

IZA DP No. 2434

Expanding Schooling Opportunities for 4-Year-Olds

Edwin Leuven
Mikael Lindahl
Hessel Oosterbeek
Dinand Webbink

November 2006

Expanding Schooling Opportunities for 4-Year-Olds

Edwin Leuven

*University of Amsterdam, SCHOLAR,
Tinbergen Institute and IZA Bonn*

Mikael Lindahl

*SOFI, Stockholm University
and IZA Bonn*

Hessel Oosterbeek

*University of Amsterdam, SCHOLAR
and Tinbergen Institute*

Dinand Webbink

CPB Netherlands Bureau for Economic Policy Analysis

Discussion Paper No. 2434
November 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Expanding Schooling Opportunities for 4-Year-Olds^{*}

This study presents quasi-experimental estimates of the effect of expanding early schooling enrollment possibilities on early achievement. It exploits two features of the school system in Holland. The first is rolling admissions; children are allowed start school immediately after their 4th birthday instead of at the beginning of the school year. The second is that children having their birthday before, during and after the summer holiday are placed in the same class. These features generate sufficient exogenous variation in children's maximum length of schooling to identify its effects on test scores. Making available one additional month of time in school increases language scores of disadvantaged pupils by 0.06 of a standard deviation and their math scores by 0.05 of a standard deviation. For non-disadvantaged pupils we find no effect.

JEL Classification: I21, I28, J24

Keywords: early childhood intervention, early test scores, early schooling, achievement, policy, identification

Corresponding author:

Edwin Leuven
Amsterdam School of Economics
University of Amsterdam
Roetersstraat 11
Room 10.06, Building E
1018 WB Amsterdam
The Netherlands
E-mail: e.leuven@uva.nl

^{*} We acknowledge comments from seminar participants in Amsterdam, Padova, Paris and Stockholm.

1 Introduction

Impact estimates of early childhood interventions are of particular interest as it has been argued that early interventions have dynamic complementarities in human capital production; early learning makes subsequent learning easier (Heckman, 1999). Most evidence on the effectiveness of early interventions comes from targeted early childhood programs in the US.¹ This paper investigates whether these results carry over to the more general setting of increasing the amount of regular education for young children.

Many advanced countries provide free high-quality education for 4-year-olds.² This is also the case for the Netherlands where children can start school when they are 4 years old, and where compulsory schooling begins when children turn five. The program consists of 24 hours per week during 41 weeks per year. The staff in front of the classroom are certified primary school teachers and are paid according to standard primary school wage scales. The curriculum consists of structured learning activities, and typically children will have started to read and write by the age of six. Over the period 1995-2001 the expenditure per pupil was about 3,500 Euro. To compare, in the United States state spending on public prekindergarten programs was \$3,551 per child in 2004/2005 (Barnett et al., 2005).

We estimate the impact of expanding enrollment opportunities (i.e. the intention to treat) around age 4 on subsequent achievement at age 6. For identification we exploit institutional features of the Dutch early schooling system that generate arguably exogenous variation in the time young children can spend in school. The first feature is that children are allowed to start school immediately after their fourth birthday as opposed to the beginning of the school year as most other countries. The second feature is that

¹Currie (2001) summarizes the beneficial short-term and long-term effects of small-scale intensive interventions such as the Perry Preschool project, the Chicago Child-Parents Centers and the Carolina Abecedarian Project. Large-scale targeted programs with lower per pupil expenditures also appear to have positive short-term and long-term effects. Garces et al. (2002) report beneficial effects on various later outcomes from participation in Head Start.

²Some countries, like France for example, provide state funded preschooling for 3-year-olds.

children having their birthday before, during and after the summer holiday are placed in the same class. Conditional on age these features generate a difference of up to 11 weeks in the time children can spend in school, which constitutes about 15 percent of the amount of schooling they may have had at the moment of the tests. The key identifying assumption is that birth patterns are independent of other factors that affect test scores. We present evidence that birth patterns are unrelated to observed background characteristics, suggesting that the identifying assumption is satisfied.

Our main finding is that making available one additional month of time in school increases language scores of disadvantaged pupils by 0.06 of a standard deviation and their math scores by 0.05 of a standard deviation. For non-disadvantaged pupils we find no effect.

In the next section we briefly review the related literature and discuss how these analyses relate to our analysis. Section 3 documents the details of the Dutch regulations regarding school enrollment age and the scheduling of holiday periods and describes how we use this in our estimation framework. Section 4 describes the data and presents descriptive statistics. Section 5 presents the estimated effects of increasing early school availability on achievement. Section 6 concludes.

2 Related literature

This study is most closely related to the literature that studies the impact of school entry, and schooling at young ages. To put our contribution into perspective it is useful to assume that a person's amount of human capital is an additive function of two components.³ The first component is solely determined by calendar age (age). The second component relates to human capital accumulation while in school, which is assumed to depend on the amount of time spent in school (s) and the age at school entry (age_e). A

³The separability is representative of the empirical literature reviewed here.

general specification capturing this is the following:⁴

$$Y(\textit{age}, s, \textit{age}_e) = Y^0(\textit{age}) + Y^1(s, \textit{age}_e)$$

where $Y^0(\cdot)$ captures the effect of chronological age. For young people it seems reasonable to assume this function to be increasing and possibly concave (after a certain age this function may however be decreasing). The amount of human capital that has been accumulated in school, $Y^1(\cdot)$, obviously depends on the amount of time spent in school, s . In addition it may depend on the age at which pupils enroll in school, measured by their age at school entry, \textit{age}_e . Various papers argue that the effectiveness of time in school depends on age at school entry, for instance because a pupil's span of attention varies with age or teachers' instruction is aimed at mean age in the class.

When estimating the impact of these components on outcomes, one important complication arises from the fact that age, age at entry and length of schooling are linearly dependent: $\textit{age} = \textit{age}_e + s$. The second complication is that age at entry and length of schooling are choice variables, and this endogeneity needs to be accounted for in the analysis. The key challenge is thus to identify sources of exogenous variation that affect one of the determinants of Y without affecting at the same time one of the other determinants. How do previous papers deal with this?

Bedard and Dhuey (2006); Datar (2006); Strøm (2004) use data from school systems characterized by one annual school entry date. Moreover these papers use data from pupils placed in the same grade level and tested at the same date. By implication, pupils will all have identical amounts of time in school, meaning that there is no variation in s . Consequently, these studies estimate functions of the following form:

$$Y(\textit{age}, \textit{age}_e|s) = Y^0(\textit{age}) + Y^1(\textit{age}_e|s)$$

and find that older children perform better. The fundamental problem faced

⁴This parsimonious model omits student, school and family characteristics in order to highlight the identification issues.

by these studies is that with the data at hand, a and a_e are perfectly correlated. Hence, it is impossible to disentangle the effects from these two variables. Datar (2006) formulates this explicitly where she states that "the entrance age effect [age_e] is not identified separately from the age effect [age] when we are interested in outcomes of school-age children". Both Datar and Strøm interpret their results in terms of differences in school entry age, whereas Bedard and Dhuey frame their interpretation in terms of age. Note that testing pupils with different birth dates at different dates in order to keep age at test date constant, does not help identification because this results in a perfect (negative) correlation between age at school entry (age_e) and amount of time in school (s).

Mayer and Knutson (1999) use data from the US and thus face the same problem as the previous papers. They attempt to disentangle the effects of chronological age (at the test date) and of age at school entry by jointly including a linear age term and quarter of birth dummies. The quarter of birth dummies are interpreted as the effect of age at school entry. Starting school a year younger (but having the same amount of schooling and age at the test date) results in a reading score increase of 0.403 of a standard deviation and a math score increase of 0.261 of a standard deviation (p.92). A potential complication with this approach is that the quarter of birth dummies may pick up non-linear age effects.

Also in Israel, children born in the same calendar year start school at the same day. Cahan and Cohen (1989) collected test score data for over 12,000 pupils in grades 4, 5 and 6 in Israel. Unlike the other studies, which are based on data from a single grade, Cahan and Cohen have test score data from adjacent grades which are tested at the same time using the same test. This allows them to disentangle the effects of age and time in school. Pupils placed in the same grade level have the same amount of schooling but differ in age, whereas pupils born in adjacent months but placed in different grade levels have (almost) the same age but differ in their amounts of schooling. The findings indicate that the effect of an additional year of schooling on test scores is about twice the effect of being one year older. Notice that in this design the effect of age captures both the effect of chronological age and

the effect of age at school entry. Cahan and Cohen discuss their age effects in terms of the effects of chronological age, not in terms of effects of age at school entry.

Recently William T. Gormley and Gayer (2005); William T. Gormley et al. (2005) used the same design as Cahan and Cohen and estimate the impact of Oklahoma’s pre-K program for 4-year-olds in Tulsa on cognitive/knowledge test scores, motor skills and language scores. Attendance increases test scores by approximately 0.4 of a standard deviation. Because there is no full compliance (as in Cahan and Cohen’s analysis for Israel) the analysis recovers impact estimates for the treated.

The next section explains how the timing of school holidays in Dutch primary schools causes variation in the amount of time in school independent of age and allows us to estimate the effect of making (pre)school available at age 4.

3 Background and identification strategy

Dutch primary schools consist of 8 grades covering the age groups of 4 to 12-year-old children. While in most countries children typically enter primary school at the same date, in the Netherlands the rule is that children are *allowed* to enroll in primary school the first school day after their 4th birthday, while enrollment is *compulsory* from the first school day of the month after the child reached the age of 5 onwards. About 98 percent of the children start school before their 5th birthday. When exactly between their 4th and 5th birthday a child actually enrolls is up to the parents. The total number of schooldays a child has attended at a given date is therefore to some degree a choice variable. The rule that enrollment is permitted at age 4 and compulsory at age 5 determines the maximum and minimum amounts of time a child can spend in primary school.

The second important feature where our identification builds on, is that a school year cohort in the Netherlands consists of everyone born between October 1 of a given year and September 30 of the next year. At the same time a school year runs from summer holiday to summer holiday. The formal

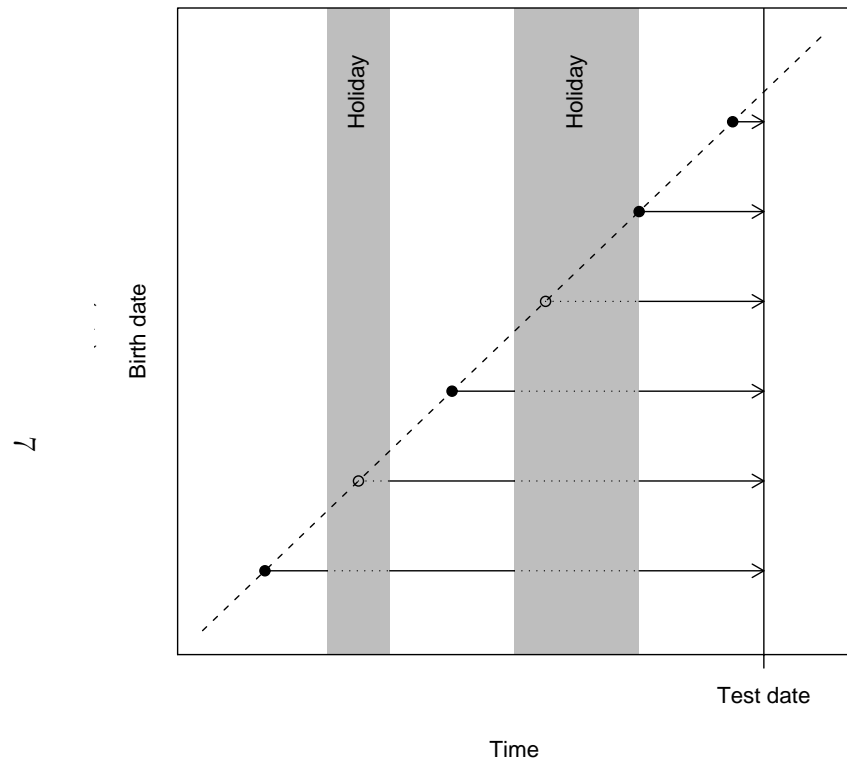
rule is that a child who enrolls in school on the first school day after its 4th birthday spends the period until October 1 in grade 1. Then it spends the period from October 1 until the (next) summer holiday again in grade 1. After the summer holiday the child continues in grade 2. Together these features produce variation in the amount of time a child can spend (maximum length of schooling) in school conditional on age.

Panel (a) of Figure 1 provides a stylized illustration of how these institutions generate maximum length of schooling.⁵ Length of schooling at the time of testing is at its maximum when enrollment takes place at the earliest possible moment which is at age four, i.e. $age_e = 4$. For any given birth date this is then the length of the segment between entry at age four and the moment of the test. However, children are out of school during holidays and we therefore need to subtract the holidays from this schooling spell. This amounts to adding up the solid segments in panel (a). Doing this with the actual school holiday data for every birth date (for a given cohort) leads to a relationship between a child's birthday and its maximum length if enrollment in school as shown in panel (b) in Figure 1.⁶ The flat segments correspond to holiday periods and the downward sloping segments represent school periods. For children having their fourth birthday on the same downward sloping segment, maximum length of schooling in school varies one-to-one with age; being one day older adds one day to the maximum length of schooling. Differences in test scores between two otherwise identical pupils within a segment are attributable to their difference in age as well as to their difference in maximum length of schooling. Maximum length of schooling does not vary across children having their fourth birthday on the same flat segment. Consequently, differences in test scores between two otherwise identical pupils from these segments are solely attributable to differences in their age.

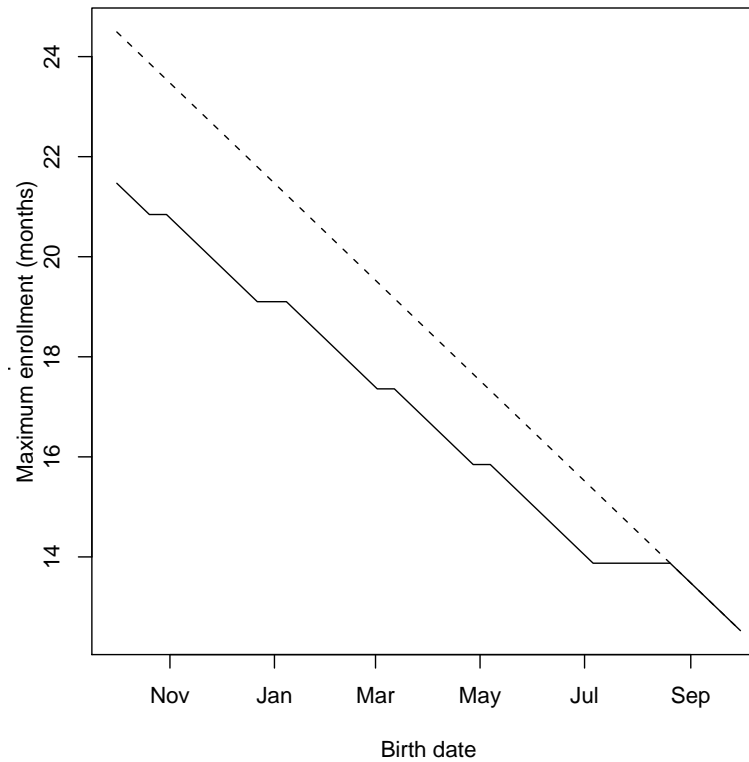
Children in the after-summer-group can have, conditional on their age,

⁵This figure abstracts from weekends.

⁶In practice the exact timing of the summer holidays varies somewhat from year to year and is different for three different regions (North, Middle and South). In all cases, however, the summer holiday ends well before October 1, and hence there are always children at a grade level who have their 4th birthday between the end of the summer holiday and October 1.



(a) Timing of enrollment and max. length of schooling



(b) Relation between birth date and max. length of schooling for a given cohort

Figure 1: Illustration of identification strategy

at most eleven weeks more time in school than children born in other periods. This is most easily seen when we compare the extrapolation of the line segment for the after-summer-group (the dashed line in Figure 1 which ignores holidays) with the solid line segments for the before-summer-group. The vertical distance between these two lines is the difference in maximum length of schooling given age.

Maximum length of schooling at the test date is by definition equal to the difference between age at the test date, age at school entry and the length of holiday between entry and test: $s^{max} = age - 4 - h$. Conditional on age, the amount of holidays between age 4 and the moment of the test (h), breaks the perfect correlation between maximum length of schooling and chronological age.

Since the identification in our analysis comes from a non-linearity there are limits to the extent to which we can control for age. Given that we only consider children within the same cohort this might not be an important problem. We show below that moving from a linear to a quadratic age specification does not affect our results, higher order terms however capture the non-linearity we are exploiting and our identification breaks down.

We will estimate the following model

$$Y = Y(s^{max}, age)$$

From the figure it is clear that we use cross-sectional variation within the entire school year cohort. Since the variation we exploit is conditional on age, we need to control sufficiently flexible for the effect of differences in age on test scores. Once we do this any remaining differences in test scores between children is attributable to differences in maximum length of schooling. In the analysis we will estimate the following specification

$$y_i = \alpha + \beta_{RF} \cdot s_i^{max} + \delta \cdot age_i + \lambda \cdot age_i^2 + x_i' \gamma + \epsilon_i \quad (1)$$

where y_i is the 2nd grade test score, s_i^{max} is the maximum length of schooling for pupil i , x_i are year and regions indicators and their interactions,

and individual characteristics. The identifying assumption is

$$E[s^{max} \cdot \epsilon | age, age^2, x] = 0$$

Some have argued that the timing of births during the year may depend on unobserved characteristics of the parents which have an effect of children's achievement. This point was raised by Bound et al. (1995) in their comment on the use of quarter of birth as an instrumental variable for years of schooling by Angrist and Krueger (1991). For the 18 countries in their dataset Bedard and Dhuey (2006) find no evidence that more educated mothers target birth so that their children are the oldest in the class. In Section 5 we will present similar evidence for our dataset which indicates that there is no systematic relation between parents' levels of education and children's maximum length of schooling.

4 Data and descriptive statistics

4.1 Data

This paper uses data from five waves of the PRIMA survey. This bi-annual survey contains information on Dutch pupils who were enrolled in grades 2, 4, 6 and 8 in the school years 1994/1995, 1996/1997, 1998/1999, 2000/2001 and 2002/2003. Several survey instruments have been used for collection of the data: administrative sources, tests, and questionnaires for teachers, parents and school headmasters. Each wave contains information of about 600 primary schools and around 55,000 pupils which is approximately 10 percent of the relevant age population. The survey design is such that it samples pupils from grades and not from cohorts. This is unfortunate because grade repeating is a fairly common phenomenon in Dutch primary schools and thus introduces substantial selection issues. Only for advancement from grade 1 to grade 2, grade repeating is not an issue.⁷ We therefore restrict

⁷Less than 3 percent of the 2nd graders in the survey are older than 2nd graders should be, indicating that they repeated a grade. The survey contains no information, however, that identifies whether they repeated first or second grade. Three percent is therefore an upper bound on grade repetition from grade 1 to grade 2.

the analysis to 2nd graders who are for the first time in grade 2 and, as a consequence, estimate the short term effect.

The outcome measures we use in the analysis are scores on two cognitive tests; one related to counting, classification and order and one related to the understanding of words, sentences and placement of words in sentences.⁸ These tests were developed for the Dutch government by testing agency CITO in order to measure pupils' readiness for arithmetic and reading, respectively. We will therefore refer to these tests as "arithmetic" and "language". The raw scores on these measures are based on tests which are especially designed for this data collection. From year to year the tests for the same grade levels are identical. The purpose of this is to compare achievement levels over time. Because the scales of the raw scores have no clear meaning, we transformed these scores for each test into wave specific standardized scores, having mean zero and standard deviation one.

In the estimations we control for gender, education levels of father and mother, and two dummies indicating whether the pupil belongs to a disadvantaged group. The Dutch funding scheme for primary schools distinguishes two main groups of disadvantaged pupils: native Dutch pupils with low educated parents and pupils with an ethnic minority background with low educated parents. Schools receive extra funding for pupils from these groups.

4.2 Descriptive statistics

Table 1 shows descriptive statistics for several samples. The average age of the children in 2nd grade at the moment of the test is almost 6 years (70.5 months). On average they could have spent almost 17 months in school at the day of the test. Note that there are no differences between disadvantaged and non-disadvantage children in this respect, as expected if birth patterns are comparable.

Not surprisingly, the parents of disadvantaged pupils have lower levels of education than the non-disadvantaged as can be seen in columns (1) to (3).

⁸Documentation containing the test items can be found here: <http://www.dans.knaw.nl/nl/data/danssearch/advancedsearch?inputall=prima>

Table 1: Descriptive statistics

	Non-	Disadv.		Girls	Boys
	Disadv.	Dutch	Minority		
	(1)	(2)	(3)	(4)	(5)
Age (months)	70.33 (3.40)	70.55 (3.40)	70.63 (3.35)	70.40 (3.39)	70.49 (3.40)
Max. length of Schooling	16.67 (2.58)	16.82 (2.59)	16.89 (2.56)	16.72 (2.57)	16.78 (2.58)
Education Mother					
-Missing	0.09	0.06	0.09	0.08	0.09
-Primary	0.01	0.12	0.53	0.16	0.16
-Lower Secondary	0.16	0.73	0.23	0.30	0.30
-Upper Secondary	0.50	0.08	0.12	0.32	0.32
-Higher	0.23	0.01	0.03	0.13	0.14
Education Father					
-Missing	0.11	0.15	0.18	0.14	0.13
-Primary	0.01	0.10	0.40	0.13	0.12
-Lower Secondary	0.19	0.70	0.27	0.31	0.32
-Upper Secondary	0.41	0.04	0.11	0.26	0.26
-Higher	0.28	0.00	0.05	0.16	0.17
Not disadvantaged	1	0	0	0.54	0.55
Disadv. Dutch	0	1	0	0.21	0.21
Disadv. minority	0	0	1	0.24	0.24
Girl	0.49	0.50	0.50	1	0
Boy	0.51	0.50	0.50	0	1
Language	0.32 (0.95)	-0.04 (0.89)	-0.69 (0.83)	0.10 (1.02)	-0.09 (0.97)
Arithmetic	0.28 (0.98)	-0.14 (0.92)	-0.52 (0.86)	0.04 (1.00)	-0.03 (1.00)
Number of obs	28,942	11,149	12,744	26,129	26,706

The missing values for mother’s education is much lower than for fathers, which could be because mothers are perhaps more likely to have filled in the parent questionnaire. Moreover, mother’s education is missing as often for disadvantaged as for non-disadvantaged children.

The bottom rows of Table 1 presents average test scores for the various subgroups. In column (1) we see that as early as in 2nd grade, non-disadvantaged children score about 1/3 of a standard deviation above average. The difference between non-disadvantaged and disadvantaged minority children is 1 standard deviation on the language test, and a bit less (0.80) on the arithmetic test. Comparing non-disadvantaged to disadvantaged Dutch pupils, we observe a difference on both tests of around 0.40. Girls score on average 0.19 standard deviation higher than boys on the language scores. They also score higher on the arithmetic test, but here the difference (0.07) is small.

5 Results

5.1 Pooled results

Before discussing the estimates of the intention to treat effects that are the focus of this paper, we first examine the exogeneity of maximum length of schooling by regressing this on four dummies for mothers’ education, four dummies for fathers’ education, a gender dummy, and dummies for the two disadvantaged groups, as well as on age, age squared and dummies for years, regions and year-region interactions. Table 2 reports results from regressions of maximum length of schooling on background characteristics controlling for other covariates. These results show that there is no systematic relation between maximum length of schooling and observed background characteristics. As reported in the main text, the F-test for the joint significance of the background variables has a p-value of 0.321. This supports (but does not prove) the identifying assumption that maximum length of schooling is orthogonal to unobserved characteristics related to achievement. This conclusion holds also for the various subsamples. We can therefore reject that

maximum length of schooling varies systematically with these background variables which affect test scores.

Table 3 shows our results for three specifications for the pooled sample. It is important to properly control for pupil's age, since the variation we exploit is conditional on age. We present results for both the language test and the arithmetic test. The standard errors are corrected for clustering at the school level and are heteroscedasticity robust.

We first present the estimate of the regression of the language score on maximum length of schooling with only a linear control for age. The effect of increasing maximum length of schooling by one month equals 0.029 of a standard deviation but is insignificant. Adding a quadratic age term (column 2) increases the point estimate somewhat but the resulting effect remains not significantly different from zero. Subsequently adding additional background variables (column 3) cuts the point estimate in half, but the difference with the estimate obtained without the additional controls is insignificant. When the language score is replaced by the arithmetic score as outcome variable (columns 4 to 6), the results are very similar.

From these results we conclude that, although the point estimate is positive, we do not find statistically significant effects of increasing enrollment opportunities on test scores for the whole population of 2nd graders.

5.2 Results for subgroups

The fact that we do not find significant results for the whole population does not imply that no group benefits from changes in the age at which parents can enroll their children. It has been argued that disadvantaged pupils are likely to benefit more from attending school at a young age because the difference between the school environment and the home environment is probably large. Moreover, gender differences in child development may also lead to a differential impact of the intervention between boys and girls. We report the effects of maximum length of schooling on language and arithmetic scores for different subgroups in Table 4. These estimates control

Table 2: Maximum length of schooling and background characteristics

	All	Non-	Disadv		Girls	Boys
	(1)	Disadv. (2)	Dutch (3)	Minority (4)	(5)	(6)
Education mother (reference category = Missing)						
- Primary	-0.002 (0.005)	0.005 (0.014)	-0.002 (0.011)	-0.005 (0.007)	-0.007 (0.007)	0.003 (0.007)
- Lower secondary	0.004 (0.004)	0.006 (0.007)	-0.003 (0.009)	0.007 (0.008)	0.000 (0.006)	0.009 (0.006)
- Upper secondary	0.004 (0.004)	0.006 (0.007)	0.000 (0.011)	0.002 (0.009)	0.003 (0.006)	0.005 (0.006)
- Higher	0.009 (0.005)	0.011 (0.007)	0.028 (0.022)	0.005 (0.014)	0.004 (0.007)	0.014 (0.007)
Education father (reference category = Missing)						
- Primary	0.002 (0.004)	-0.010 (0.013)	-0.004 (0.009)	0.006 (0.006)	0.000 (0.006)	0.004 (0.006)
- Lower secondary	-0.004 (0.003)	-0.006 (0.006)	-0.004 (0.006)	-0.003 (0.006)	-0.002 (0.005)	-0.007 (0.005)
- Upper secondary	-0.005 (0.004)	-0.007 (0.006)	-0.007 (0.011)	-0.005 (0.008)	-0.002 (0.006)	-0.009 (0.005)
- Higher	-0.006 (0.004)	-0.007 (0.006)	-0.033 (0.030)	-0.007 (0.011)	-0.001 (0.006)	-0.011 (0.006)
Disadv. Dutch	-0.003 (0.003)				-0.004 (0.005)	-0.003 (0.004)
Disadv. minority	-0.003 (0.003)				0.002 (0.004)	-0.009 (0.004)
Girl	0.002 (0.002)	0.003 (0.002)	-0.006 (0.004)	0.004 (0.003)		
N	52,835	28,942	11,149	12,744	26,129	26,706
F-test joint sign.	0.321	0.834	0.819	0.490	0.447	0.243

Table 3: Estimation results - Full sample, various specifications

	Language			Arithmetic		
	(1)	(2)	(3)	(4)	(5)	(6)
Max. length of Schooling	0.029 (0.017)	0.036 (0.021)	0.016 (0.018)	0.027 (0.016)	0.032 (0.020)	0.016 (0.019)
Age	0.030 (0.013)	0.068 (0.065)	0.102 (0.056)	0.041 (0.012)	0.071 (0.065)	0.093 (0.058)
Age sq/100		-0.031 (0.052)	-0.041 (0.045)		-0.024 (0.052)	-0.028 (0.047)
R-squared	0.032	0.032	0.234	0.049	0.049	0.189
Marginal effect age		0.024 (0.016)	0.045 (0.014)		0.036 (0.015)	0.054 (0.014)
Background controls	No	No	Yes	No	No	Yes

Note: All regressions include 4 year dummies, 2 region dummies and their interactions. The background variables are: 4 dummies for mother's education, 4 dummies for father's education, 1 disadvantaged Dutch dummy, 1 disadvantage minority dummy and 1 gender dummy variable. The standard errors are corrected for clustering at the school level and are heteroscedasticity robust. The "Marginal effect age" is the derivative with respect to age of the estimated model, evaluated at the mean age. Standard errors of the marginal effect of age are calculated using the delta-method.

for age, age-squared and background characteristics.⁹

As above we examined the exogeneity of maximum length of schooling by regressing this, for each subgroup, on four dummies for mothers' education, four dummies for fathers' education, a gender dummy, and dummies for the two disadvantaged groups, as well as on age, age squared and dummies for years, regions and year-region interactions. The F-tests for the joint significance of the background controls have p-values ranging from 0.243 for boys to 0.834 for non-disadvantaged Dutch. Also for the subgroups we can therefore reject that maximum length of schooling varies systematically with these background variables which affect test scores.

The top panel of table 4 reports the results for language, the bottom panel for arithmetic. The first to fourth columns report estimates for socio-economic groups: non-disadvantaged pupils, disadvantaged Dutch pupils, disadvantaged minority pupils, and pupils from any disadvantaged group (taking disadvantaged Dutch and disadvantaged minority together). The last two columns report separate estimation results for girls and boys.

For non-disadvantaged pupils, the intervention has no impact. For both outcome variables, the point estimates are even negative. The results in columns (2)-(4) indicate that children from the two disadvantaged groups benefit from expanding their enrollment opportunities. In columns (2) and (3) all four point estimates for the effect of maximum length of schooling are positive, but they lack precision. Only the effect on language for disadvantaged minority pupils is significant at conventional levels. Since the effects for the two disadvantaged groups are of fairly similar magnitudes (equality cannot be rejected), we merge the two groups in order to gain precision. This gives us the results in column (4). For both outcomes the effect of maximum length of schooling is significantly positive (for language at the 2%-level, for arithmetic at the 10%-level). Expanding enrollment opportunities by one month increases performance of disadvantaged pupils on the language test by about 0.06 of a standard deviation, and on the arithmetic test by about 0.05 of a standard deviation. Notice that similarity of the results for two independent outcome measures makes it unlikely that the results can be at-

⁹Excluding the background characteristics gives very similar results.

tributed to randomness.¹⁰

That we find a significant effect for minority pupils on language and not on arithmetic can be attributed to the fact that the two largest minority groups in the Netherlands are from Turkish and Moroccan origin. Many parents of these pupils still use Turkish or Moroccan as language at home.

The findings differ somewhat between boys and girls. The effect of increasing enrollment opportunities is for both outcomes larger for boys than for girls. For language the effects are insignificant, but for arithmetic it appears to be the case that boys benefit from increased enrollment opportunities, whereas girls don't.

As we mentioned, our identification strategy assumes that a quadratic specification of the age-achievement profiles suffices. While this is restrictive it is important to note that the related studies discussed in Section 2, typically include only a linear age term. To further probe the issue of the age controls, we regressed outcome measures on (1) year dummies and other controls but excluding region dummies and the region-year interactions, and on (2) region dummies and other controls but excluding year dummies and region-year interactions. The first specification also uses variation in the timing of holidays between regions to identify the effect of maximum length of schooling (but does not control for region effects), the second specification also exploits variation in the timing of holidays between years to estimate the effect of interest (but does not control for year effects). For all subgroups the estimates from these alternative specifications are very similar to the estimates reported in Table 4, and in no single case could we reject equality of the effects from the alternative specification and from the reported specification. This is reassuring and adds to the credibility of the findings.

To put the results in perspective it is important to understand what the learning environment would have been in the absence of enrollment in school. The exogenous variation in maximum length of schooling is caused by the incidence of up to 11 weeks of school holidays. Six of these 11 weeks are

¹⁰With multiple outcomes and multiple subgroups, it is likely that for at least one combination of outcome and subgroup a significant effect is found, even if the true impact is zero.

in the summer, the other 5 are during the year. Most Dutch families spend 3 to 4 weeks of the summer holiday away from their home (often abroad). The other holiday weeks are typically spend at home. In most cases one of the parents will look after the child(ren) during the holiday weeks that are spend at home. In fewer cases, grandparents or other family will play a role. This implies that around 35 percent of the holiday weeks are spend away from home and the remainder at home with one of the parents or another family-member. This will often also be the situation for children whose school attendance is postponed.

5.3 *The effect of time in school on test scores*

The previous section reports the effect of expanding enrollment opportunities on test scores. Parents can decide to take up the enrollment possibility or they may decide to wait. We have found that expanding enrollment opportunities increases performance of disadvantaged pupils. Given this result one might be interested in the effect of making enrollment compulsory.

To address this question one would need to estimate the following outcome equation using 2SLS

$$y = \alpha + \beta_{IV} \cdot s_i + \delta \cdot age_i + \lambda \cdot age_i^2 + x_i' \gamma + \epsilon_i \quad (2)$$

where s_i is actual length of schooling. Since actual enrollment is endogenous and under the discretion of the parents it is likely to be correlated with unobserved determinants of pupil achievement. If we want a reliable estimate of the causal effect of making enrollment compulsory at an earlier age we would need to find a good instrument; something that affects enrollment yet satisfies the exclusion restriction that it is uncorrelated with unobservables affecting test scores.

We have such an instrument, namely maximum length of schooling. The corresponding first stage equation for $enroll_i$ would then be

$$s_i = \eta + \pi \cdot s_i^{max} + \zeta \cdot age_i + \vartheta \cdot age_i^2 + x_i' \varphi + \omega_i \quad (3)$$

If maximum length of schooling significantly affects actual schooling we

Table 4: Estimation results - Subgroups

	Non- Disadv.	Disadv.			Girls	Boys
	(1)	Dutch	Minority	All	(5)	(6)
<i>Language</i>						
Max. length of schooling	-0.021 (0.027)	0.053 (0.039)	0.066 (0.034)	0.060 (0.026)	-0.004 (0.027)	0.036 (0.024)
Age	0.168 (0.078)	0.057 (0.116)	-0.045 (0.102)	0.008 (0.075)	0.070 (0.082)	0.135 (0.075)
Age sq/100	-0.067 (0.063)	-0.027 (0.092)	0.036 (0.080)	0.002 (0.060)	-0.006 (0.066)	-0.075 (0.059)
R sq	0.082	0.084	0.090	0.195	0.235	0.221
Marginal effect age	0.074 (0.020)	0.019 (0.030)	0.003 (0.026)	0.010 (0.020)	0.061 (0.020)	0.029 (0.018)
<i>Arithmetic</i>						
Max. length of schooling	-0.007 (0.025)	0.065 (0.042)	0.031 (0.037)	0.047 (0.027)	-0.029 (0.026)	0.059 (0.026)
Age	0.192 (0.084)	-0.009 (0.122)	-0.095 (0.105)	-0.049 (0.078)	0.031 (0.083)	0.150 (0.078)
Age sq/100	-0.083 (0.067)	0.018 (0.097)	0.092 (0.086)	0.053 (0.063)	0.041 (0.066)	-0.091 (0.063)
R sq	0.093	0.087	0.086	0.123	0.194	0.183
Marginal effect age	0.075 (0.019)	0.017 (0.032)	0.035 (0.029)	0.026 (0.021)	0.089 (0.020)	0.021 (0.020)
N	28,942	11,149	12,744	23,893	26,129	26,706

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, a quadratic in age and the full set of background variables listed in the note in Table 3. The standard errors are corrected for clustering at the school level and are heteroscedasticity robust. The “Marginal effect age” is the derivative with respect to age of the estimated model, evaluated at the mean age.

could then estimate β_{IV} . Unfortunately the PRIMA survey contains only limited and unreliable information on actual enrollment.¹¹ No questions with regard to actual enrollment have been asked in the 1994 and 2000 and 2002 waves. In 1996 and 1998 parents were asked how old their child was when it entered school. Here parents are supposed to report age in years (4, 5 or 6) and months (0 to 11). Only 40.1 percent of the observations in 1996 and 1998 have non-missing values on both year and month. For the disadvantaged groups, this figure is even worse; 31.2 percent for disadvantaged native pupils and 19.3 percent for disadvantaged minority pupils. Schools have also been asked to report for each pupil the year and the month in which they started to attend school. We thus have two measures of the same variable. Regressing the parents measure on the school measure and vice versa gives the reliability ratios of both measures. The reliability ratio of the parents measure equals 0.62 for all groups together, but reduces to 0.26 for minority pupils. The reliability ratio of the school measure equals 0.23 for all groups together (and 0.27 for minority pupils). The low response rates together with the low reliability ratio's among those who responded make the information on actual enrollment useless for further analysis.¹² For this reason we only reported estimates of the reduced form model and recuperated $\hat{\beta}_{RF}$. As we argued, these results are interesting from a policy point of view in their own right.

Even without information on actual enrollment, we can, however, infer something about β_{IV} . We know that the IV estimate is given by the following expression

$$\hat{\beta}_{IV} = \frac{\hat{\beta}_{RF}}{\hat{\pi}_{FS}} \quad (4)$$

The question then is, what value $\hat{\pi}_{FS}$ has. Suppose we increase enrollment opportunities by 1 month. Those who enroll immediately when they turn 4

¹¹Unfortunately there are also no other sources with information on precise age at which young children start school. For funding purposes schools only have to administer how many pupils attend at October 1st of each year.

¹²As the results from the previous section show, even with data from five waves and many more observations per wave, precision is not overwhelming. Estimating an IV-model with just a fraction of the original number of observations used for the reduced form approach will give very large standard errors.

before the expansion takes place - the constrained group c -, can increase their enrollment by at most 1 month. For this group we can thus infer that $0 \leq \hat{\pi}_{FS,c} \leq 1$. Those who delay enrollment before the expansion - the unconstrained group u -, can increase their enrollment by at most 1 month plus the delay. While possible in principle, this seems an unlikely response because it requires that people respond to loosening a constraint that was not binding for them. A more likely scenario is that those who enter at, say, age 4 years and 3 months when they are allowed to start at age 4 years and 0 months, will also enter at age 4 years and 3 months when they are allowed to start at the age of 3 years and 11 months. For this group we then would have $\hat{\pi}_{FS,u} = 0$. According to the responses on actual enrollment in the parents' questionnaires (the least unreliable source), 68 percent of the pupils belong to the constrained group. For Dutch disadvantaged pupils this percentage equals 65 and for disadvantaged minority pupils 43.

Under the behavioral assumptions implicit in the above discussion, the overall estimate of $\hat{\pi}_{FS}$ is thus a weighted average of 0 and a value between 0 and 1, so that we have that $0 \leq \hat{\pi}_{FS} \leq 1$. It then follows from (4) that

$$\hat{\beta}_{IV} \geq \hat{\beta}_{RF}$$

and our reduced form estimate is a lower bound on the effect of making schooling compulsory at a lower age.

6 Conclusions

This study introduced a novel way to estimate the effect of expanding enrollment opportunities on test scores and identify this separately from the age effect. This was possible due to the specific feature of the Dutch schooling system that allows children to start school when they turn 4. Together with the incidence of school holidays and the fact that a school year cohort children born between October 1 and September 30 of the next year, this generates exogenous variation in enrollment opportunities conditional on age.

For disadvantaged pupils we find that increasing enrollment opportuni-

ties by one month increases language scores on average by 0.06 standard deviation and arithmetic scores on average by 0.05 standard deviation. Non-disadvantaged pupils do not benefit in test scores from expanded enrollment opportunities. This suggests that at age 4 school and home environment are close substitutes in the production of achievement for non-disadvantaged children, whereas for disadvantaged children school provides better learning opportunities than the home environment.

Although these effects are reduced from effects and as such do not estimate the causal effect of enrollment, we argue that they are lower bounds of the effects of making enrollment in primary education compulsory at a younger age.

The test scores are measured around two years later and the effects we measure are therefore relatively short-term effects. Yet, as the results of Garces et al. (2002) show, even if intervention effects on test scores fade out over time there may be long-term effects on other outcome variables.

The 0.05-0.06 standard deviation increase in test scores reported here come at a cost of (depending on the type of disadvantaged pupil) 354 to 541 euro per pupil. This compares favorably to the costs and effects of Head Start. Currie and Thomas (1995) report an effect of Head Start participation on early test scores of 0.203 of a standard deviation for disadvantaged white children. For Afro-American children they find no significant effects. Participation in Head Start costs approximately \$3,500 per child per year. Increasing opportunities to enroll into primary school at younger ages are therefore an interesting policy alternative to targeted programs such as Head Start.

References

- Angrist, J. D. and Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics*, 106(4):979–1014.
- Barnett, W. S., Hustedt, J. T., Robin, K. B., and Schulman, K. L. (2005).

- The State of Preschool: 2005 State Preschool Yearbook.* The National Institute for Early Education Research, New Brunswick, NJ.
- Bedard, K. and Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *Quarterly Journal of Economics*, forthcoming.
- Bound, J., Jaeger, D. A., and Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430):443–450.
- Cahan and Cohen (1989). Age versus schooling effects on intelligence development. *Child Development*, 60:1239–1249.
- Currie, J. (2001). Early childhood interventions. *Journal of Economic Perspectives*, 15(2):213–238.
- Currie, J. and Thomas, D. (1995). Does Head Start make a difference? *American Economic Review*, 85(3):341–364.
- Datar, A. (2006). Does delaying kindergarten entrance give children a head start? *Economics of Education Review*, 25:43–62.
- Garces, E., Thomas, D., and Currie, J. (2002). Longer-term effects of Head Start. *American Economic Review*, 92:999–1012.
- Heckman, J. J. (1999). Policies to foster human capital. Working Paper 7288, NBER.
- Leuven, E., Lindahl, M., Oosterbeek, H., and Webbink, D. (2004). New evidence on the effect of time in school on achievement. Scholar wp 47/04, Department of Economics, University of Amsterdam.
- Mayer, S. E. and Knutson, D. (1999). Does the timing of school affect how much children learn? In Mayer, S. E. and Peterson, P. E., editors, *Earning and Learning: How School Matters*, pages 79–102. Brookings Institution and Russell Sage Foundation.

Strøm, B. (2004). Student achievement and birthday effects. Unpublished manuscript, Norwegian University of Science and Technology.

William T. Gormley, J. and Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-K program. *Journal of Human Resources*, 60:533–558.

William T. Gormley, J., Gayer, T., Phillips, D., and Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology*, 41(6):872–884.