IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11807

The Timing of Instruction Time: Accumulated Hours, Timing and Pupil Achievement

Paul Bingley
Eskil Heinesen
Karl Fritjof Krassel
Nicolai Kristensen

SEPTEMBER 2018

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11807

The Timing of Instruction Time: Accumulated Hours, Timing and Pupil Achievement

Paul Bingley
VIVE

Nicolai Kristensen
VIVE and IZA

Eskil Heinesen
The ROCKWOOL Foundation Research Unit

Karl Fritjof Krassel
VIVE

SEPTEMBER 2018

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Timing of Instruction Time: Accumulated Hours, Timing and Pupil Achievement[1]

Instruction time varies among schools, subjects, pupils and grades. This variation is positively associated with test scores and has been used to identify modest positive causal effects for instruction hours in certain grades. We exploit administrative data on delivered and timetabled instruction time in each grade throughout compulsory school for three full cohorts of Danish children, and find positive marginal hours effects on 9th grade test scores using accumulated time which are twice as large as when using 9th grade time only. Effects are largest for low SES households and for boys with non-western immigrant background, especially in the early grades.

JEL Classification: C21, I2, I21

Keywords: school resources, test scores, time in education

Corresponding author:
Nicolai Kristensen
VIVE - The Danish Centre of Applied Social Science
Herluf Trolles Gade 11
1052 Copenhagen K
Denmark
E-mail: nikr@vive.dk

1 We thank the National Agency for IT and Learning for providing the data on planned instruction hours. We thank Maria Humlum, Lars Skipper and seminar participants at the third CEN workshop for helpful comments.

1. Introduction

Education is a key determinant of inequality (Katz & Autor 1999), economic growth (Goldin & Katz 2009), individual earnings (Harmon & Walker 1995) and other life trajectories (Oreopoulos & Salvanes 2011)[2]. In most societies, education is publicly provided and constitutes one of the largest government investments; the average OECD expenditure on primary and lower secondary education was 2.5 per cent of GDP in 2012.[3] Improving the efficiency of educational resource allocation could have large societal impacts, and economists have made progress in identifying the causal relationships that drive educational outcomes (Hanuschek, 2006). In this paper, for the first time, we identify how instruction time throughout compulsory school affects test scores.

Formal education around the world differs greatly with regard to the amount of classroom instruction pupils receive, how much teaching time is devoted to various subjects and the ages at which children receive schooling[4]. This variation reflects different beliefs as to how much time learning requires and different preferences as to what should be taught and when. The more pupils are exposed to school resources, the more impact those resources might have on pupil outcomes; there is a positive association between instruction time and test scores, and many studies have gone further by exploiting exogenous sources of variation in instruction hours for selected grades to find modest positive causal effects on test scores[5]. We contribute to this literature by using Danish administrative data providing exogenous variation in instruction hours for each grade throughout compulsory school, enabling us to estimate the effects of total instruction time and the effects of age of instruction.

In order to identify the effect of instruction time on achievement, it is essential to overcome selection issues that might otherwise bias estimates. For example, if pupils are sorted into schools by aptitude and specialisation, estimates would be upward biased; whereas compensatory teaching for under-achievers would downward bias estimates of instruction time effects. In the absence of experimental evidence, researchers have used a variety of quasi-experimental empirical strategies – differences-in-differences, cohort comparisons, natural experiments and fixed effects – in order to overcome identification issues.

---

[2] Such as health (Clark & Royer, 2013), intergenerational mobility (Currie & Moretti, 2003), teenage pregnancy (Black et al., 2008), marriage (Lefgren & McIntyre 2006), trust and civic participation (Milligan et al., 2004) and criminal activity (Moretti & Lochner, 2004).

[3] OECD (2015).

[4] Compulsory instruction time in general education varies greatly around the world. On average, children spend 7,500 hours under instruction, the national hours of instruction ranging from 5,700 hours in Hungary to 11,000 hours in Australia (OECD, 2016). In our sample, mean total instruction time is 7,500 hours.

[5] See Hanushek (2015) for an introduction to an Economic Journal feature on time in education, and references therein.

Card and Krueger (1992) were the first to estimate the causal effect of instruction time on outcomes. Using variation in length of the school year between US states, which also varies over time, they found that instruction time increases subsequent earnings. However, in this differences-in-differences (DiD) framework effects become insignificant after controlling for other school characteristics. Pischke (2007) used the short West-German school year of 1966-7 in a cohort comparison design, and found reduced educational attainment.[6] Marcotte (2007) and Marcotte & Hemelt (2008) exploited the natural experiment of snow-induced school closure days in Maryland, and found reduced test scores. Finally, Lavy (2015), with the approach that is closest to ours, used variation in instruction time in 50 countries from PISA 2006, estimating within-pupil between-subject fixed effects and found significantly positive effects on test scores.[7] [8]

There are two studies of instruction time effects in Denmark; an experimental study by Andersen and co-authors (2016) and an observational study by Jensen (2014). Andersen and co-authors (2016) report findings from a 2013 randomised trial of Danish language instruction time in 4th grade. A 15 per cent increase in instruction time for 16 weeks caused a 0.15 standard deviation increase in Danish reading test scores measured shortly afterwards. Jensen (2014) uses a 2003 instruction time reform in a difference-in-differences framework, finding a 2 per cent increase in maths instruction time throughout 9th grade caused a 0.21 standard deviation increase in maths test scores at the end of grade 9. However, the 3 per cent increase in Danish language instruction time had no effect on Danish test scores.

The common feature of the Danish and international studies is that instruction time is only observed for some grades. In contrast to these partial or snapshot measures, we observe instruction time throughout the schooling career. These comprehensive measures enable us to make three main contributions: estimation of effects of the timing of instruction time, effects of complementarities between instruction times in different grades, and the effect of accumulated instruction time.

In addition to these main contributions, our data also allow for two data-based tests of the existing literature. Firstly, while the focus of the literature has been on creative methods of identification less attention has been paid to interpretation of inferences based on different measures of hours. Several different measures

---

[6] Parinduri (2014) examined the longer school year in Indonesia in 1978-79, and finds, in line with Pischke (2007), increased educational attainment.

[7] Rivkin & Schiman (2015) used PISA 2009 data to compare between grades within school and subject, yielding findings similar to those of Lavy (2015).

[8] The studies referenced in the main text exemplify the range of identification strategies that have been used. Two recent studies used a reform which reduced the length of the academic track of German schools after grade 4 from 9 to 8 years, while compensating for this with a proportional increase in instruction time in the remaining 8 years to cover a similar curriculum. For grades 5-9, the reform increased average weekly instruction time by about 2 hours, and at the same time increased curriculum demands for these years. Using DID analysis, Huebener et al. (2017) found that the reform increased PISA test scores in grade 9, but significantly more so for high-performing pupils. Dahmann (2017) found that the reform had a positive effect on crystallised intelligence at age 17 (for boys only), but no effect at the end of the academic track.

of instruction time have been considered in the literature – nationally or regionally mandated, school planned, teacher taught, pupil received – and interpretation of effects depends upon the nature of the measure, and who is being asked to report. While mandates at the national or regional level provide plausibly exogenous variation, the extent to which schools and teachers comply or pupils attend is open to question. Similarly, instruction time timetabled by schools might not be delivered by teachers.

Both mandated and timetabled instruction time allows estimation of intention to treat (ITT) effects. Pupil reports of instruction time received are the most direct measure but may give rise to biased estimates because of truancy, if test score differences reflect pupil absence from lessons. We use school timetabled instruction as our primary measure of instruction time, and for two years we combine timetabled hours with school reports of delivered instruction time, finding a reliability ratio of 0.91 – which implies a modest up-scaling of our timetabled hours ITT estimates – to obtain an average treatment effect (ATE).

Secondly, in Danish administrative registers we observe timetabled hours at the school level in each of three subjects throughout compulsory schooling for three cohorts of pupils, together with test scores in these subjects for all pupils at the end of compulsory schooling. These data allow us to follow the approach of Lavy (2015), using within-pupil between-subject differences and thereby removing pupil-specific and school-specific factors, which might otherwise correlate with instruction time and bias estimates. While Lavy (2015) considered instruction time in the 9th grade only, we take a more expansive view, using instruction time in *each* grade throughout compulsory school. Our data allow for a test of the underlying assumption of the absence of correlation between lagged instruction time and the error term.

Our findings relate to existing grade-specific studies, which we are able to nest within accumulated instruction hours measures. New to the literature, we can identify effects of instruction in specific grades and test for complementarities between early and late instruction. Finally, we can consider heterogeneity in instruction timing effects by gender, ethnicity and socio-economic status (SES).

Compared to previous studies, which have used 9th grade instruction hours only, we find that all-grade effects are twice as large. Because hours are correlated grade-to-grade, 9th grade hours *per se* contribute 1/3, and 1st-8th grades contribute 2/3 of the effect of accumulated instruction hours on 9th grade test scores.

Beyond this reinterpretation of 9th grade hours effects, we find that the timing of instruction hours between grades matters. Girls benefit significantly from instruction time in primary (grades 1-3) and lower secondary school (grades 7-9), whereas boys benefit from instruction time in junior school (grades 4-6). Testing interactions of instruction time in different grades, we find no evidence of complementarities; earlier instruction time does not significantly enhance the effect of later instruction time.

Accumulated instruction hours effects have a steep SES gradient, with one more hour of instruction per week increasing low (high) SES scores by 0.090 (0.051) standard deviations. These SES gradients in hours effects become steeper for higher grades, the difference in high-low SES effects in primary (lower secondary) school being 0.022 (0.053) standard deviations. For pupils with Non-Western immigrant background, instruction hours have stronger effects on test scores than for natives, especially for boys, and they are driven by early grade effects for both genders.

Overall, instruction timing matters, with hours in later grades having the largest effects on 9th grade test scores. However, heterogeneous effects show that the largest of our effects stems from boys with immigrant background in primary school, and that the SES gradient in instruction hours effects increases with increasing grades. While we found no evidence of instruction time complementarity between grades, we see instruction time effects differing according to pupil endowments. These heterogeneous effects of instruction time between groups and between grades offer the prospect of a re-allocation of instruction time, which would improve test scores without increasing total hours taught.

The remainder of the paper is organised as follows. In Section 2, the institutional details are outlined, followed by a data description in Section 3. The empirical model and the identification strategy are presented in Section 4, while Section 5 includes estimation results. Discussion and perspectives are given in Section 6, and, finally, conclusions and suggestions for further research are provided in Section 7.

2. Institutional setting

Compulsory schooling in Denmark runs from the 1st through 9th grade (age 7-15).[9] Public schools are free, and expenditure on public schools is financed by municipal taxes, primarily income tax, but extensive grants and equalisation schemes eliminate the greater part of the financial inequalities among the 98

[9] Prior to 1st grade, parents could previously choose to enroll their child into grade K (kindergarten). Enrolment was very high (about 99%). From 2009, grade K became mandatory.

municipalities.[10] Private schools are heavily subsidised; parents pay a tuition fee covering about 15% of the costs, on average.

Almost 90% of children in the 1st grade are enrolled in public schools. However, a few per cent of children who start in public schools change to private schools later, and in the 9th grade some pupils change to private boarding schools. For the three cohorts analysed in this paper, 70% attended public schools in the 9th grade, 13% attended boarding schools, and 17% attended other private schools in the 9th grade.

The number of instruction hours by subject and grade are regulated to a certain extent at the national level by minimum and guideline hours provided by the Ministry of Education. A reform implemented from the school year 2003/2004 tightened this regulation by increasing classroom hour minima and introducing more subject-specific regulation. This reform, which reduced the variation in instruction time among schools, was utilised in Jensen (2014) to estimate effects of instruction time in the 9th grade on 9th grade test scores. A major reform was implemented from the school year 2014/2015, which increased instruction time considerably in all grades. Between these two reforms, only minor changes were made in minimum and guideline hours of instruction. It is the variation in instruction time among schools and over time during this "quiet" period that is used in this paper.

Even after the 2003 reform, municipalities had a large degree of freedom to decide upon levels and allocation of financial resources for public schools, and typically individual schools also had discretion to allocate resources to specific grades and subjects. Thus, there is considerable variation in the number of instruction hours across municipalities, and across schools within municipalities. The large variation applies for total hours from 1st to 9th grade and for hours by grade and subject. At the end of compulsory schooling (9th grade), there are mandatory exams for all pupils in four subjects: Danish, Maths, English and Physics/Chemistry.[11] In this paper, we focus on the relationship between test scores (exam marks) and instruction hours in three subjects: Danish, Maths and English. We do not include the subject Physics/Chemistry in the analysis, since this is taught only from the 7th grade onwards and our focus is on effects of instruction hours throughout the whole school career.[12] We show that our main results are robust to including this extra subject in the analysis. For each pupil, the test scores in Danish used in this paper are an average of three individual test scores (for

[10] There were 271 municipalities from 2003-2006, and 275 prior to 2003.

[11] All pupils are to take two additional exams; one in Biology or Geography and one in History, Religion or Social studies (or an extra exam in English). The Ministry of Education decides upon the distribution of these exams among classes. There is also an examination in a second foreign language (German or French), which is taught from the 7th grade onwards, but many academically weak pupils are exempted from attending classes in German/French, and more than 10% of the pupils attending classes do not take the exam.

[12] In grades 1-6, pupils are taught a subject called "Nature and technology", which is weakly related to Physics/Chemistry.

reading, spelling and writing an essay). Maths test scores are an average of two individual test scores, while, in English, there is a single oral examination. [13]

3. Data and descriptive statistics

Our main explanatory variable is based on data for planned instruction hours at the school-subject-grade-year level for public schools, obtained from the Ministry of Education[14] for the school years 2003/2004-2013/2014. For three cohorts, we have data on planned instruction hours from school-start in 1st grade (in 2003-2005) to the end of 9th grade (in 2012-2014).[15] The relationship between cohorts, school years and grades is shown in Appendix Figure A.1. Planned instruction hours are registered before the commencement of the school year. They may deviate from the number of hours actually taught, due to teacher sickness or in-service training, for instance. For two years we also observe school reports of delivered instruction time.

These data on instruction hours are linked to individual pupils, using administrative registers from Statistics Denmark. Thus, for each pupil we have information on the grade attended in each school year and, from 2007/2008 also on the school attended. Moreover, through unique personal identification numbers, we can link up to information on personal characteristics (including gender, age, ethnicity and test scores in 9th grade), parental characteristics (including education, earnings and labour market status), family status and place of residence. Our initial sample consisted of 166,771 pupils who were in the population register each year until they completed 9th grade in 2012-2014. From this sample, we omitted pupils who repeated or skipped a grade (about 5%), since it is not clear how instruction hours should be assigned for this group. In the school year 2007/2008, the three cohorts attended 3rd, 4th and 5th grade, respectively. For earlier years, where we do not observe the school attended, we assume that pupils attended the same school as in 2007/2008. However, we omit from the sample pupils who previously lived in other municipalities than the municipality of the school attended in 2007/2008 (about 7%), since the majority of this group changed school during the earlier years (and therefore we cannot assign instruction hours for these years). We also omit pupils who attended private schools in one or more school years (about 23%), as we do not have data on

[13] Oral examinations, the Danish essay and one of the Maths tests are evaluated by the teacher and an external examiner. Pupils also receive marks for the year's work, which are given by the teacher.

[14] The Danish National Agency for IT and Learning (STIL).

[15] Children enrol in 1st grade in the calendar year in which they turn 7. To allow for early and late school-start age, the three school start cohorts are based on the population of children born in 1995-1999.

instruction hours for private schools.[16] Finally, we omit pupils who did not have 9th grade test scores in Danish, Maths and English (about 3%). The final estimation sample consists of 102,875 pupils.[17] In our main analysis, we distinguish between pupils who changed school during their school career (34,324) and pupils who stayed at the same school from 1st to 9th grade (68,551).

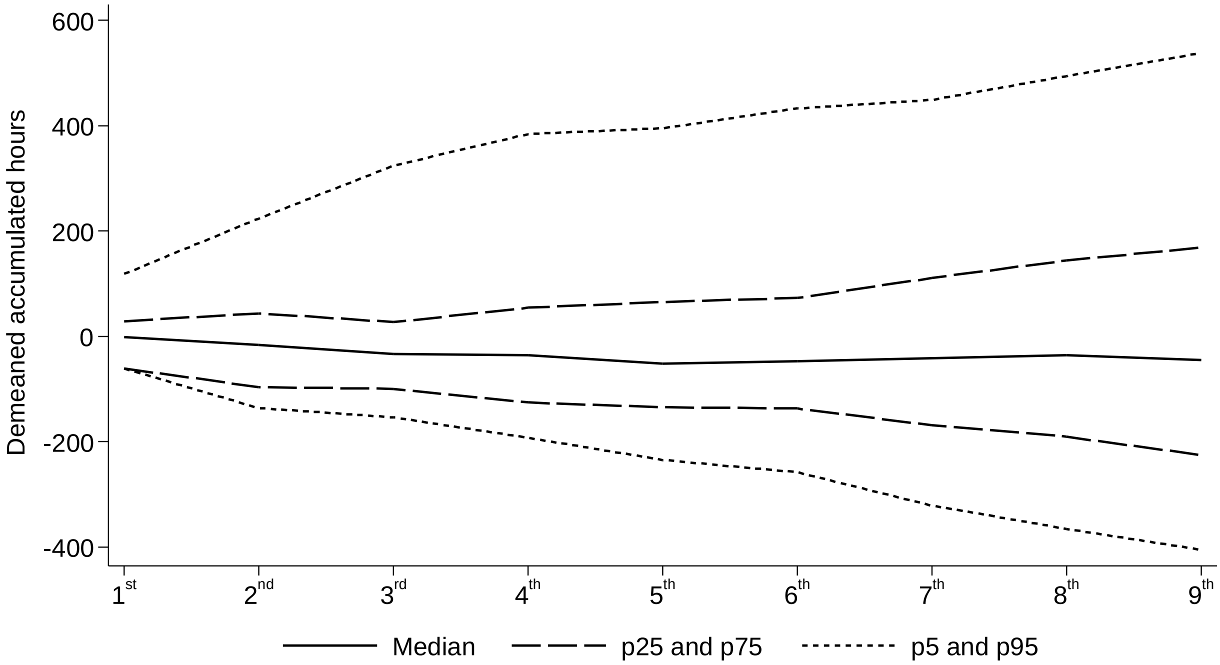
The variation in instruction hours is very large in Danish public schools generally – not merely in specific subjects or in specific grades – and the variation increases as hours are accumulated over several grades. To illustrate this, Figure 1 shows percentiles of accumulated and demeaned total hours of instruction in all subjects combined for all pupils in the dataset. For instance, at the point on the x axis for 1st grade we show on the y axis percentiles of instruction hours for 1st grade only; at the x axis point for 2nd grade we show the percentiles for accumulated hours for 1st and 2nd grade on the y axis; and at the x axis point for 9th grade we show the percentiles for accumulated total hours of the whole school career from 1st through 9th grade on the y axis. The difference between the 95th and 5th percentile in accumulated hours from 1st to 9th grade is about 940. This means that pupils at the 95th percentile receive more than one extra year of instruction during their school career, compared to pupils at the 5th percentile (a difference of about 12%).[18]

[16] In the early grades, for which we do not have school information (grades 1-2 for the '05 cohort, grades 1-4 for the '03 cohort), some pupils may have attended private schools. However, it is uncommon in Denmark for children to begin their school career in a private school and then transfer to a public school (the opposite is more common). The cohort starting in 1st grade in 2007 (which is not part of our estimation sample) is the first cohort for which we have data on the school attended in all grades. Analysis for this cohort indicates that among those enrolled in a public school in 3rd and 5th grade, respectively, and who did not move municipality between this grade and 1st grade, 5% and 9%, respectively, attended a different school in 1st grade. This early school mobility thus gives rise to measurement error of instruction time in one or more of the early grades of about 7% of our sample.

[17] About 5 per cent of pupils receive some teaching because of special needs. In our data, we are unable to distinguish special needs instruction hours from regular instruction hours.

[18] The overall average of accumulated instruction time from 1st to 9th grade is about 7500 hours, i.e. 830 hours per year. The number of instruction hours increases with increasing grades. In 9th grade, the average is about 1000 hours. In 1st grade, the average is about 700. For 1st grade, the difference between the 95th and 5th percentiles is 180 hours, i.e. 25% of the average instruction hours for 1st grade.

Figure 1. Percentiles of accumulated total instruction hours in all subjects by grade, demeaned



As will be discussed in detail in Section 4, we exploit variation in instruction time across subjects (Danish, Maths and English) to estimate pupil fixed effects models. We construct subject-specific accumulated instruction hours from 1st through 9th grade for each pupil and convert these into average weekly instruction hours.[19] Similarly, for each pupil we construct the average number of weekly instruction hours in grades 1-3, 4-6 and 7-9, and do this for each grade separately in order to investigate separate effects for different parts of the school career. Table 1 shows means and standard deviations for weekly instruction time and test scores. The variation in instruction time is largest for Danish, especially in the early grades. The small mean for English hours for grades 1-3 is due to the fact that English is typically taught from 3rd grade onwards, although some schools provide English lessons from the 1st or 2nd grade. In the analysis, we use test scores that are standardised for each of the three subjects for the total sample of pupils used in the estimations.[20] In the main analysis, we use test scores as the dependent variable, but we also show results using scores for the year's work.

[19] Dividing by 9 times 40, since a school year consists of 40 weeks.

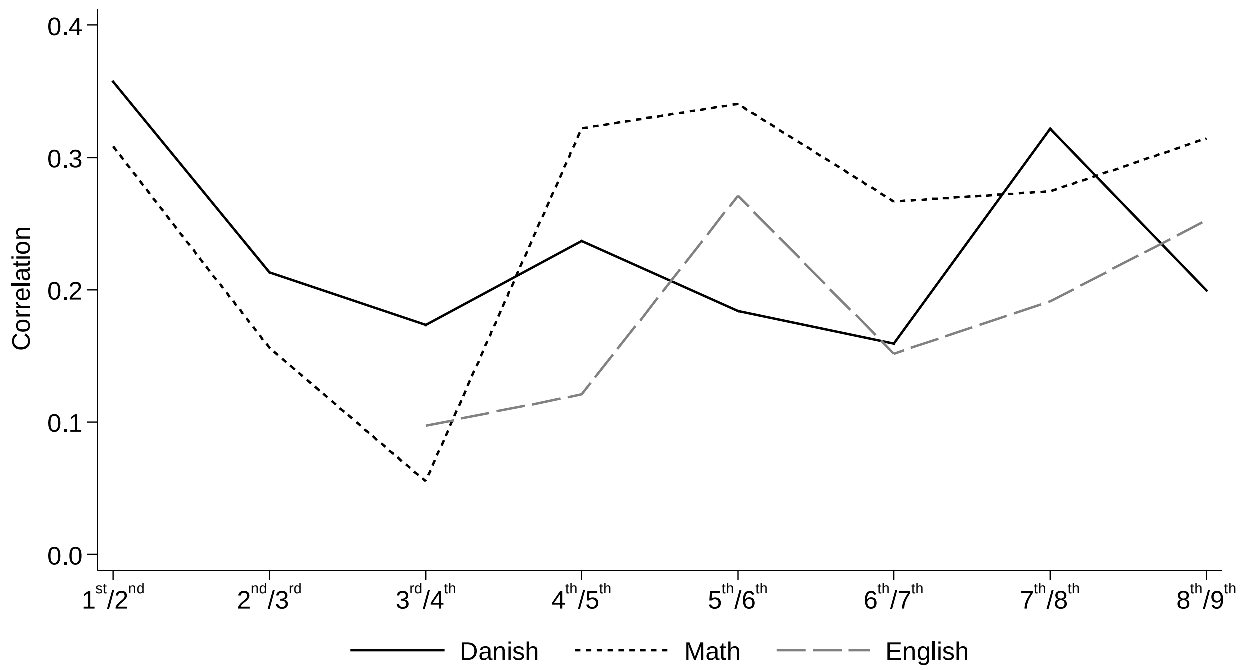
[20] Test scores are given according to a 7-point scale with the values: -3, 0, 2, 4, 7, 10 and 12.

Table 1. Summary statistics of weekly instruction hours and test scores by subject

	Danish		Maths		English	
	Mean	SD	Mean	SD	Mean	SD
Standardised test scores	0.00	1.00	-0.00	1.00	0.00	1.00
Test scores	6.62	3.00	6.48	3.18	7.51	3.54
Standardised scores for the year's work	-0.00	1.00	-0.00	1.00	0.00	1.00
Scores for the year's work	6.96	2.89	7.15	3.12	6.87	3.10
Weekly hours 1st-9th grade	5.47	0.29	3.31	0.13	1.60	0.07
Weekly hours 1st-3rd grade	7.19	0.64	3.83	0.30	0.50	0.10
Weekly hours 4th-6th grade	4.74	0.45	3.07	0.19	2.03	0.15
Weekly hours 7th-9th grade	4.48	0.21	3.02	0.11	2.28	0.12
Weekly hours 9th grade	4.49	0.28	3.03	0.18	2.30	0.21
Pupils			102,875			
Schools			988			

For each of the three subjects, Figure 2 shows correlations in the number of instruction hours between consecutive grades for pupils who did not change school, and the correlations thus only reflect variation within schools. The correlations are positive and mostly around 0.2-0.3. Correlations based on the whole sample, including pupils changing school, are similar. These modest positive correlations reflect two opposing mechanisms related to resource allocation in schools: some schools provide persistently high or low levels of instruction time across grades (which would give rise to high positive correlations), and, given budget constraints, schools may decide to provide more instruction time in some grades at the expense of others (which would tend to give rise to negative correlations).

Figure 2. Correlation in instruction hours between consecutive grades, by subject



Most studies investigating the effect of instruction time on pupil achievement have cross-sectional data for instruction time only in the school year in which achievement is measured, and have no information about the correlation in instruction time from grade to grade. For instance, this limitation applies to studies using the international PISA data (e.g., Lavy, 2015, and Rivkin & Schiman, 2015). Figure 3 shows correlations in our data between instruction hours in 9th grade and instruction hours in each of the earlier grades. For all three subjects, correlations in hours between 9th grade and grades 1-6 are below 0.1, and correlations between hours in 9th grade and grades 1-4 are close to zero. Thus, the variation in hours in 9th grade does not capture the variation in hours in earlier grades. This underscores the importance of investigating the effect of accumulated hours and of investigating potentially different effects of instruction time in earlier and later grades.

We identify effects of instruction time on test scores using within-pupil variation (see Section 4). Therefore, it is important that we have considerable within-pupil across-subject variation in instruction time and test scores. This is indeed the case. For average weekly instruction hours, the correlation coefficients between English and Danish, and between English and Maths are only about 0.2, and the correlation between Danish and Maths is 0.43-0.47 (depending on the sample); see Table A.1 in the Appendix. For test scores, the correlations are between 0.44 and 0.50; see Table A.2 in the Appendix.

Figure 3. Correlation in instruction hours between 9th grade and each of the lower grades, by subject

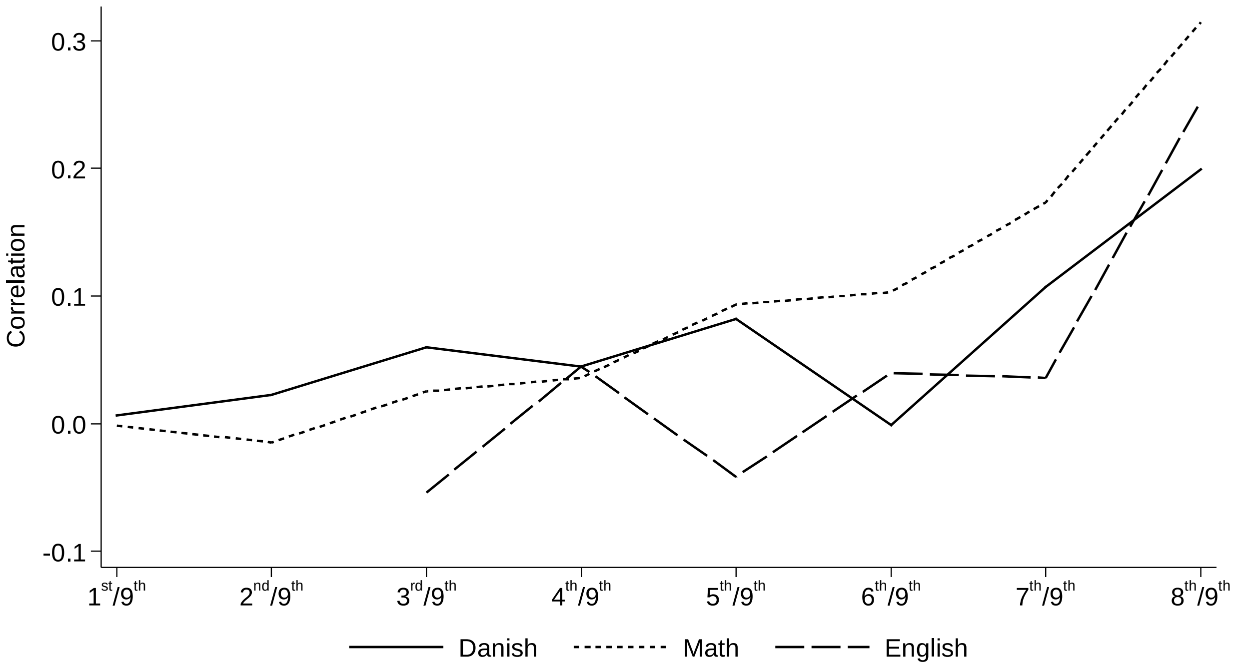


Table 2 shows summary statistics for control variables. Parental background variables are measured in the year the child turns 5, which is typically the year before school start. The "main control variables" listed in the table are used in most estimations. The "additional control variables" are used in interactions with instruction time (or in selecting the estimation sample) in some estimations. The variable "Low SES" is a dummy for both parents having a low level of education (no education beyond compulsory school), and the variable "High SES" is a dummy for at least one of the parents having a higher (i.e. post-secondary) level of education. The table shows statistics for the whole sample and for two subsamples: those who stay in the same school for their entire school career from 1st through 9th grade, and those who move school at least once. In some of the estimations, we allow for separate effects of instruction time for these two subgroups.

The control variables do not vary across subjects within pupils, but in the pupil fixed effects models we include interaction terms between these control variables and subject dummies, since they may explain some of the variation across subjects within pupils. For instance, boys typically have a comparative advantage in Maths; see Marks (2008). On average, pupils who do not change school have more advantaged backgrounds: They are more likely to live with both parents and to have parents with higher education, higher earnings and stronger labour market attachment. One important reason for moving school is that some schools are feeder schools, i.e. smaller schools (typically in non-urban areas) that only have pupils in earlier grades (typically grades 1-6) and from which pupils have to transfer to other schools for the final grades of compulsory schooling.

Table 2. Summary statistics of control variables

	All		Same school		Move school	
	Mean	SD	Mean	SD	Mean	SD
Main control variables						
'03 cohort	0.332		0.352		0.293	
'04 cohort	0.335		0.334		0.339	
'05 cohort	0.332		0.314		0.368	
Girl	0.489		0.479		0.509	
Living with both parents	0.785		0.800		0.754	
Living with one parent and the parent's new partner	0.066		0.059		0.082	
Living with one parent	0.145		0.138		0.159	
Living without parents	0.004		0.003		0.005	
Non-Western immigrant background	0.078		0.076		0.081	
Father not observed in registries	0.007		0.006		0.007	
Basic parental education (father)	0.253		0.239		0.282	
Vocational parental education (father)	0.457		0.444		0.485	
Higher parental education (father)	0.290		0.317		0.234	
Log wage (father)	10.981	(4.266)	11.134	(4.146)	10.674	(4.481)
Unemployment in per cent (father)	0.038	(0.138)	0.036	(0.135)	0.042	(0.144)
In workforce (father)	0.971		0.973		0.967	
Not in workforce (father)	0.029		0.027		0.033	
Welfare beneficiary (father)	0.039		0.035		0.046	
Mother not observed in registries	0.000		0.000		0.000	
Basic parental education (mother)	0.259		0.243		0.291	
Vocational parental education (mother)	0.411		0.404		0.426	
Higher parental education (mother)	0.330		0.353		0.283	
Log wage (mother)	10.408	(4.358)	10.592	(4.217)	10.040	(4.605)
Unemployment in per cent (mother)	0.064	(0.171)	0.059	(0.165)	0.073	(0.182)
In workforce (mother)	0.961		0.964		0.955	
Not in workforce (mother)	0.039		0.036		0.045	
Welfare beneficiary (mother)	0.097		0.086		0.120	
Additional control variables						
Low SES	0.120		0.108		0.143	
High SES	0.432		0.461		0.374	
Urban area	0.301		0.348		0.207	
Copenhagen and Aarhus	0.127		0.154		0.073	
Change from any institution	0.334		0.000		1.000	
Change from any institution except from feeder school	0.234		0.000		0.701	
Change from feeder school	0.100		0.000		0.299	
Pupils	102,875		68,551		34,324	
Schools	988		655		988	

4. Empirical model

Like other measures of school resources, instruction time may be correlated with observed and unobserved variables that affect pupil educational outcomes. Parents choose where to live, and they choose schools for their children. This non-random selection into schools may partly be determined by perceived school quality, including instruction time. As described in Section 3, instruction time differs a lot among schools in Denmark, both because of municipal decisions and school decisions. If high-ability pupils are sorted into schools with more instruction time, OLS estimates of the effect of instruction time would tend to be upward biased. Similarly, compensatory assignment by municipalities of more instruction time to schools with many disadvantaged pupils would tend to produce downward bias. To take account of these types of sorting, we identify effects of instruction time by estimating pupil fixed effects models as in Lavy (2015) and Rivkin & Schiman (2015). Thus, we use within-pupil across-subjects variation in instruction time and achievement. The strategy of using within-pupil across-subjects variation has also been applied in the literature on effects of teacher characteristics on pupil outcomes; see e.g. Dee (2005) and Clotfelter et al. (2010).

We estimate models of the form

$$y_{ijst} = \alpha + \beta h_{ijst} + \gamma^s X_{ijt} D_s + \mu_i + \eta_{st} + u_{ijst} \quad (1)$$

where y_{ijst} is the 9th grade test score for pupil i of cohort t at school j (in 9th grade) in subject s, h_{ijst} is accumulated instruction time from 1st to 9th grade in subject s for pupil i, X_{ijt} is a vector of characteristics of pupil i (including gender and parental background variables), D_s is a subject dummy that is equal to one if the subject is s (there is one dummy for each subject except for a reference category), μ_i is unobserved individual fixed effects, η_{st} is a subject by cohort fixed effect, and u_{ijst} is the remaining error term. In some models, we estimate separate effects of instruction time at different grade levels, in which case h_{ijst} is a vector of hours of instruction at different grade levels, and β is the corresponding parameter vector.

The pupil fixed effects control for overall pupil ability, pupil background and school characteristics, which do not vary by subject. However, pupils may have rather different abilities in different subjects. For instance, boys tend to have a comparative advantage in Maths and girls in language, and relative abilities in different subjects may also vary by parental socioeconomic variables. To take account of this, we include interaction terms between these individual background variables (X_{ijt}) and subject dummies in the model. In our regressions with three subjects (Danish, Maths and English), there are two subject dummies D_s, and the coefficients γ^s reflect the effects of the controls on test scores in subject s relative to the reference subject. Note that the pupil fixed effects μ_i also control for unobserved characteristics of school j for cohort t, or for unobserved characteristics of the sequence of schools that pupil i attended in his school career, if he did not stay in the same school in all years. Since all subjects are taught in the same class across all Danish schools, the pupil fixed effects also control for characteristics of the specific classes attended in the whole school career (including class size and peer quality).

One threat to our identification strategy is the possibility that sorting into schools is based on subject-specific characteristics. For instance, pupils with high ability and interest in Maths relative to other subjects might sort into schools with relatively more instruction time in Maths (producing upward bias), or pupils with a relatively low ability in Maths might sort into such schools because their parents are aware that they require more instruction time to learn Maths (producing downward bias). However, such sorting is not likely to be a major problem. Firstly, Danish public schools do not provide a particularly high quality in specific subjects at the expense of other subjects. Secondly, while parents may be aware of general indicators of school quality, such as class size and the overall level of instruction time, it is unlikely that they are aware of the relative amount of planned instruction time in different subjects when they choose a school for their children. Thirdly, schools do not admit pupils based on subject-specific considerations, and schools must not refuse to admit pupils living in their catchment area. Finally, there is no tracking within schools, and at the school-subject-grade-year level all pupils will be offered the same number of instruction hours.

One limitation of the identification strategy is that we have to assume that the effect of an extra hour of instruction time is the same in all subjects, i.e. the parameter β in (1) does not vary by subject. However, it is possible to allow β to vary by grade, and we do this in some of our model specifications. It is also important to note that there may be spill-over effects between subjects. For instance, improved skills in Danish may enhance skills in Maths and vice versa, and our estimates of β therefore may tend to be downward biased (since they are net of such spill-over effects).[21]

Most of our estimations are conducted for the subgroup of pupils who did not change school. There are several reasons for this. Firstly, pupils who attended feeder schools in the early grades, i.e. smaller schools (typically in non-urban areas) that only have pupils in earlier grades (typically grades 1-6), must transfer to other schools for the final grades of compulsory schooling. Feeder schools typically have small class sizes and few instruction hours. At these schools, an instruction hour may be systematically more efficient for pupil

[21] Rivkin & Schiman (2015) find much smaller effects when including language subjects, Maths and Science in the estimation than when including only language subjects and Maths, and they interpret this difference as indicating substantial spill-over effects between Maths and Science. However, Lavy (2016) finds no indication of important spill-over effects between subjects in his analysis using Israeli data on a large number of subjects, and his results using pupil fixed effects models are very similar to an alternative identification strategy using exogenous variation from a policy reform.

learning due to the small class size,[22] or systematically less efficient due to less specialised teachers at the smaller schools. This means that the interpretation of accumulation of teacher hours over grades may be problematic, especially if the distribution of teacher hours over grades differs across subjects. This is indeed the case for the average pupil, for whom most instruction time in Danish lies in the earlier grades, whereas the opposite is true for English; see Table 1. Thus, even though our identification strategy is to use pupil fixed effects accumulated instruction time may be a relatively noisy measure of school inputs for pupils who have attended feeder schools compared to pupils who have always attended the same school.[23]

Another reason for focussing on pupils who do not change school is that moving due to parental behaviour may be associated with other confounding changes in household resources and school resources; changes of address may be because of parental divorce, job loss or career opportunities; changes of school may be due to parents being dissatisfied with school quality in general or instruction time in particular.

In the regressions, standard errors are clustered on municipality as there is a lot of municipal discretion involved in decisions on the level and allocation of resources for public schools (see Section 2).

5. Results

5.1 Baseline results

Table 3 reports results from our baseline pupil fixed effects specification. The dependent variable is test scores in 9th grade. The main explanatory variables of interest are the average weekly hours (in Danish, Maths and English) over the entire nine years of school-life. In all specifications, we include indicator variables for subjects (reference is English) and subject indicators interacted with both cohort and the set of control variables on individual and parental background described in Table 2.[24] In model (1) of Table 3, we include all observations and disregard whether a pupil has changed school during the 9 years. This gives rise to a parameter estimate of 0.036, which is statistically significant at the 5% level. In model (2), we constrain the sample to pupils who do not change school. This results in a loss of one third of the observations, but gives a cleaner interpretation of the effect of instruction time, as previously discussed.

---

[22] Many studies find positive effects on pupil test scores of reducing class size, e.g. Krueger (1999), Angrist & Lavy (1999) and Heinesen (2010).

[23] Beuchert et.al. (2018) find negative test score effects for pupils changing schools due to school consolidations. Effects decrease over time, suggesting disruption shortly after school change is the cause.

[24] The estimates of hours effects do not change in any significant way if we leave out the interaction terms between subject dummies and parental background variables.

Restricting the sample to pupils with no change of institution increases the estimated effect of instruction time on 9th grade test scores considerably. The point estimate rises to 0.066 and becomes more significant. The interpretation of this estimate is that one extra hour of instruction time per week for nine years causes a rise in 9th grade test scores of 6.6% of a standard deviation, and this effect is statistically significant at the 1% level. The magnitude of this parameter is discussed further below.

Subsequently, in columns (3) and (4) of Table 3 we estimate alternative specifications, using all observations and including interaction terms between instruction time and dummies for change of school. Firstly, we include an interaction term between a dummy for change of institution, model (3), and secondly, we divide this change into a change in institutions that arises either because the initial school is a feeder school with classes going up to 5th or 7th grade only, or as a result of some other disruption. Feeder school changes are planned and likely represent much less disruption in the life circumstances of the pupil in question compared to the other type of changes, where the disruption might be expected to be more severe and have a negative impact on the pupil's school performance. For pupils who do not change school, the parameter of interest, the impact of average weekly hours in the 1st-9th grade, is virtually constant across the three types of specification of changes, models (2)-(4). It is highly significant in all three models, and the parameter estimate is 0.065-0.066. For pupils who change school, the estimated effect of instruction time is not significantly different from zero (the point estimate is about -0.015 when the main effect and the interaction effects are added). We do not find any difference in coefficients of the two interaction terms (see column (4)), and therefore the hypothesis regarding feeder-school interruptions does not find empirical support. The feeder school interaction is insignificant and the non-feeder school interaction is significant, which probably reflects the fact that the number of non-feeder school changes is more than twice as large as the number of feeder school changes (see Table 2).

It is no surprise that the estimated effect of instruction time is smaller for pupils who change school (though it is perhaps surprising that it is approximately zero). As discussed above, this difference likely reflects that accumulated instruction time is a poorer measure of school resources for those who change school compared to those who stay at the same school throughout their school career. Therefore, in the remainder of this paper we focus on the sample of pupils who do not change school.[25]

---

[25] The results reported below are robust to inclusion of all observations and interaction terms.

Table 3. Effect of instruction time on test scores in 9th grade, baseline specification

	(1) All pupils	(2) No change of school	(3) With interaction	(4) With interactions
Avg. weekly hours 1st-9th grade	0.036**	0.066***	0.065***	0.065***
	(0.016)	(0.022)	(0.023)	(0.023)
Avg. weekly hours 1st-9th grade			-0.082**	
x Change from any institution			(0.034)	
Avg. weekly hours 1st-9th grade				-0.079
x Change from feeder school				(0.049)
Avg. weekly hours 1st-9th grade				-0.079**
x Change except from feeder school				(0.036)
R^2	0.670	0.667	0.670	0.671
Observations	308,625	205,653	308,625	308,625

Note. Pupil fixed effects regressions. All regressions include subject dummies and their interactions with the main control variables, which are listed in Table 2. Standard errors clustered by municipalities.
* p < 0.10, ** p < 0.05, *** p < 0.01.

As a comparison, Table A.3 in the Appendix shows estimates of the hours effect using OLS (without pupil fixed effects) for the specification without interaction terms. When we include subject-by-cohort fixed effects and the full set of control variables, point estimates are positive but small and either insignificant (for the sample of all pupils, for which the point estimate is 0.01) or only marginally significant at the 10% level (for the sample of pupils who do not change school, for which the point estimate is 0.05). The relationship between our OLS and pupil fixed effects estimates contrasts with many other estimates in the literature. For instance, Lavy (2015) reports OLS estimates that are 3-4 times larger than his pupil fixed effects estimates and are statistically highly significant. The fact that our OLS estimates tend to be smaller than our pupil fixed effects estimates may indicate that there is very little sorting (by unobservable variables) of high-ability pupils into schools with more instruction time, or that such sorting is more than offset by other sorting mechanisms, for instance compensatory assignment by municipalities of more resources for instruction time to schools with a large proportion of pupils from disadvantaged backgrounds.

5.2 Learning dynamics and dynamic complementarity

An important advantage of the data applied here is that we observe the number of instruction hours at each grade level over all nine years. This enables us to answer questions related to not only "how much" but also "when". In order to understand the importance of "when", we split the sample by average weekly hours for subsets of the nine grades, see Table 4. Furthermore, we seek to understand whether there are any gender differences in learning dynamics. For instance, among 7 year olds (the normal school starting age in Denmark) girls are often more mature and 'ready' for learning (see e.g. Lim et al. (2015)). This could manifest itself in girls being more able to harvest benefits from more instruction hours in the early school years, with boys, possibly, catching up later in their school life.

Models (1)-(3) in Table 4 show estimates of the effect of instruction time for the first three-year interval only in total and by each gender separately. Similarly, Models (4)-(6) and (7)-(9) show estimates of the effects of instruction time for consecutive three-year intervals, while in models (10)-(12) all nine years are included but with a separate parameter for each three-year interval.

All point estimates are positive, although many are insignificant. Interestingly, and in line with the hypothesis that girls are more mature and more ready to learn in the early years, we find that the number of instruction hours in grades 1-3 is statistically significant for girls' test scores at the end of 9th grade. In comparison, boys seem to harvest more benefits from increased instruction time in grades 4-6. For grades 7-9, girls, again, show a statistically significant gain on their 9th grade test score, while boys do not. These results underscore the importance of having data on the entire school life.[26]

A main finding here is that instruction hours earlier may indeed have effects later in life, and that pupils' learning curves may differ widely over the course of their schooling life. Econometrically, this is also important, since it indicates that estimates based on cross-section data, which dominate the literature, may suffer from bias. Identification of causal effects of instruction time on pupil achievements based on cross-section data relies on the assumption that there is no correlation between lagged instruction time and the error term. In the presence of learning dynamics, however, this exogeneity assumption is violated.

As demonstrated in Figure 3, the correlation between instruction hours from one year to the following year is positive, and, on average, pupils with relatively many instruction hours prior to 9th grade also receive many instruction hours in the 9th grade. If the number of instruction hours in grades 1-8 matter for learning from additional instruction time in the 9th grade, but the effect is estimated based on instruction hours in the 9th grade only, the estimated effect will be biased.[27]

---

[26] Several of these gender and grade specific parameter estimates are only significant at the 10% level. However, it appears reasonable to expect a non-negative impact on test scores from more instruction time, and, consequently, the hypothesis to be tested is one-sided. In light of this, the significance levels reported are conservatively low.

[27] Rivkin & Schiman (2015: F433) acknowledge this potential problem. Reassuringly, they note that any bias will unambiguously be towards zero.

Table 4. Effects of hours across grades by gender

	(1) All	(2) Girls	(3) Boys	(4) All	(5) Girls	(6) Boys	(7) All	(8) Girls	(9) Boys	(10) All	(11) Girls	(12) Boys
Avg. weekly hours 1st - 3rd grade	0.017*	0.022*	0.012							0.014	0.020*	0.008
	(0.010)	(0.011)	(0.013)							(0.010)	(0.012)	(0.013)
Avg. weekly hours 4th - 6th grade				0.035**	0.030	0.041**				0.030*	0.021	0.038*
				(0.017)	(0.019)	(0.020)				(0.018)	(0.019)	(0.020)
Avg. weekly hours 7th - 9th grade							0.048	0.070*	0.028	0.041	0.064*	0.019
							(0.032)	(0.037)	(0.038)	(0.034)	(0.038)	(0.040)
R^2	0.667	0.666	0.661	0.667	0.666	0.661	0.667	0.666	0.661	0.667	0.666	0.661
Observations	205,653	98,526	107,127	205,653	98,526	107,127	205,653	98,526	107,127	205,653	98,526	107,127

Note. Pupil fixed effects regressions with test scores in 9th grade as the dependent variable, based on the sample of pupils who did not change school. All regressions include subject dummies and their interactions with the main control variables, which are listed in Table 2. Standard errors clustered by municipalities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To further investigate the importance of learning dynamics, and the potential bias, we estimate the model based on instruction time in 9th grade only and compare this to a model-specification that includes both the instruction time in 9th grade and the accumulated instruction time prior to 9th grade, i.e. in 1st-8th grade, see Table 5.

Table 5. Learning dynamics, effect of hours in 9th grade

	(1)	(2)	(3)
Avg. weekly hours 1st -9th grade	0.066***		
	(0.022)		
Avg. weekly hours 9th grade		0.030**	0.026*
		(0.014)	(0.014)
Avg. weekly hours 1st-8th grade			0.054**
			(0.021)
R^2	0.667	0.667	0.667
Observations	205,653	205,653	205,653

Note. Pupil fixed effects regressions with test scores in 9th grade as the dependent variable, based on the sample of pupils who did not change school. All regressions include subject dummies and their interactions with the main control variables, which are listed in Table 2. Standard errors clustered by municipalities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The first column of Table 5 simply shows the baseline result again, based on all 9 grades. In the second column, the causal effect of more instruction time in 9th grade is reported based on cross-section data for that grade only. The estimate in column (2) appears to be rather high, considering that it is based on 9th grade instruction time only. This estimate reflects the combined impact of the number of instruction hours in 9th grade and how these are correlated with instruction hours in the first 8 grades. In column (3), the same parameter is estimated, but here the model also includes a parameter for instruction hours in the 8 years prior to 9th grade.

With no control for the first 8 years, the point estimate is 0.030, while the same parameter estimate becomes 0.026 once we control for the first 8 years of accumulated instruction time. This indicates an upward bias of about 15% in the estimate of the effect of instruction time in 9th grade, when we do not control for instruction time in earlier grades. This difference is rather small (and not statistically significant), reflecting the modest positive correlation across grades in instruction time in our data, as illustrated by Figure 3.

The findings above turn around the (potential) importance of learning dynamics, which relates closely to the notion of *dynamic complementarity.* In a series of papers, James Heckman and co-authors argue for the existence of complementarity in learning, e.g. Cunha et al. (2006). Central to their analysis is the observation that childhood has more than one stage, and that skills beget skills through a multiplier process. Investments in skills at one stage increase the return on investments in skills at later stages. The policy implication is that early investments are essential, but also that early investments are productive only if they are followed up by later investments. The theory of dynamic complementarity relates to the entire life cycle skill formation, from the cradle to the grave, and thus we are not able to test this theory in our setting. However, inspired by this theory we test whether complementarity exists in the first 9 years of primary and lower secondary schooling by interacting the number of instruction hours received early in school life with the number of instruction hours received at later stages during the first 9 years, see Table 6.

Schools have been allocated into tertiles of their instruction hours, and here we compare the lower tertile (lowest 33%) to the upper tertile (highest 33%). In column (1) of Table 6, estimates are presented of the impact of hours in grades 4-6 and grades 7-9, conditional on the schools being in the lower tertile (receiving relatively few instruction hours) in the first three years of schooling. Similarly, column (2) shows the same parameters conditional on being in the upper tertile in the first three years of schooling. Signs of complementarity would manifest themselves in smaller parameters for the lower tertile vis-à-vis the upper tertile. The point estimates, however, are all highly insignificant. Further conditioning on both grades 1-3 and grades 4-6 makes no difference (columns (3) and (4)). Alternative specifications include interactions with continuous variables or interactions with dummy variables (columns (5) and (6)) with similar highly insignificant results. We find no sign of dynamic complementarity.

Table 6. Dynamic complementarity, effects of instruction time

	(1) 1st-3rd grade 1st tertile	(2) 1st-3rd grade 3rd tertile	(3) 1st-3rd grade 1st tertile & 4th-6th grade 1st tertile	(4) 1st-3rd grade 3rd tertile & 4th-6th grade 3rd tertile	(5) Interactions with continuous variables	(6) Interactions with dummy variables
Hours 1st- 3rd grade					-0.063	0.021*
					(0.043)	(0.013)
Hours 4th- 6th grade	-0.007	0.035			0.071	0.055
	(0.023)	(0.025)			(0.060)	(0.034)
Hours 1st- 3rd grade					0.007	
× Hours 4th - 6th grade					(0.005)	
Hours 1st- 3rd grade 1st tertile						-0.046
× Hours 4th - 6th grade						(0.034)
Hours 1st- 3rd grade 3rd tertile						-0.031
× Hours 4th - 6th grade						(0.035)
Hours 7th- 9th grade	0.052	0.041	0.046	0.044	0.067	0.014
	(0.050)	(0.061)	(0.074)	(0.062)	(0.076)	(0.038)
Hours 1st- 3rd grade					0.010	
× Hours 7th- 9th grade					(0.008)	
Hours 4th- 6th grade					-0.020	
× Hours 7th- 9th grade					(0.017)	
Hours 1st- 3rd grade 1st tertile						0.061
× Hours 7th- 9th grade						(0.040)
Hours 1st- 3rd grade 3rd tertile						0.032
× Hours 7th- 9th grade						(0.043)
Hours 4th- 6th grade 1st tertile						-0.005
× Hours 7th- 9th grade						(0.009)
Hours 4th- 6th grade 3rd tertile						-0.002
× Hours 7th- 9th grade						(0.010)
R^2	0.670	0.667	0.664	0.673	0.667	0.667
Observations	69,867	59,964	40,911	26,127	205,653	205,653

Note. Pupil fixed effects regressions with test scores in 9th grade as the dependent variable, based on the sample of pupils who did not change school. Hours are average weekly hours. Schools have been allocated into tertiles of their instruction hours in grades 1-3 and 4-6. Column (1) shows estimates of the impact of hours in grades 4-6 and grades 7-9, conditional on being in the lowest tertile (receiving relatively few instruction hours) in the first three years of schooling. Similarly, column

(2) shows the same parameters conditional on being in the upper tertile in the first three years of schooling. Columns (3) and (4) show results for effects of hours in grades 7-9 conditioning on the 1st and 3rd tertiles, repectively, in both grades 1-3 and grades 4-6. Columns (5) and (6) show results for the whole sample using specifications, including interactions with continuous variables or interactions with dummy variables. All regressions include subject dummies and their interactions with the main control variables listed in Table 2. Standard errors clustered by municipalities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.3 The importance of socioeconomic background and ethnicity

The amount of instruction time is frequently mentioned as a policy-instrument for fighting economic inequality through education. For this argument to hold, returns should be higher for socially disadvantaged groups or groups where parents are less able to assist with homework. Here, we identify two groups that might have particularly high benefits from increased instruction time: pupils with non-western immigrant background and pupils from families with low socioeconomic status (SES) vis-à-vis high SES.[28]

The SES-gradient in the impact of instruction hours on achievement is strikingly clear in all five model-specifications in Table 7, i.e. both in total and at various grade levels over the 9 years of primary and lower secondary schooling. The number of hours is statistically significant at the 5% level for all interactions with SES, and in many cases the effect of instruction hours for pupils from low SES families is 2-3 times higher than the effect for high SES families. While not fully comparable, these results are in line with Battistin & Meroni (2013), who find that additional instruction time for pupils in low achieving lower secondary schools in Southern Italy increases maths test scores in schools characterised by pupils from less advantaged backgrounds. The results underscore the importance of family background for learning, and how learning for pupils from relatively socioeconomically affluent families are insulated from school characteristics to some degree: the school matters much less when the family background is high SES. However, the results also emphasise that more instruction hours, at least to some extent, can make up for relatively poor family background characteristics.

[28] Low SES is defined as families where both parents have no more than a lower secondary education. The results are robust to a definition where only one of the parents belongs in this group. High SES is defined as families where at least one parent has tertiary/further education.

Table 7. Learning dynamics, effects of instruction time by socioeconomic background

	(1)	(2)	(3)	(4)	(5)
Avg. weekly hours 1st-9th grade	0.071***				
	(0.023)				
Avg. weekly hours 1st-9th grade	0.018***				
x Low SES	(0.006)				
Avg. weekly hours 1st-9th grade	-0.020***				
x High SES	(0.005)				
Avg. weekly hours 1st-3rd grade		0.021**			
		(0.010)			
Avg. weekly hours 1st-3rd grade		0.011***			
x Low SES		(0.003)			
Avg. weekly hours 1st-3rd grade		-0.012***			
x High SES		(0.003)			
Avg. weekly hours 4th-6th grade			0.040**		
			(0.017)		
Avg. weekly hours 4th-6th grade			0.021***		
x Low SES			(0.008)		
Avg. weekly hours 4th-6th grade			-0.021***		
x High SES			(0.007)		
Avg. weekly hours 7th-9th grade				0.056*	
				(0.032)	
Avg. weekly hours 7th-9th grade				0.027**	
x Low SES				(0.011)	
Avg. weekly hours 7th-9th grade				-0.026***	
x High SES				(0.009)	
Avg. weekly hours 9th grade					0.036**
					(0.015)
Avg. weekly hours 9th grade					0.024**
x Low SES					(0.010)
Avg. weekly hours 9th grade					-0.021**
x High SES					(0.009)
R^2	0.667	0.667	0.667	0.667	0.667
Observations	205,653	205,653	205,653	205,653	205,653

Note. Pupil fixed effects regressions with test scores in 9th grade as the dependent variable, based on the sample of pupils who did not change school. All regressions include subject dummies and their interactions with the main control variables listed in Table 2 and the High/Low SES dummies. Standard errors clustered by municipalities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8 shows results for the group of pupils with non-western immigrant background. For both boys and girls from this ethnic group, the effect of instruction hours from 1st to 9th grade is larger than the overall sample average of 0.066. The point estimates are 0.112 for boys and 0.093 for girls. The gains from more instruction hours are significant in the early years (grades 1-3) for both genders. This is in contrast to the overall estimates for boys in Table 4, where there is no statistically significant effect of more hours in grades 1-3. Intuitively, this result is not surprising, since many of these children come from families where Danish is not the main language, see also Andersen et al. (2016).

Table 8. Learning dynamics of pupils with non-western immigrant background, effects of instruction time by gender

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Boys						Girls					
Avg. weekly hours 1st-9th grade	0.112**						0.093***					
	(0.053)						(0.030)					
Avg. weekly hours 1st-3rd grade		0.055***			0.050***			0.038***			0.033***	
		(0.015)			(0.016)			(0.013)			(0.012)	
Avg. weekly hours 4th-6th grade			0.048		0.033				0.043*		0.026	
			(0.050)		(0.054)				(0.024)		(0.024)	
Avg. weekly hours 7th-9th grade				0.001	-0.023					0.056	0.037	
				(0.085)	(0.086)					(0.099)	(0.104)	
Avg. weekly hours 9th grade						-0.008						0.008
						(0.048)						(0.068)
R^2	0.672	0.672	0.671	0.671	0.672	0.671	0.680	0.680	0.679	0.679	0.680	0.679
Observations	7,395	7,395	7,395	7,395	7,395	7,395	8,280	8,280	8,280	8,280	8,280	8,280

Note. Pupil fixed effects regressions with test scores in 9th grade as the dependent variable, based on the sample of pupils with non-western immigrant background who did not change school. All regressions include subject dummies and their interactions with the main control variables listed in Table 2 (except the dummy for immigrant background). Standard errors clustered by municipalities, * p < 0.10, ** p < 0.05, *** p < 0.01.

5.4 Timetabled hours and hours taught

Several different measures of instruction time have been considered in the literature – e.g. nationally or regionally mandated, school planned, teacher taught or pupil received – and the interpretation of the effects depends upon the nature of the measure and who is being asked to report. At issue are the goals of the inference and the nature of the measures available. Regionally mandated or school timetabled instruction would allow estimation of ITT effects because reduced form estimates do not require information about the level of compliance. Teacher and pupil reports might mismeasure instruction time because of absence, where a substitute teacher took the class or a pupil did not attend a class. School reporting would account for teacher substitution and pupil absence, but might be biased towards compliance with mandates.

While it is clear that different measures have their pros and cons, studies typically use the single measure that is available, paying scant attention to the consequences of inference for the sake of comparison with other papers. Cattaneo et al. (2017) is unique in comparing estimates from two measures of instruction time – PISA pupil reports and Swiss Canton instruction mandates – and finds that estimates are remarkably similar between measures.

We use school timetabled instruction as our primary measure of instruction time, and for two years we also observe school reports of delivered instruction in each grade, summed across subjects. Ideally, we would estimate a causal effect, instrumenting school reports of delivered with timetabled instruction. However, schools do not report delivered instruction time separately by subject, meaning that we could no longer use our preferred specification to transform the data within-pupil between-subject. Instead, we estimate illustrative “first stage” regressions, explaining delivered with timetabled instruction time in Table 9. Schools report instruction time performed by class teachers and substitute teachers separately, and estimates using these measures are shown in panels A and B in Table 9.

Table 9. Delivered instruction time

	(1) 1st grade	(2) 2nd grade	(3) 3rd grade	(4) 4th grade	(5) 5th grade	(6) 6th grade	(7) 7th grade	(8) 8th grade	(9) 9th grade	(10) All grades pooled
					PANEL A – Delivered instruction time					
Planned weekly hours	0.981***	0.985***	0.950***	0.961***	0.982***	0.962***	0.962***	0.967***	0.925***	0.932***
	(0.005)	(0.008)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)	(0.010)	(0.013)	(0.002)
R^2	0.954	0.900	0.910	0.899	0.911	0.896	0.899	0.833	0.740	0.947
N	1,851	1,853	1,859	1,859	1,859	1,859	1,868	1,860	1,856	16,724
					PANEL B – Class teacher-delivered instruction time					
Planned weekly hours	0.912***	0.877***	0.886***	0.822***	0.919***	0.823***	0.765***	0.881***	0.847***	0.912***
	(0.024)	(0.026)	(0.035)	(0.036)	(0.031)	(0.033)	(0.027)	(0.029)	(0.025)	(0.006)
R^2	0.434	0.380	0.257	0.224	0.322	0.256	0.300	0.331	0.385	0.608
N	1,851	1,853	1,859	1,859	1,859	1,859	1,868	1,860	1,856	16,724

Estimates from separate OLS regressions. The dependent variable is planned hours x share taught in November 2010 and 2011. Coefficients of interest are presented in the table. Intercepts are included in the regressions but not shown.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

From panel A, we can see that planned explains about 95% of total delivered instruction time, and from panel B about 61% of class teacher delivered instruction time. There are no obvious patterns in the associations grade-to-grade. We can interpret these estimates in two complementary ways. One, as causal effects of delivered class teacher instruction time due to timetabling; shifting up our ITT estimates by 1/0.91 or 10%. Two, as measurement error-corrected estimates of instruction time – assuming that timetabled and class teacher-delivered instruction time are independent measures of the same object, we find a reliability ratio of 0.91.

5.5 Robustness tests

In this section, we demonstrate the robustness of our baseline estimate of the effect of instruction time by conducting five robustness checks.

Firstly, we restrict the sample by excluding the two major cities, Copenhagen and Aarhus, where selection into specific schools is arguably higher than in small towns or rural areas, because of the greater choice of public and private schools. For instance, the share of pupils in private schools is 30-40% in Copenhagen, while it is less than 20% for the entire country. This reduces the sample by 15%. Secondly, and in addition to the first restriction, we also exclude the greater Copenhagen area and other cities, i.e. Odense and Aalborg. This reduces the sample size by a further 20%. Both restrictions tend to slightly increase the point estimate, and both estimates are significant at the 1% level, see Table 10.

Next, as an alternative to exam scores we use continuous assessments given by the teachers for each pupil in their class. This grade is intended to reflect how well the pupil performed throughout the school year. *A priori*, one might hypothesise that the test score best captures actual learning, while the grade for the year's work also reflects the liaison with the teacher and peers in the class. The point estimate of 0.053 in column (3) of Table 10 represents a small decrease from the baseline estimate of 0.066, and this possibly supports the hypothesis. In any event, the difference is modest, and the overall result is unaltered.

In the next robustness check, we exclude data for the subject English – which has fewer hours of instruction than Danish and Maths and is typically only taught from 3rd grade onwards – and estimate the causal effect on data for Danish and Maths only. This reduces the sample size by a third, but the baseline parameter remains significant (at the 10% level) and the point estimate becomes 0.072, i.e. once again close to the baseline result of 0.066. In the latter robustness check, we add Physics, which is only taught in grades 7-9, as an extra subject. This increases the number of observations compared to the baseline by almost a third. The point estimate of the effect of average weekly hours is 0.055 and is significant at the 1% level.

Table 10. Robustness checks

	(0) Baseline	(1) Not CPH and Aarhus	(2) Non-urban	(3) Year marks	(4) Excluding English	(5) Including Physics
Avg. weekly hours 1st-9th grade	0.066***	0.076***	0.085***	0.053**	0.072*	0.055***
	(0.022)	(0.026)	(0.030)	(0.022)	(0.038)	(0.020)
R^2	0.667	0.665	0.666	0.731	0.763	0.622
Observations	205,653	173,922	134,178	205,391	137,102	273,616

Note. Pupil fixed effects regressions with test scores in 9th grade as the dependent variable, based on the sample of pupils who did not change school. All regressions include subject dummies, and their interactions with the main control variables listed in Table 2. Standard errors clustered by municipalities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6. Discussion and perspective

Our results indicate that, on average, an increase in weekly instruction time in a given subject of 1 hour throughout the school career leads to an increase in test scores in this subject of about 0.06 standard deviations in the distribution of individual test scores.

6.1 Comparison to existing studies

Using PISA 2006 data for OECD countries and similar methods, Lavy (2015) finds that an increase in weekly instruction time of 1 hour also increases PISA test scores by about 0.06 SDs in the test score distribution. However, in the PISA data, instruction time is measured for the year of the PISA tests only (not for the whole school career), and instruction time is based on pupils' self-reported hours of school attendance in the subject (not the number of hours offered by the school). The extent to which the estimates of Lavy (2015) merely reflect effects of increasing weekly instruction time in 9th grade, or also in earlier grades, depends on the size of the correlations of instruction hours across grades in the OECD countries included in the analysis.[29] Since these correlations are presumably much larger than zero, the estimates in Lavy (2015) indicate larger effects than our estimates.

Based on PISA 2009 data for 72 countries and similar methods, Rivkin & Schiman (2015) also find significant positive effects of instruction time, but point estimates indicate smaller effects than those found in Lavy

[29] Rivkin & Schiman (2015) report that the correlation between school average instruction time differences in 9th and 10th grade is 0.41 in the 2009 PISA data, which indicates higher correlations than in our data, see Section 3.

(2015).[30] One reason for the smaller estimates may be that Rivkin & Schiman pool data for all countries, including developing countries, for which Lavy finds smaller effects than for OECD countries.

When comparing our results to the results of Lavy (2015) and Rivkin & Schiman (2015), it is important to note that measures of instruction time differ. We use administrative data on planned hours at the school-subject-grade-year level. Since planned and delivered instruction time differ, our estimates may be interpreted as ITT effects. The analysis in Section 5.4 indicates that the causal effect of class teacher-delivered instruction time is about 10% larger than our ITT estimates.[31] Instruction time in the PISA data is self-reported by the individual pupils and reflects instruction time attended by each pupil, not instruction time offered by the school. The question (which pupils answered for each subject) from PISA 2006 that Lavy (2015) used is: "How much time do you typically spend per week studying the following subjects in regular lessons at school?". If pupils do not attend all lessons offered by the school, for instance because they play truant from some of the classes, their answers might reflect this. In their analysis, both Lavy and Rivkin & Schiman use average instruction time in each subject at each school reported by the pupils (and Rivkin & Schiman also differentiate by grade, i.e. 9th or 10th grade). However, since only a few classes of pupils at each school participate in PISA, subject-specific truancy might be important at the school level, e.g. it might reflect the relative quality of teachers in different subjects at the school. This might tend to produce upward biased estimates of the effects of instruction time.

6.2 Policy considerations

From a policy perspective it is interesting to compare estimates of the effect of increasing instruction time with effects of alternative school interventions, and also to compare costs associated with these different interventions. One example of an alternative intervention is class size reduction. Clearly, this is a very different intervention since, given the number of instruction hours in each subject, a class size reduction affects learning conditions in all subjects proportionately, whereas instruction time may be adjusted differently for different subjects. Using data from the STAR experiment, Krueger (1999) finds that reducing class size by 7-8 pupils from a level of 22 from kindergarten class to 3rd grade increases test scores in both reading and maths by about 0.20 SDs in the distribution of individual test scores. Similar effect sizes are found in Angrist & Lavy (1999), using data for 5th graders in Israel, and in Heinesen (2010) using data for 9th graders in Denmark. The class size reduction in the STAR experiment corresponds to increasing the number of classes

[30] Their main estimates are for the language subjects and Maths only, and here the estimated effects are about half the size of those found in Lavy (2015) for languages, Maths and Science, but their estimates are even smaller when they include science in the estimations.

[31] This interpretation assumes that substitute teacher instruction has no effect on test scores. If it has the same effect as class teacher instruction, our analysis indicates that the ITT estimates should only be increased by about 7%.

in the first four grades by about 50%. This intervention is presumably at least as costly as increasing instruction time in all subjects by 50% for the same grades. The number of teacher hours required for the two interventions would be about the same, but increasing the number of classes would presumably be more expensive in terms of investment in new school buildings, since increasing instruction time may be possible to a larger extent if the existing building capacity is used for more hours each day. Increasing the number of instruction hours by 50% in the first four grades corresponds to an increase in accumulated instruction time from 1st to 9th grade of 15-20%, since the number of classroom hours is larger at higher grades. In our data, the average number of weekly classroom hours in Danish, Maths and English combined from 1st to 9th grade is about 10.4. An increase of 15% would correspond to about 0.5 extra hours per subject per week on average throughout the school career. Our estimates indicate that this would increase test scores in each subject by about 0.03 SDs. This is only about 15% of the effect of the corresponding class size reduction, but it would be possible to increase hours in, for instance, Maths (or Maths, Danish and English) much more, if instruction time in other subjects were held constant, thereby targeting skill improvement in specific subjects that are considered especially important. It is also important to keep in mind that our estimates of effects of instruction time are potentially downward biased due to spill-over effects between subjects.

These back-of-the-envelope comparisons suggest that, while statistically significant, the marginal learning impact of an increase in instruction hours may be relatively modest, except for more vulnerable groups (pupils with low SES or non-western immigrant backgrounds), who benefit significantly more than the average pupil.

7. Conclusion

Schooling is one of the most important areas for public service; it is a pathway to increased individual earnings and better career prospects as well as to economic growth in society. Public involvement is also a reflection of distributional concerns. With stakes this high, it is important to understand the impact of school quality on pupil achievement. One important aspect of school quality is instruction time, and, in light of the large cross-country differences among OECD countries, it is important to understand the impact of the amount and timing of instruction time for pupil achievement. A rather extensive literature sheds light on the impact of instruction time, but our study is the first to investigate the impact of accumulated instruction time over the entire span of compulsory schooling years. Using administrative data for Denmark, we find remarkable differences in the accumulated number of instruction hours from 1st to 9th grade: the difference between the 5th and the 95th percentiles is equivalent to more than the average instruction time in an entire school year.

Using within-pupil across-subject variation in hours and test scores, our main findings are the following. On average, an increase in weekly instruction time of 1 hour from 1st to 9th grade increases test scores at the end of 9th grade by about 0.06 SD. Compared to previous studies, which have used 9th grade instruction hours only, we find that all-grade effects are twice as large. In addition, we find that the timing of instruction hours between grades matters. Girls benefit significantly from instruction time in primary (grades 1-3) and lower secondary school (grades 7-9), whereas boys benefit from instruction time in junior school (grades 4-6). Testing interactions of instruction time in different grades, we find no evidence of complementarities – earlier instruction time does not significantly enhance the effect of later instruction time. We also find that accumulated instruction hours effects have a steep SES gradient, with a larger positive impact of instruction time for pupils from low SES families and pupils with non-western immigrant background, especially boys. Our measure of instruction time is based on planned weekly hours by subject, grade and school. Since planned hours and hours taught may differ, our estimates may be interpreted as ITT effects. For a subset of the years, we are able to compare delivered hours with planned hours, and their relationship indicates that effects of delivered hours are approximately 10% higher than our ITT effect estimates.

The conclusion from this study is that, overall, we find that an increase in instruction time has positive and statistically significant effects on pupil achievement. The effects are rather small for the average pupil, but considerably larger for pupils with low SES or a non-western immigrant background. Effects seem to be small compared to many estimates of class size effects, but changing instruction time is a more flexible intervention, which may be more easily targeted at specific subjects and grades.

References

Andersen, S. C., Humlum, M. K. & A. B. Nandrup (2016). Increasing Instruction Time in School Does Increase Learning. *Proceedings of the National Academy of Sciences of the United States of America*, **113**(27): 7481-7484.

Angrist, J.D. & V. Lavy (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics* 114(2): 533-575.

Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *Quarterly Journal of Economics*, **106**(2): 407-43.

Battistin, E. & E. C. Meroni (2013). Should We Increase Instruction Time in Low Achieving Schools? Evidence from Southern Italy. *IZA DP* No. 7437.

Bellei, C (2009). Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile, *Economics of Education Review* **28**(5): 629 – 640.

Beuchert, L., M. K. Humlum, H. S. Nielsen & N. Smith (2018). The short term effects of school consolidation on student achievement: Evidence of disruption?, *Economics of Education Review* **65**(1): 31-47.

Bils, M. & P. Klenow (2000). Does Schooling Cause Growth? *American Economic Review*, **95**: 1653-1672.

Black, S., P. J. Devereux & K. G. Salvanes (2008). Staying in the classroom and out of the maternity ward? The effects of compulsory schooling laws on teenage births'. *Economic Journal*, **118**(530): 1025-1054.

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

Card, D. & Krueger, A.B. (1996). Labor market effects of school quality: theory and evidence, in (G. Burtless, ed.), Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success, pp. 97–140, Washington, D.C: Brookings Institution Press.

Clark, D. & H. Royer (2013). The Effect of Education on Adult Mortality and Health: Evidence from the United Kingdom. *American Economic Review*, **103**(6): 2087-2120.

Clotfelter, C.T., H.F. Ladd & J.L. Vigdor (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Journal of Human Resources* 45 (3): 655–81.

Cunha, F., J. J. Heckman, L. Lochner & D. V. Masterov (2006). Interpreting the Evidence on Life Cycle Skill Formation. In E. Hanushek & F. Welch (eds.) *Handbook of the Economics of Education*, **1**: 697-812.

Currie, J. & E. Moretti (2003). Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings. *Quarterly Journal of Economics*, **118**, 1495-1532.

Dahmann, S.C. (2017). How does education improve cognitive skills? Instruction time versus timing of instruction. *Labour Economics* 47: 35-47.

Dearden, L., Ferri, J. & Meghir, C. (2002). The effect of school quality on educational attainment and wages, *Review of Economics and Statistics*, **84**(1), pp. 1–20.

Dee, T.S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review* 95 (2): 158-165.

Dustmann, C., Rajah, N. & van Soest, A. (2003). Class size, education, and wages, *Economic Journal*, **113**, pp. F99–120.

Fitzpatrick, M. D., D. Grissmerb & S. Hastedt (2011). What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment. *Economics of Education Review,* **30:** 269–279.

Goldin, C. & L. F. Katz (2009). *The Race Between Education and Technology*. Harvard University Press.

Grossman, M. (2006). Education and Nonmarket Outcomes. Chapter 10 in Handbook of the Economics of Education, vol. 1, ed. Eric Hanushek and Finis Welch. Elsevier.

Hansen, B. (2011). School Year Length and Student Performance: Quasi Experimental Evidence. Manuscript, University of Oregon.

Hanushek, E.A. (1996). School resources and student performance, in (G. Burtless, ed.), Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success, pp. 43–73, Washington, D.C: Brookings Institution Press.

Hanushek, Eric A. (2015). Time in Education: Introduction, *The Economic Journal* **125**(588): F394–F396.

Harmon, C. & I. Walker (1995). Estimates of the Economic Return to Schooling for the United Kingdom. *American Economic Review*, **85**(5): 1278–86.

Hedges, L.V. & Greenwald, R. (1996). Have times changed? The relation between school resources and student performance, in (G. Burtless, ed.), Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success, pp. 74–92, Washington, D.C: Brookings Institution Press.

Heinesen, E. (2010). Estimating class-size effects using within-school variation in subject-specific classes. Economic Journal 120: 737-760.

Huebener, M., S. Kuger & J. Marcus (2017): Increased instruction hours and widening gap in student performance. *Labour Economics* 47: 15-34.

Huebener, M. & J. Marcus (2014). Adverse effects of increased education efficiency? The impact of shortening high school tenure on graduation age, grade repetitions and graduation rates, Berlin: Deutsches Institut für Wirtschaftsforschung.

Jensen, V. M. (2013). Working longer makes students stronger? The effects of ninth grade classroom hours on ninth grade student performance, *Educational Research* **55**(2): 180–194.

Katz, L. F. & D. Autor (1999). Changes in the Wage Structure and Earnings Inequality, in *Handbook of Labor Economics*, **3A**, edited by Orley Ashenfelter and David Card. Amsterdam and New York: North Holland.

Krueger, A.B. (1999). Experimental estimates of educational production functions. *Quarterly Journal of Economics* 114(2): 497-532.

Lavy, V. (2015). Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries, *The Economic Journal* **125**(588): F397–F424.

Lavy, V. (2016). Expanding school resources and increasing time on task: Effects of a policy experiment in Israel on student academic achievement and behavior. Manuscript, University of Warwick.

Lefgren, L. & F. McIntyre (2006). The Relationship between Women's Education and Marriage Outcomes. *Journal of Labor Economics*, **24**(4): 787-830.

Lim, S., C. E. Han, P. J. Uhlhaas & M. Kaiser (2015). Preferential Detachment During Human Brain Development: Age- and Sex-Specific Structural Connectivity in Diffusion Tensor Imaging (DTI) Data. *Cerebral Cortex*, **25**(6): 1477-89.

Marcotte, D. E. (2007). Schooling and Test Scores: a Mother-Natural Experiment. *Economics of Education Review,* **26**(5): 629–640.

Marcotte, D. E. & S. W. Hemelt (2008). Unscheduled School Closings and Student Performance. *Education Finance and Policy*, **3**(3): 316-338.

Marks, G. (2008). Accounting for the gender gaps in student performance in reading and mathematics: Evidence from 31 countries. *Oxford Review of Education*, 34: 89-109.

Milligan, K., E. Moretti & P. Oreopoulos (2004). Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics*, **88**:1667-95.

Moretti, E. & L. Lochner (2004). The effect of education on criminal activity: evidence from prison inmates, arrests and self-reports. *American Economic Review*, **94**(1): 155-89.

OECD (2015). *Education at a Glance 2015: OECD indicators*. OECD Publishing, Paris.

OECD (2016). *Education at a Glance 2016: OECD indicators*. OECD Publishing, Paris.

Oreopoulos, P. & K. G. Salvanes (2011). Priceless: The Nonpecuniary Benefits of Schooling. *Journal of Economic Perspectives*, **25**(1): 159-184.

Parinduri, R. A. (2014). Do children spend too much time in schools? Evidence from a longer school year in Indonesia. *Economics of Education Review*, **41**(2014): 89-104.

Pischke, Jörn-Steffen (2007). The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years, *The Economic Journal* **117**(523): 1216–1242.

Rivkin, Steven G. & Jeffrey C. Schiman (2015). Instruction Time, Classroom Quality, and Academic Achievement, *The Economic Journal* **125**, 588:F425–F448.

UNI-C (2012). Undersøgelse af planlagt, gennemført og aflyst undervisning i folkeskolen, november 2011 (Investigation of planned, accomplished and cancelled teaching in public schools, November 2011). UNI-C Statistik & Analyse.

Appendix

Figure A.1. Cohorts, school years and grades for the period with data for planned instruction hours: 2003/2004-2013/2014

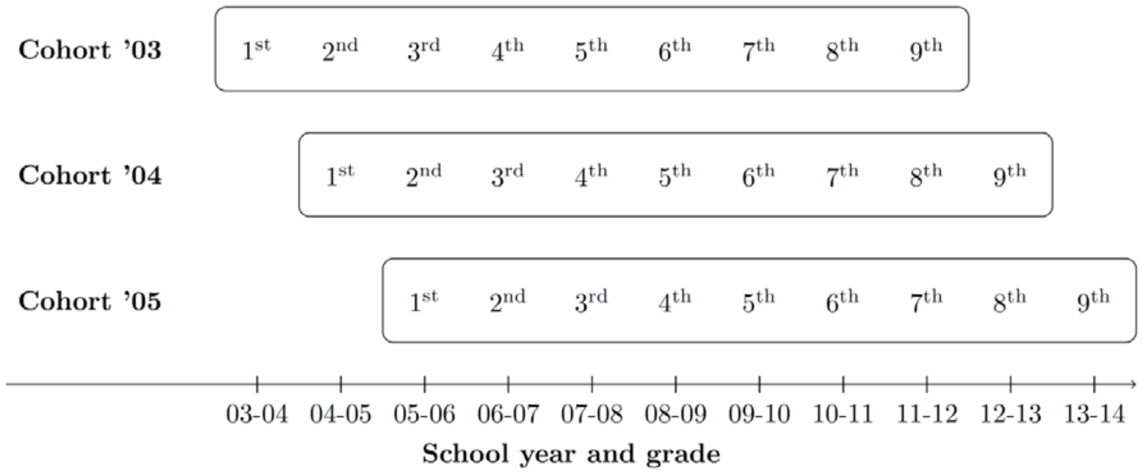


Table A.1. Within-student between-subject correlations in average weekly hours 1st-9th grade

	All			No change of school		
	Danish	Maths	English	Danish	Maths	English
Danish	1.000	–	–	1.000	–	–
Maths	0.473	1.000	–	0.428	1.000	–
English	0.172	0.198	1.000	0.180	0.210	1.000

Table A.2. Within-student between-subject correlations in test scores

	All			No change of school		
	Danish	Maths	English	Danish	Maths	English
Danish	1.000	–	–	1.000	–	–
Maths	0.480	1.000	–	0.477	1.000	–
English	0.498	0.448	1.000	0.495	0.438	1.000

Table A.3. OLS estimates of the effect of instruction time on test scores in 9th grade

	(1) All	(2) No change of school	(3) All	(4) No change of school	(5) All	(6) No change of school
Avg. weekly hours 1st-9th grade	-0.052	-0.007	0.007	0.045*	0.009	0.048*
	(0.052)	(0.058)	(0.025)	(0.027)	(0.024)	(0.026)
Subject-by-cohort fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	No	No	Yes	Yes	Yes	Yes
Controls and subject interactions	No	No	No	No	Yes	Yes
R^2	0.002	0.002	0.110	0.110	0.130	0.129
Observations	308,625	205,653	308,625	205,653	308,625	205,653

Standard errors clustered by municipalities, * p < 0.10, ** p < 0.05, *** p < 0.01.

Table A.4. Baseline specification, pupil fixed effects, 9th grade hours

	(1) All	(2) No change	(3) With one interaction	(4) With interactions
Avg. weekly hours 9th grade	0.010	0.030**	0.030**	0.030**
	(0.014)	(0.014)	(0.014)	(0.014)
Avg. weekly hours 9th grade			-0.065**	
x Change from any institution			(0.030)	
Avg. weekly hours 9th grade				-0.090**
x Change from feeder school				(0.042)
Avg. weekly hours 9th grade				-0.056
x Change except from feeder school				(0.039)
Controls	Yes	Yes	Yes	Yes
Subject and cohort interactions	Yes	Yes	Yes	Yes
R^2	0.670	0.667	0.670	0.671
Observations	308,625	205,653	308,625	308,625

Standard errors clustered by municipalities, * p < 0.10, ** p < 0.05, *** p < 0.01.