1948 and 1953. Data on test scores on a standardized math test and family background variables are from when the pupils were in sixth grade. So even if the standardized math test is not the same as the one I use in previous sections in this paper, the results should be comparable. If the estimates regarding math scores on earnings have changed over time, the comparison is harder to make. If the effect of test scores on earnings in Sweden has increased over time, as it has in the US, the causal effect of math scores on earnings will be underestimated when this estimate is generalized to later cohort of pupils.[36]

In Table 14, I report estimates from a regression of the logarithm of yearly earnings on the standardized scores on the math- and IQ tests. Because the math score might be related to exogenous family background variables, I have (in all estimations) controlled for gender, father's education, and mother's education. In column 1, the estimates on math score are 0.13 and 0.09. In column 3, I also control for scores on IQ tests taken in the same grade. Controlling for IQ should give an underestimation of the effect of math score on earnings, because a longer school year probably also affects the IQ score. Maybe somewhat surprisingly, the estimate on the math score only decreases somewhat to 0.11 and 0.06. In columns 2 and 4, a quadratic in the math score is tested. The quadratic is positive and cannot be rejected for the 1953 cohort, which implies that a standard deviation gain in the math score might have higher payoff at higher levels of math skill. In columns 5 and 6, one's own education, which clearly is endogenous with respect to math and IQ scores, is controlled for. The estimates indicate that roughly half of the total math-score effect goes through education. Assuming that these estimates will be about the same now, we at least get a

[36] See Murnane et al. (1995), for evidence of test-score trends for the US.

hint that an increase on math test scores by, for example, one standard deviation gives, on average, 6-13% higher earnings.

B. *The effect of a longer school year on subsequent earnings*

By using the results from the previous sections, a tentative evaluation of the effect of a longer school year on test scores and on subsequent earnings can be made. The goal with the analysis in this section is to calculate a likely lower-bound estimate of the effect of a one-week increase in the school year on test scores and subsequent earnings. In Sweden, the summer vacation is 10 weeks and the (effective) school year is 38 weeks.[37] Suppose the summer break is reduced to 9 weeks. In Section 3, the difference between 10 weeks of school-year learning and 10 weeks of summer learning were estimated to give a raw points increase between 3.03-3.44 if scores from the fall of the sixth grade were predicted (last row of columns 4-5 of Table 7). This estimate accounts for non-linear learning, because only the fall of the sixth grade (and not the spring of the fifth grade) scores were predicted. However, the estimate of the school-year learning was an estimate of the average learning during the school year. Because the analysis in Section 2.3.2 indicated that learning was positive, but decreasing during the school year, here we assume that pupils' learn nothing during the last 7.2 weeks at school in fifth grade. So a lower-bound estimate of the causal effect, for the average pupil, of a one-week increase in the school year might be estimated to give a 0.23 raw point increase (from row 4 of column 5 of Table 7), assuming that the adjustments for re-test bias are correct. This corresponds to a test-score increase by 0.02 standard deviations

[37] About four weeks of vacation, spread over three occasions, are included in the 42 weeks of the school year.

Assuming that the effect of a standard deviation increase in math scores on log earnings is 6% (from column 3 of Panel B in Table 14), we find that earnings will increase by 0.12%. It is important to remember that this calculation is done only for an increase in the school year in one grade. If these increases were made in grades 1-9 (or if the school year increase by 8 weeks in one year) and if the effect of time in school on math score and the effect of math score on earnings are roughly the same as in sixth grade, then we would expect a 0.96% increase in yearly earnings from this policy.[38]

To illustrate the magnitude of these effects, we can compare them to previous estimates of the causal effect of years in school on earnings for Sweden. In previous studies that use the same earnings/math score data as in this study, an increase in years of schooling by one year was estimated to increase yearly earnings by about 4%.[39] An increase in the compulsory number of years in school by one year is comparable to an increase in time in school by 38 weeks. Dividing these 38 weeks among the nine years of schooling gives a 4.2-week increase of the school year, each year in school. Multiplying this number with the causal effect of one week spent in school on earnings reported previously, we get 4.03%. So estimating the causal effect of an increase in years of schooling by one year gives a number very close to estimates from the literature on the rate of return to education.

In Lindahl (2000), that uses the same data as in this paper, the lower bound effect of decreasing class size by one pupil was about one-third percentile rank higher scores on the math test.[40] An increase in the school year by one week is roughly a 2.6% (1 week divided by the 38 weeks that constitute a school year) increase in the

---

[38] Nine years of schooling were the number of compulsory years in school for most of the sample.
[39] See Kjellström (1999), Meghir & Palme (1999). Also, see Isacsson (1999), who uses a large sample of Swedish twins to estimate the returns to education, which produces quantitatively similar results.
[40] Most of the estimates of class-size effects in Lindahl (2000) were however much higher.

school year. A 2.6% lowering of the average class size would compare to about 0.5 fewer pupils in the average class.[41] This would give about a 0.17 percentile point (or 0.007 standard deviations) higher score on the math test in the sixth grade, which gives about 0.04% higher earnings, using the estimation result from Table 14 in this paper (from Column 3 of Panel B in Table 14). This is lower compared to the gain from an increase in the school year by one week in one grade, which was 0.12%.

VI. Conclusion

In Sweden, time spent in school seems to affect learning positively, compared to time spent at home. That is, schooling has a positive effect on test scores. There is also evidence that pupils lose math skills during the summer. But the evidence does not suggest that pupils from disadvantaged social backgrounds lose relatively more during the summer in Sweden. However, there is evidence that pupil's with non-Swedish parents learn more during the school year. Because having Swedish parents is strongly associated with pupils scoring high on the math tests, it seems that the school does a good job in narrowing this gap in test scores. When a variable measuring class size is added to the school-year regressions, the effect of non-Swedish parents on learning decreases and becomes statistically insignificant. Since in Stockholm, class sizes are lower in neighborhoods with many non-Swedish individuals, it seems that the re-distribution of school resources toward these neighborhoods has paid off. More speculatively, if the results are interpreted in light of the differences between the Swedish and the US societies, it could be the case that since in the US, social backgrounds are more heterogeneous, this leads pupils with disadvantaged social backgrounds to significantly losing skills when schools are

[41] If teacher time were perfectly divisible, this comparison would be correct.

closed. Instead, the Swedish society might be homogenous enough to prevent this from happening. It should however be remembered that the sample used in this study is small and that test taking occurred quite long into the semesters. Although several attempts to correct for this were made, both factors would lead to bias toward insignificant estimates of family background variables on summer learning. Another explanation could be that results differ depending on whether math or reading skills is measured. When a test part, measuring a combination of math and reading skills were used, the results for the socioeconomic index were the same as before, but the results for non-Swedish parents changed to showing that pupils with non-Swedish parents lose skills relatively, but gained equally during the school-year.

This paper has made several attempts to improve upon existing studies of summer and school-year learning. We tested for the presence of re-test bias in the test scores, and found that the estimates of absolute test score changes needed to be corrected by an amount sufficient to change the direction of the summer estimate. Although many of the other studies use similar but non-identical tests (which has the drawback that they are unlikely to be equally difficult), there is still likely to be re-test bias present even then. Since all previous studies of summer learning built on tests that are taken sometime before and after the actual start and ending of the summer vacation, we tested for whether linearity in learning would be a reasonable assumption to correct for this (by predicting scores). We did however find evidence indicating that learning (at least on this test) increases but at a decreasing rate. Since summer programs often extend the school year for those pupils attending them, the non-linear learning result may explain why studies of these programs usually have found small or non-existent effects on test scores. In this paper it was also found that teachers in classes where the pupils comes from families with low socioeconomic levels, the teachers in these classes seem to adversely try to affect the test results positively.

I also laid out an alternative way to estimate the rate of return to time in school, on earnings. This was done by combining an estimate of the causal effect on math test scores of increasing the school year, with an estimate of the effect of math test score on earnings. Despite strong assumptions underlying this calculation, this number was surprisingly similar to previous estimates of the returns to years of schooling in Sweden, using the traditional Mincer approach. Also, a comparison with an estimate of the causal effect of class size on test scores, from Lindahl (2000) was made. The benefits from an increase in the length of the school-year seemed more beneficial compared to an equivalent decrease in class sizes.

Appendix 1: Questionnaire to Teachers

1. What class is your responsibility?

2. How many pupils do you now have in your class?

3. How many pupils did you have in the same class during the last school year?

4. Do you have assistance from other teachers in your teaching? (extra teachers, special teachers, subject teachers)

5. How many pupils in your class do you think do not have parents whose native language is Swedish?

6. How many years have you spent in formal education?

7. What is your highest academic degree?

8. Which type of degree was this?

9. Are you a certified teacher?

10. How long have you been teaching?

11. How many years have you been teaching in your present class?

12. How many pupils in your class do you estimate were attending some sort of summer school course?

13. Did your school organize some education over the summer for pupils in your class?

14. During what period was the national test in mathematics distributed to students in grade 5 during the spring semester 1998?

15. Did you practice on old national tests in mathematics with the class before this point in time?

16. How much time do you estimate that your class spent on preparations for this particular test?

17. Have you gone through the solutions and the answers on the math test done in spring semester 1998 with the pupils? Was this done during the spring semester 1998 or the fall semester 1998? How well did you go through the math tests with the pupils? Describe.

18. Did you follow the time recommendations that were given by the National Agency of Education regarding the different individual test parts in mathematics?

19. If you did not follow these time recommendations, did the pupil get much longer, longer, or shorter time? If you can, state how much time they were allowed for doing the different individual test parts in mathematics.

20. How much active help did you give the pupils when the different test parts were taken?

Appendix 2: Prediction of test scores at start and end of school years

The predicted scores at the start and end of the sixth grade are calculated in the following way. The weekly test-score change for pupil i in sixth grade can be expressed as:

$$(A-1) \qquad w_{6,i} = \frac{T_{6_2,i} - T_{6_1,i}}{t^a_{6_2} - t^a_{6_1}},$$

where $T_{6_2,i}$ and $T_{6_1,i}$ are test scores in spring and fall of sixth grade, for pupil i; $t^a_{6_2}$ and $t^a_{6_1}$ are number of weeks into the sixth grade that the tests were taken, in spring and fall of the sixth grade, for pupil i.

Assuming that weekly learning rate is constant during the school year, the predicted test scores at the first and last week of sixth grade, can be calculated as:

$$(A-2) \qquad T^p_{6_2,i} = T_{6_2,i} + (w_{6,i} \times (t_6 - t^a_{6_2,i}))$$

$$(A-3) \qquad T^p_{6_1,i} = T_{6_1,i} - (w_{6,i} \times t^a_{6_1,i}),$$

where t_6 is the length of the school year, in weeks, which is set to 38.

Because we do not have two test occasions in fifth grade, we cannot use the same method to calculate predicted test scores at the end of fifth grade. One alternative is to assume that the rate of weekly learning is the same in fifth grade, as in sixth grade. Even though many characteristics are constant between these grades, some are not. For example, some school characteristics probably changed. So I try to account for this by correcting the fifth-grade learning rate for the changes in class size between the grades. I first estimate the following equation:

$$(A-4) \qquad w_{6,i} = c_6 + b_6 CS_{6,i} + e_{6,i},$$

where $CS_{6,i}$ is class size in sixth grade for pupil i and c_6 and b_6 are parameters to be estimated. Estimation of (A4) gives estimates $\hat{c}_6$, $\hat{b}_6$ and $\hat{e}_{6,i}$, where $\hat{e}_{6,i}$ is the estimated residual. I then calculate the weekly learning rate in fifth grade as:

$$(A-5) \qquad \hat{w}_{5,i} = \hat{c}_6 + \hat{b}_6 CS_{5,i} + \hat{e}_{6,i},$$

where $CS_{5,i}$ is the class size in fifth grade, for pupil i.

Then predicted test scores the last week in sixth grade is calculated as:

$$(A-6) \qquad T^{p}_{5_2,i} = T_{5_2,i} + (\hat{w}_{5,i} \times (t_6 - t^{a}_{5_2,i})),$$

where $T_{5_2,i}$ is the test scores in spring fifth grade for pupil *i*.

Observe that in this paper, when adjustments to non-linear learning are made, only the predicted test scores at the fall of the sixth grade (*A3*) are used, in addition to the raw scores for the spring of the fifth and the sixth grade.

References

Björklund, Anders. 1998. "Income Distribution in Sweden: What is the Achievement of the Welfare State." *Swedish Economic Policy Review* 5(1): 39-80.

Burtless, Gary (ed.). 1996. *Does Money Matter? The Link Between Schools, Student Achievement and Adult Success* (Washington, D.C.: Brookings Institution).

Card, David and Alan B. Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy* 100: 1-40.

Cooper, Harris, Barbara Nye, Kelly Charlton, James Lindsey, and Scott Greathouse. 1996. "The Effect of Summer Vacation on Achievement Test Scores: A Narrative and Meta-Analytic Review." *Review of Educational Research* 66(3): 227-268.

Cronbach, Lee J. 1951. "Coefficient alpha and the internal structure of tests," *Psychometrika* 16: 297-334.

Entwisle, Doris R., Karl L. Alexander. 1992. "Summer Setback: Race, Poverty, School Composition, and Mathematics Achievement in the First Two Years of School." *American Sociological Review* 57(February): 72-84.

---------. 1994. "Winter Setback: the Racial Composition of Schools and Learning to Read." *American Sociological Review* 59(June): 446-460.

Entwisle, Doris R., Karl L. Alexander and Linda Steffel Olson. 1997. *Children, Schools and Inequality.* Boulder, CO: Westview Press.

Erikson, Robert, and Jan O. Jonsson. Eds. 1996. *Can Education Be Equalized?*. Westview Press.

Frazier, Julie A. and Frederick J. Morrison. 1998. "The Influence of Extended-Year Schooling on Growth of Achievement and Perceived Competence in Early Elementary School." *Child Development* 69(2): 495-517.

Heyns, Barbara. 1998. *Summer Learning and the Effects of Schooling*. New York: Academic Press.

---------. 1987. "Schooling and Cognitive Development: Is There a Season for Learning?." *Child Development* 58(5): 1151-1160.

Härnqvist, Kjell, Ingemar Emanuelsson, Sven-Eric Reuterberg and Allan Svensson. 1994. "Dokumentation av Projektet "Utvärdering Genom Uppföljning." Report No. 1994:03, Institute for pedagogy, Gothenburg University.

Isacsson, Gunnar. 1999. "Estimates of the Return to schooling in Sweden from a Large Sample of Twins." *Labour Economics* 6(4): 471-490.

Kjellström, Christian. 1999. *Essays on Investment in Human Capital*, Doctoral Dissertation, Swedish Institute for Social Research, Stockholm University.

Krueger, Alan B. 1998. "Reassesing the View That American Schools Are Broken." *Federal Reserve Bank of New York Economic Policy Review* 4(1): 29-43.

Lindahl, Mikael. 2000. "Home vs. School Learning: A New Approach to Estimating the Effect of Class Size on Achievement." Study 3 in: *Studies of Causal Effects in Empirical Labor Economics* Doctoral thesis, Swedish Institute for Social Research, Stockholm University.

Meghir, Costas and Mårten Palme. 1999. "Assessing the Effect of Schooling on Earnings Using a Social Experiment." SSE/EFI Working Paper Series in Economics and Finance, No. 313.

Moffitt, Robert A. 1996. "Symposium on School Quality and Educational Outcomes: Introduction," *Review of Economics and Statistics* 78(4): 559-561.

Murnane, Richard J. 1975. *The Impact of School Resources on the Learning of Inner-City*

Children. Cambridge, MA: Ballinger.
Murnane, Richard J., John B. Willett, and Frank Levy. 1995. “The Growing Importance of Cognitive Skills in Wage Determination.” *The Review of Economics and Statistics* 77(2): 251-266.
O’Brien, Daniel M. 1998. ”Does Minority and Disadvantaged Children Suffer Summer Fallback in Cognitive Skills?” Paper prepared for presentation at the Meetings of the Southern Economic Association in November 1998.
Pischke, Jörn-Steffen. 1999. “Does Shorter Schooling Hurt Student Performance and Earnings? Evidence from the German Short School Years.” Mimeo, MIT.
Reuterberg, Sven-Eric, and Jan-Eric Gustafsson. 1992. “Confirmatory Factor Analysis and Reliability: Testing Measurement Model Assumptions,” *Educational and Psychological Measurement* 52(4): 795-811.

Table 1: Descriptive statistics

	Mean	St. Dev	Min	Max
Test scores (raw scores)				
Fifth-grade, spring	42.32	12.93	8.63	68.87
Sixth-grade, fall	44.12	13.43	5.00	70.50
Sixth-grade, spring	50.74	12.28	12.00	71.00
Demographic and family background variables				
Girl	0.50	0.50	0	1
Non-Swedish parents	0.23	0.42	0	1
Parents' years of education	12.36	1.96	7.53	19.67
Log (family income) [Family income in SEK]	12.60 [345,300]	0.54 [240,434]	11.19 [72,499]	14.75 [2,557,953]
Socioeconomic index	0.00	0.93	-2.51	2.52

Notes: Number of observations is 556. Test scores pertain to the national mathematics tests. Missing values are imputed. The socioeconomic index is calculated as the average of the standardized values of parents' years of education and log of family income.

Table 2: Correlation matrix of the test scores

	Fifth-grade, spring	Sixth-grade, Fall	Sixth-grade, spring
Fifth-grade, spring	1.00		
Sixth-grade, fall	0.73	1.00	
Sixth-grade, spring	0.73	0.78	1.00

Notes: Number of observations is 556.

Table 3: Correlation matrix for demographic and family background variables

	Gender (girl=1)	Non-Swedish parents	Parents' education	Log (family income)	Socio-economic index
Girl	1.00				
Non-Swedish parents	-0.03	1.00			
Parents' education	0.03	-0.50	1.00		
Log(family income)	0.01	-0.55	0.72	1.00	
Socioeconomic index	0.02	-0.57	0.93	0.93	1.00

Table 4: Features of pupils who are not but should have been included in the sample used in this study

	Mean	St. Dev.	Min	Max
Not in sample (n=145)				
Girl	0.46	0.50	0	1
Non-Swedish parents	0.41	0.49	0	1
Parents' education	11.79	2.14	7.02	16.64
Log (family income)	12.45	0.60	11.19	14.75

Table 5: Test for re-test bias

	Dependent Variable: Summer, test-score change, (raw scores)			
	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)
Dummy for not taking spring fifth-grade test (D)	-0.94 (0.28)		-0.96 (0.34)	
D*Dummy for test part b		-0.67 (0.56)		0.36 (0.81)
D*Dummy for test part c		-0.58 (0.86)		-0.81 (0.92)
D*Dummy for test part d		-1.11 (0.55)		-0.54 (0.72)
D*Dummy for test part e		-1.00 (0.37)		-2.36 (0.66)
Pupil dummies	No	No	Yes	Yes
R^2	0.013	0.013	0.410	0.419
No. of pupils (n)	556	556	142	142
No. of observations (n*no. of test parts)	2101	2101	542	542

Notes: Mean (Std) of dependent variable in columns 1-2 is 0.42 (3.98). Mean (Std) of dependent variable in columns 3-4 is –0.07 (3.75). All columns include dummies for the separate test parts. The first two columns include controls for non-Swedish parent, girl, parents' education and family income.

Table 6: Test for non-linear learning

Dependent variable: Weekly school-year learning (raw scores)						
	Unadjusted			Adjusted for re-test bias		
	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.247 (0.020)	0.647 (0.176)	0.739 (0.554)	0.119 (0.021)	0.563 (0.190)	0.464 (0.575)
Weeks between tests		-0.015 (0.007)	-0.013 (0.006)		-0.016 (0.007)	-0.012 (0.007)
Additional controls	No	No	Yes	No	No	Yes
R^2	-	0.010	0.045	-	0.012	0.047

Notes: Number of observations is 556. Additional controls are girl, non-Swedish, parents' education, family background and schooling variables (class size, teacher experience, teacher experience squared and teacher experience in class). Standard errors allow for within school correlated regression errors.

Table 7: Absolute test-score changes

Test scores	Raw scores	Scores corrected for re-test bias	Predicted scores (assuming linear learning)	Predicted scores (allowing non- linear learning)	Predicted scores (allowing non-linear learning, and correcting for re-test bias)
	(1)	(2)	(3)	(4)	(5)
Fifth-grade, spring test score	42.32 [12.93]	42.32 [12.93]	43.86 [12.97]	42.32 [12.93]	42.32 [12.93]
Sixth-grade, fall test score	44.12 [13.43]	41.19 [13.26]	41.67 [15.07]	41.67 [15.07]	40.04 [14.96]
Sixth grade, spring test score	50.74 [12.28]	44.48 [12.19]	51.04 [12.40]	50.74 [12.28]	44.42 [12.11]
Summer, test score change	1.80 (0.41) [9.75]	-1.13 (0.41) [9.68]	-2.18 (0.56) [13.26]	-0.65 (0.49) [11.59]	-2.28 (0.49) [11.56]
School-year, test score change	6.62 (0.37) [8.62]	3.30 (0.36) [8.57]	9.37 (0.52) [12.15]	9.06 (0.50) [11.75]	4.38 (0.50) [11.73]
School-year change minus summer change	4.82 (0.67) [15.89]	4.42 (0.67) [15.75]	-	-	-
School-period change (10 weeks) minus summer-period change (10 weeks)	0.66 (0.49) [4.64]	2.00 (0.46) [10.96]	4.64 (0.67) [15.84]	3.03 (0.59) [13.88]	3.44 (0.59) [13.85]

Notes: Number of observations is 556. Missing values imputed. Standard errors are in parentheses and standard deviations are in brackets.

Table 8: Level regressions

Dependent Variable	Fifth grade, spring test scores		Sixth grade, fall test scores		Sixth grade, spring test scores	
	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Girl	-0.35	-0.36	-0.58	-0.57	0.33	0.35
	(1.01)	(1.01)	(0.98)	(0.99)	(0.92)	(0.92)
Non-Swedish parents	-4.60	-4.67	-5.32	-5.28	-3.09	-2.99
	(1.41)	(1.41)	(1.84)	(1.84)	(1.55)	(1.55)
Parents' education	0.61		0.98		1.17	
	(0.41)		(0.43)		(0.37)	
Log (family income)	3.96		2.55		1.88	
	(1.42)		(1.55)		(1.47)	
Socioeconomic index		3.31		3.31		3.34
		(0.68)		(0.83)		(0.64)
R^2	0.121	0.120	0.123	0.122	0.105	0.104

Notes: Number of observations is 556. Standard errors (in parentheses) are adjusted for blocks-clustering. Test scores are measured in raw scores.

Table 9: Correlation matrix of scores on parts of the fifth-grade spring test

	Part b	Part c	Part d	Part e
Part b	1.00 (527)			
Part c	0.57 (512)	1.00 (529)		
Part d	0.48 (453)	0.50 (457)	1.00 (478)	
Part e	0.50 (445)	0.49 (451)	0.43 (434)	1.00 (469)

Notes: The scores are from parts of the spring fifth-grade test occasion. Number of observations used for each correlation is in parentheses

Table 10: Seasonal regressions

A. Summer Regressions

	Dependent variable: Summer, test-score change (raw scores)					
	(1)	(2)	(3)	(4)	(5)	(6)
Girl	-0.20			-0.21	-0.20	-0.23
	(0.85)			(0.83)	(0.83)	(0.85)
Non-Swedish parents		-0.60		-0.61	-1.16	-2.32
		(1.01)		(1.15)	(1.20)	(1.53)
Socioeconomic index (SES)			0.15	-0.00	-0.17	0.39
			(0.49)	(0.54)	(0.53)	(0.73)
Socioeconomic index squared					0.52	
					(0.46)	
SES*Non-Swedish parents						-2.27
						(1.36)
R^2	0.000	0.001	0.000	0.001	0.003	0.005

B. School-year regressions

	Dependent variable: School-year, test-score change (raw scores)					
	(1)	(2)	(3)	(4)	(5)	(6)
Girl	0.87			0.92	0.93	0.91
	(0.71)			(0.71)	(0.71)	(0.71)
Non-Swedish parents		2.23		2.29	1.57	0.84
		(0.82)		(1.09)	(1.18)	(1.48)
Socioeconomic index (SES)			-0.55	0.03	-0.20	0.35
			(0.42)	(0.53)	(0.55)	(0.56)
Socioeconomic index squared					0.68	
					(0.35)	
SES*Non-Swedish parents						-1.93
						(1.23)
R^2	0.003	0.012	0.004	0.015	0.020	0.019

Notes: Number of observations is 556. Standard errors (in parentheses) are adjusted for blocks-clustering. The dependent variable Panel A is the change in raw scores between spring of the fifth grade and fall of the sixth grade. This variable has a mean (std) of 1.80 (9.75). The dependent variable in Panel B is the change in raw scores between spring of the fifth grade and fall of the sixth grade. This variable has a mean (std) of 6.62 (8.62).

Table 11: Seasonal regressions, sensitivity analysis

A. Summer Regressions

	Dependent variable: Summer, test-score change (raw scores)			Dependent variable: Summer, test-score change (predicted scores at sixth grade, fall)		
	(1)	(2)	(3)	(4)	(5)	(6)
Girl	-0.21	-0.31	-0.23	-0.60	-0.70	-0.61
	(0.83)	(0.77)	(0.65)	(1.01)	(0.94)	(0.84)
Non-Swedish parents	-0.61	-1.94	-0.87	-1.29	-2.53	-1.43
	(1.15)	(1.39)	(0.96)	(1.56)	(1.62)	(1.24)
Socioeconomic index (SES)	-0.00	0.94	0.18	0.01	0.89	0.11
	(0.54)	(0.65)	(0.44)	(0.75)	(0.77)	(0.57)
Fifth grade, spring test score		-0.28	-0.05		-0.27	-0.03
		(0.03)	(0.04)		(0.04)	(0.05)
Measurement error adjustment	Not needed	No	Yes, R=0.7861	Not needed	No	Yes, R=0.7861
R^2	0.001	0.126	-	0.003	0.080	

B. School-year regressions

	Dependent variable: School-year, test-score change (raw scores)				Dependent variable: School-year, test-score change (predicted scores at sixth grade, fall)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girl	0.92	0.75	0.88	0.88	1.31	1.08	1.23
	(0.71)	(0.64)	(0.50)	(0.69)	(0.97)	(0.87)	(0.91)
Non-Swedish parents	2.29	0.68	1.87	1.94	2.96	0.79	2.43
	(1.09)	(0.93)	(0.73)	(1.07)	(1.49)	(1.29)	(1.44)
Socioeconomic index (SES)	0.03	1.03	0.29	0.25	0.02	1.38	0.32
	(0.53)	(0.43)	(0.34)	(0.51)	(0.71)	(0.59)	(0.67)
Sixth grade, fall test score		-0.30	-0.08	-0.07		-0.41	-0.09
		(0.03)	(0.03)	(0.05)		(0.04)	(0.06)
Measurement error adjustment	Not needed	No	Yes, R=0.7861	Yes, IV	Not needed	No	Yes, IV
R^2	0.015	0.212	-	-	0.014	0.209	-

Notes: Number of observations is 556. Standard errors (in parentheses) are adjusted for blocks-clustering. The dependent variable in the first three columns of Panel A is the change in raw scores between spring of the fifth grade and fall of the sixth grade. This variable has a mean (std) of 1.80 (9.75). The dependent variable in columns 1-4 of Panel B is the change in the raw scores in fall and spring of the sixth grade. This variable has a mean (std) of 6.62 (8.62). The dependent variable in columns 4-6 of Panel A is the change in the raw scores in spring of the fifth grade and the predicted raw scores in the fall of the sixth grade. This variable has a mean (std) of -0.65 (11.59). The dependent variable in columns 5-7 of Panel B is the change in the predicted raw scores in fall of the sixth grade and the raw scores in spring of the sixth grade. This variable has a mean (std) of 9.06 (11.75).

Table 12: Seasonal regressions, using the combined math-reading test

Dependent variable:	Summer, test-score change (raw scores on test part e)	School-year, test-score change (raw scores on test part e)
	OLS	OLS
	(1)	(2)
Girl	0.07	-0.10
	(0.34)	(0.28)
Non-Swedish parents	-1.04	-0.03
	(0.48)	(0.40)
Socioeconomic index (SES)	-0.18	0.05
	(0.20)	(0.17)
R^2	0.010	0.001

Notes: Number of observations is 556. Standard errors (in parentheses) are adjusted for blocks-clustering. The dependent variable in column 1 is the change in raw scores on test part e between spring of the fifth grade and fall of the sixth grade. This variable has a mean (std) of 0.23 (3.65). The dependent variable in column 2 is the change in raw scores on test part e between spring of the fifth grade and fall of the sixth grade. This variable has a mean (std) of 1.06 (3.25). When scores on test part e were missing values are out-of-sample predicted. This were done by first regressing the level of score on test part e on Girl, Non-Swedish parents and the socioeconomic index for the pupils with existing scores. Then, the values on the estimates where used to predict scores on test part e for the pupils that lacked scores on this test part. This out-of-sample prediction was done separately for the test scores on the spring of the fifth grade, the fall of the sixth grade and the spring of the sixth grade.

Table 13: Descriptive statistics, Gothenburg data

	Mean	St. Dev	Min	Max
1948 cohort				
Log(earnings)	12.02	0.59	6.40	14.42
Standardized math test scores grade 6, in standard deviations	0.00	1	-2.98	2.95
IQ test grade 6, in average standard deviations	0.00	0.79	1	2.37
Gender	0.50	0.50	0	1
Fathers education (5 levels)	1.22	0.72	0	4
Mothers education (5 levels)	1.17	0.49	0	4
Own education (6 levels)	3.77	1.44	2	7
1953 cohort				
Log(earnings)	11.93	0.62	5.30	14.09
Standardized math tests, grade 6, in percentile ranks.	0.00	1	-2.52	2.58
IQ test, grade 6, in percentile ranks.	-0.00	0.79	-2.52	2.25
Gender	0.51	0.50	0	1
Fathers' education (5 levels)	1.33	0.78	0	4
Mothers' education (5 levels)	1.16	0.60	0	4
Own education (6 levels)	3.82	1.39	2	7

Notes: Number of observations is 7760 for 1948 cohort and 6687 for 1953 cohort. IQ test is the average of the standard deviations of the scores in three tests regarding opposite, number series and metal folding. Log(earnings) is the logarithm of yearly earnings for all individuals with more than SEK 100 in annual earnings. Earnings are calculated as the sum of earnings from employment and the earnings from own company. Mothers' and fathers' education levels are junior secondary, elementary, unknown, upper secondary, and post-secondary school. Own education levels are <=9 years, upper secondary <= 2 years, upper secondary >2 years, post secondary <= 2 years, post secondary >2 years, and doctoral degree. Observe that descriptive statistics for this earnings variable are not comparable to the family income variable in Table 1.

Table 14: Regressions of earnings on test-scores

	Dependent variable Log(earnings)					
	(1)	(2)	(3)	(4)	(5)	(6)
A. 1948 cohort						
Math test score (in standard deviations)	0.13 (0.01)	0.13 (0.01)	0.11 (0.01)	0.11 (0.01)	0.08 (0.01)	0.07 (0.01)
Math test score squared		0.004 (0.006)		0.004 (0.006)		
IQ test (in average standard deviations)			0.05 (0.01)	0.05 (0.01)		0.02 (0.01)
Controls for gender, mothers' and fathers' education	Yes	Yes	Yes	Yes	Yes	Yes
Controls for own education	No	No	No	No	Yes	Yes
R^2	0.178	0.178	0.180	0.180	0.216	0.217
B. 1953 cohort						
Math test score (in standard deviations)	0.09 (0.01)	0.08 (0.01)	0.06 (0.01)	0.06 (0.01)	0.03 (0.01)	0.03 (0.01)
Math test score squared		0.02 (0.01)		0.02 (0.01)		
IQ test (in average standard deviations)			0.04 (0.01)	0.04 (0.01)		0.01 (0.01)
Controls for gender, mothers' and fathers' education	Yes	Yes	Yes	Yes	Yes	Yes
Controls for own education	No	No	No	No	Yes	Yes
R^2	0.150	0.151	0.151	0.152	0.186	0.185

Notes: Number of observations is 7760 for 1948 cohort and 6687 for 1953 cohort. IQ and math test scores are in standard deviations. Robust standard errors are in parentheses. Controls are four dummies for mothers' education, four dummies for fathers' education and a dummy for gender, and five dummies for own education.

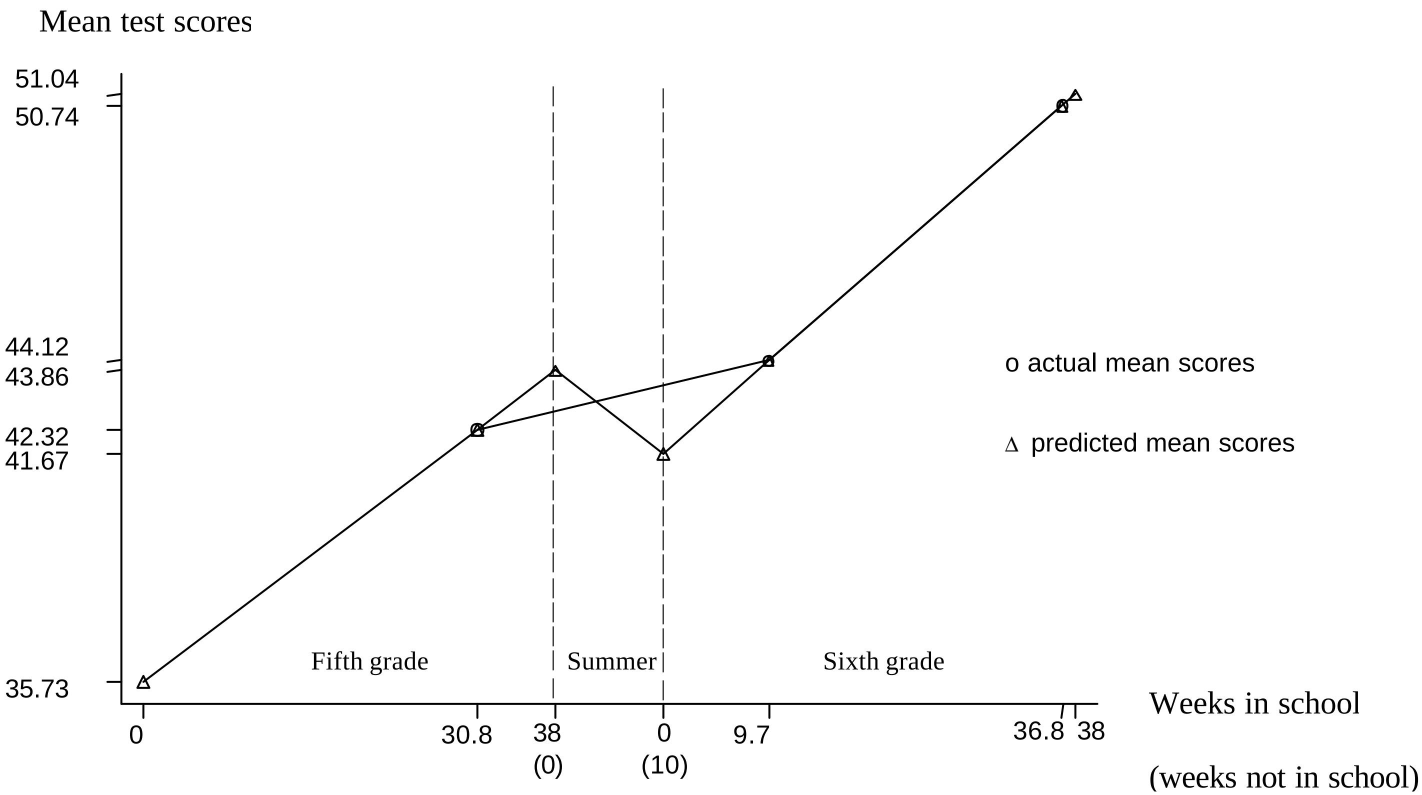

Figure 1: Learning profiles for the mean pupil