DISCUSSION PAPER SERIES

IZA DP No. 9797

The More, the Better? The Impact of Instructional Time on Student Performance

Maria A. Cattaneo
Chantal Oggenfuss
Stefan C. Wolter

March 2016

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The More, the Better? The Impact of Instructional Time on Student Performance

Maria A. Cattaneo
Swiss Coordination Center for Research in Education

Chantal Oggenfuss
Swiss Coordination Center for Research in Education

Stefan C. Wolter
University of Bern,
Swiss Coordination Center for Research in Education, CESifo and IZA

Discussion Paper No. 9797
March 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 9797
March 2016

ABSTRACT

The More, the Better?
The Impact of Instructional Time on Student Performance

Although instruction time is an important and costly resource in education production, there is a remarkable scarcity of research examining the effectiveness of its use. We build on the work of Lavy (2015) using the variance of subject-specific instruction time within Switzerland to determine the causal impact of instruction time on student test scores, as measured by the international PISA test (2009). We extend the analyses in two ways and find that students must differ considerably in the time needed to learn. This difference is supported by our findings that the effectiveness of instructional time varies substantially between different school (ability) tracks and that additional instruction time significantly increases the within-school variance of subject-specific test scores.

JEL Classification: C21, I21, I25

Keywords: instruction time, PISA, fixed-effect models, tracking

Corresponding author:

Stefan C. Wolter
Centre for Research in Economics of Education
Department of Economics
University of Bern
Schanzeneckstrasse 1
P.O. Box 8573
CH-3001 Bern
Switzerland
E-mail: stefan.wolter@vwi.unibe.ch

1. Introduction

> *"Some scientists and educational scholars find the concept of instructional time to be intellectually unexciting, so commonsensical, and of such obvious importance that it only leads to trivial understandings and to findings that have the status of truisms (e.g., students who spend more time studying learn more)… "*
>
> D.C. Berliner, 1990, p.3

Although David Berliner's assessment of the state of research dates back to the beginning of the 1990s, empirical contributions to questions of the use and effectiveness of instructional time have – with several recent exceptions – remained relatively limited (see Hanushek 2015). This is astonishing, given that instructional time is not only an important but, most importantly, a scarce resource in education production. The time that students can spend on education is limited by the hours in a day, the days in a week and the weeks in a school year. Additionally, every hour of instruction comes at a high cost, and as educational budgets are limited, the money spent on instruction detracts from other potential inputs in education production.

In addition to the importance of instructional time, both in terms of real and monetary input in the educational process, the lack of research is surprising for at least one additional reason. Total instruction time and the allocation of time to specific subjects vary greatly between countries,[1] but even a very superficial glance at the correlations between average instruction time and student test scores is sufficient to raise doubt that there is a simple relationship (Scheerens 2014) between the two. It may well be that educational authorities, believing that students that have more instruction time also learn more, have made intentional decisions when allocating more or less instruction time. However, the empirical question of whether additional instruction time is really used effectively or not remains an open one.

In this paper, we analyze both the effectiveness and the heterogeneity of the effectiveness of the use of instructional time, contributing to the existing literature in five ways. First, we will provide causal evidence of the average effect of instructional time, replicating the methodology applied by Lavy (2015) for Switzerland. Contrary to Lavy, who uses international data to estimate the impact of instruction time on test scores, we take advantage of the fact that due to the cantonal responsibilities for educational matters in Switzerland, instruction time shows considerable variance within that country. In contrast to international comparisons and their limitations (see e.g., Hanushek & Woessman, 2011

[1] According to the latest OECD statistics (2015), the average hours per school year in lower secondary education range from 754 (Sweden) to 1167 hours (Mexico) with an OECD average of 916 hours.

or Goldstein, 2004), this approach offers an advantage in that we can compare the impact of different uses of instruction time in a setting in which educational goals and curricula are very similar and therefore not a source of potential biases. Second, we control for potential misreporting by students by also using the official instruction times prescribed by educational authorities. Third, we refine the analyses by controlling for extra time spent on specific subjects either during school or after school (enrichment, remedial courses or paid private tutoring). Fourth, we do separate analyses for each school track; and fifth, we analyze the impact of additional instruction time on the within-school variance of student test scores.

In this paper, we analyze the effects of variations in subject-specific instruction time on student test scores, mainly originating from variations in the number of lessons taught per week. This is, however, not the only source of variation in instruction time. Instruction time can also vary considerably due to differences in the duration of lessons, the number of school weeks or even the number of school years. These latter sources of variation in instruction time may have very different impacts on student learning (see, for an overview of the literature, OECD 2016). Due to data limitations, we are only able to analyze the impact of cumulative instruction time (weekly or yearly) on student test scores and not the impact of differences in the organization of school days and weeks.

Our analyses show that one additional hour of instruction per week increases the PISA score by between 0.05 and 0.06 standard deviations. However, the returns of one additional hour vary greatly by school track. For students attending schools with advanced requirements, one extra hour of instruction increases the PISA score by between 0.07 and 0.08, while the increase in PISA score is only approximately 0.04 for students in tracks with basic requirements. These differences can be the consequence of many factors, such as different school environments, different teachers' attitudes, or different behavioural aspects, such as school discipline. However, the differences in effectiveness of instructional time can also be the consequence of differences in pupil aptitude (time needed to learn). If this is the case, then pupils with different abilities benefit to different extents from additional instruction time. Our results point in this direction, as we can show that additional instruction time also increases the within-school and subject variance of test scores.

This paper is organized as follows: the next Section provides a brief overview of the Swiss education system and of the 2009 PISA survey in Switzerland. Section 3 reviews the existing literature. Section 4 describes the data used in more detail, and section 5 describes the methods applied. Section 6 presents our empirical findings, and section 7 presents the paper's conclusions.

2. The Swiss Education System and PISA 2009

Switzerland provides an excellent laboratory for the study of educational policies for at least two reasons. First, the fact that the Swiss education system comprises different cantons[2] with independent educational policies leads to considerable variation in many relevant parameters of the education production function, including subject-specific instruction time. Second, despite having a high degree of freedom in making educational policy, the individual cantonal systems are under pressure to produce similar outcomes at the end of compulsory schooling (see SCCRE, 2014). This is due to the small geographical size of the country and the design of the post-compulsory education system. The two main reasons for this are that first, all baccalaureate schools in Switzerland lead to nation-wide, free access to all universities, and second, that most firms that offer apprenticeship training[3] operate nation-wide or regionally across cantonal borders. In other words, considerable variation in the design and operation of the education system, paired with high pressure to achieve comparable outcomes, make Switzerland the perfect environment to study the impact and effectiveness of school resources.

In order to analyze how hours of instruction affect pupils' performance, we use data from the fourth (2009) PISA test, conducted by the Organization for Economic Cooperation and Development (OECD). PISA is a standardized test administered to 15-year-old students who are enrolled in grades seven and above and who reside in OECD member countries and other participating countries. Students are assessed in three domains: reading, math and science. The sample is drawn using a two-stage stratification design. First, schools within the country are randomly selected. Second, a random sample of students is selected from within each school. In addition to the test results, PISA includes a student questionnaire with family and socio-economic background information and a school questionnaire with information on school type and school demographics.

In Switzerland, an additional representative sample of students in grade 9, the last year of compulsory education, was collected in 2009. This approach was employed in only 12 out of 26 cantons. We use this so-called national sample of 12 cantons in our analysis; first, because a comparison of students attending the same grade is more adequate for our purposes, and second, because the over-sampling in the national sample increases the number of observations considerably. For this national sample of 12 cantons, 13,605 students were interviewed and tested in 2009, and the average scores in reading, math and science were 502 points, 536 points and 517 points *(Konsortium PISA.ch, 2011)* with standard deviations of 88, 95 and 93 points.

[2] The Swiss cantons are comparable with US states, German Länder or the Canadian provinces in terms of their degree of autonomy in educational policy.
[3] Two thirds of the school leaving cohort do apprenticeship training.

3. Short Literature Review

Lavy (2015), the OECD (2016) and Rivkin and Schiman (2015) provide recent overviews of the literature. Therefore, we only highlight a selective number of studies that have a more or less direct link to our paper.

Lavy (2015) examined international gaps in student achievement, estimating the effects of instructional time using PISA 2006 data. The study exploited within-student and within-school variation by subject (reading, mathematics and science), estimating student fixed effects. He found that instructional time has a positive and significant effect on test scores. By comparing countries, he found evidence that the effect is much lower in developing countries and that the productivity of instructional time is higher in countries with such school characteristics as accountability measures and autonomy in budgetary decisions. Lavy estimated different samples, distinguishing by school tracking policies as well. The reported effects are significantly lower for schools without any tracking than for schools that practice some type of tracking. Overall, these results are in line with Lavy (2012), which estimated the effect of increasing the length of the school week and also subject-specific instructional time per week in Israel.

Building on the method used by Lavy (2015), Rivkin and Schiman (2015) ran similar regressions using the PISA 2009 data and adding controls for school quality. They determine that school circumstances are important determinants of the benefits of additional instruction time.

Results from a study based on the international student-level database TIMSS, using a cross-country setting and controlling for standard characteristics and for institutional characteristics of different schooling systems, show that instructional time is positively related to student performance as well (Woessmann, 2003). Mandel and Suessmuth (2011) estimated state-fixed effects within Germany for cumulative instructional time and found positive effects on student performance.

However, there are also studies finding no significant relationship between instructional time and school outcomes. Woessmann (2010) used cross-state variation in Germany to eliminate unobserved country-specific factors but did not find an effect on student test scores significantly different from zero. In addition, some studies looking at the length of school years and their impact on later earnings (Grogger 1996 or Pischke 2007) did not find statistically significant results. A study analysing the duration of the academic baccalaureate in Switzerland (Skirbekk 2006) estimated the impact of the canton-based variation of these programs on the TIMSS 2006 scores and did not find an effect of time spent in school on student achievement after controlling for school and student characteristics.

Several studies focusing on tracking and ability grouping provide evidence that gains in achievement among students differ from track to track (for an overview see Betts, 2011, Robertson & Symons, 2003). This relationship certainly does not constitute causality but still suggests that the effect of instructional time varies between tracks. Nonetheless, most of the research on tracking concentrates on inequality (Betts, 2011, Hanushek & Woessmann, 2006). The relation between instructional time and ability groups is less thoroughly examined. A study from Allensworth and Nomi (2009) estimated the efficacy of an algebra program in Chicago Public Schools providing a doubling of lessons to students in the 9th grade whose test scores in the 8th grade were below the national median. The authors provide evidence that the lowest-skilled students benefit less from additional lessons than do higher-skilled students. Because the students not only received more instruction but also more difficult coursework and improved instruction (as teachers had been taught how to use the extra time), these estimates jointly measure time and instruction quality characteristics.

For Switzerland, there is only a study from Angelone and Moser (2009) on instructional time and tracks. The authors estimated the effect of instructional time on performance for Swiss pupils with PISA 2006 data using an OLS regression. These researchers observed different effects among school tracks with different ability levels. Their results suggest that students in a high-level track benefit more from additional instruction time than students in a lower-level track, although methodological difficulties might cause biases.

4. Data

Apart from information on test performance, the PISA data set provides information on hours of instruction. Specifically, students were asked, "How many minutes, on average, are there in a <class period> for the following subjects?" and "How many <class periods> per week do you typically have for the following subjects?" We combined both questions with information on the number of school weeks per year per canton to create a variable for the number of instruction hours per year and the number of hours of instruction weekly. In order to make our results comparable to Lavy (2015), we do not use individually reported times for the estimations but the school average of individual answers. While school averages are less prone to misreporting, there might still be measurement error due to self-reporting and students' recall bias.

We try to overcome this using an additional source of information for mandatory hours of instruction, coming from official administrative data provided by the Swiss Conference of Cantonal Ministers of Education (EDK). This data source contains the number of lessons and minutes per lesson per subject by canton and type of school. These data should avoid some of the individual misreporting from the PISA data. However, there are still other potential issues to consider when using these data. First, we cannot be sure that each school and teacher actually follows the cantonal guidelines; and second, some schools might choose to offer extra instruction or optional lessons for some subjects. We believe that the true measure of hours of instruction is probably somewhere in between official and reported hours. This is the reason why we will use both measures in our analysis and compare the results with both.

Our initial sample consists of 13,605 9th graders. After deleting observations with missing values for hours of instruction, we have a final analytical sample of 11,433 students with an average reading score of 511 points, an average math score of 545 points and an average science score of 528. The average score in all subjects for our sample is slightly, although not statistically significantly, higher than for the whole sample.

The average hours of instruction are very similar in both data sets. However, the information differs significantly between PISA self-reported instruction time and the official cantonal guidelines for reading and science (See Table 1). A possible reason for the difference in science is that because the number of optional subjects offered by schools is larger in that area, pupils take more lessons than the minimum suggested by the EDK. In the case of reading, pupils' and authorities' definitions of reading lessons and what they entail might vary, leading pupils to think they have fewer hours than they actually have.

Insert Table 1 here

Most pupils in grade 9 have already been split into different school tracks. These tracks vary from canton to canton or even within a canton. In the majority of cantons, pupils are sorted into different school tracks after the 6th grade, according to their intellectual abilities. There can be between two and four different tracks, but mostly three tracks are used: a track with advanced requirements, which teaches the more intellectually demanding courses; a track with intermediate requirements; and finally, one track offering basic-level courses. All tracks teach general skills, but at different levels of academic requirements. In some cantons, there is no formal splitting into tracks, but pupils might still be sorted into different ability groups depending on the subject, especially for reading and math.

5. Empirical Strategy

We will estimate the following model to calculate the effect of instructional time on PISA scores:

$$(1) \qquad y_{ijk} = \beta_j I_{jk} + \varkappa Z_i + \gamma X_{ij} + \psi S_k + a_i + \lambda_j + \sigma_k + \mu_{ijk}$$

where y_{ijk} is the score of pupil i in subject j in school k, I_{jk} is instructional time for subject j in school k, X is a vector of individual characteristics that vary by subject, Z is a vector of individual characteristics that are the same over all subjects and do not vary across schools, and S is a vector of school characteristics. α_i, σ_k and λ_j represent unobservable characteristics of the individual, the school and the subject, respectively. μ_{ijk} is the unobserved error term.

One of the greatest difficulties when estimating the effect of hours of instruction on test scores is the possibility that unobservable individual and school characteristics might be correlated with the number of school lessons. Following Lavy (2015), we take advantage of the PISA data structure and estimate within-pupil effects of instructional time on test scores. This method accounts for potential confounding factors such as students' ability and school quality.

In order to identify the effect of hours of instruction on score, we need to assume that the effect of one hour of instruction is the same for all subjects.[4] In order to test if this assumption holds, we follow Metzler and Woessmann (2012) and model the unobserved pupil effects as

$$(2) \qquad a_i = \delta_{read} I_{readk} + \delta_{math} I_{mathk} + \delta_{science} I_{sciencek} + \nu X_{iread} + \nu X_{imath} + \nu X_{iscience} + \tau Z_i + \varphi S_k + \omega_i$$

The δ parameters are allowed to differ across subjects, but we assume that the coefficients of the X characteristics are the same across subjects. Substituting equation 2 into equation 1 and rearranging the terms, we obtain

(3)

$$y_{ijk} = (\beta_j + \delta_j)\, I_{jk} + \textstyle\sum_{s\neq j} \delta_s I_{isk} + (\tau+\varkappa)\, Z_i + (\lambda+\upsilon)\, X_{ij} + \textstyle\sum_{s\neq j} \upsilon X_{is} + (\psi+\varphi)\, S_k + a_i + \lambda_j + \sigma_k + \omega_i + \mu_{ijk}$$

[4] Lavy (2012, 2015) also suggests that this assumption holds.

This allows us to test whether $\delta_{read} = \delta_{math} = \delta_{science}$ and whether $\beta_j = \beta_s$ for all $s \neq j$, i.e., whether the relationship between hours of instruction and the individual unobservable term is the same for all subjects, and whether the effect of one hour of instruction on PISA scores is the same for all subjects. After estimating model (3), as suggested in Chamberlain (1982), we cannot reject the hypothesis that the coefficients are the same (see Table A1 in the appendix). Therefore, we estimate the following restricted fixed effects model:

$$(4) \qquad y_{ijk} = \beta \ddot{I}_{ik} + \gamma \ddot{X}_{ij} + \lambda_j + \ddot{u}_{ijk}$$

where $\ddot{y}_{ijk} = y_{ijk} - \sum_j y_{ijk} / 3$ and similarly for $\ddot{I}_{jk}$, $\ddot{X}_{ij}$ and $\ddot{u}_{ijk}$

6. Results

In this chapter, we present our empirical results that examine the average impact of instruction time on student performance (6.1), the impact of instruction time per school track (6.2) and the impact of instruction time on the within-school variance of student performance (6.3), providing some robustness checks at the end.

6.1 Impact of instruction time on student performance

The results of our replication of the Lavy (2015) results are presented in Table 2. We standardize the PISA score using the national mean of 518 and standard deviation of 92.[5] Lavy (2015, p. F408) reports a student fixed-effect result of 0.058 standard deviations for an additional hour of instructional time. Our empirical results show exactly the same coefficient for Switzerland as in Lavy when using the self-reported data on school hours but a significantly smaller coefficient when the official cantonal data are used: 0.046 standard deviations instead of 0.058.[6,7]

If we make a back-of-the envelope calculation comparing the effect of an average hour of instruction with the effect of an additional hour of instruction, we find that an additional hour has only between a third and two fifths of the effect of an average hour on PISA scores. According to the OECD, an average student should progress by 42 PISA points

[5] Lavy uses the international mean of 500 and standard deviation of 100.
[6] We also controlled for the fact that learning can be a cumulative process (See Mandel and Suessmuth, 2011) by estimating the regression using official data on cumulated hours of instruction between grades 7th and 9th. The coefficient of hours of instruction is still positive and significant, even though smaller at about 0.01.
[7] We also include controls for attending enrich, remedial or extra private lessons.

per year of schooling. On average (all subjects), a Swiss student had 3.4 hours of subject specific weekly schooling. Therefore, if an average Swiss student would also progress by 42 PISA points per year, an average weekly hour of instruction would generate a gain of 12.35 PISA points. Our FE results indicate, however, only a gain of between 4.2 and 5.3 PISA points for an additional hour of instruction.

Insert Table 2 here

6.2 Impact of instruction time on student performance per track

Switzerland is one of the countries where tracking takes place rather early. Although it is now taking place after the 6th grade, in earlier decades, certain cantons already tracked students after the 4th grade. This feature of the Swiss system allows us to compare the efficiency of one extra hour of instruction for pupils with different academic abilities and school environments.

Even if one hour of instruction has the same productivity for all subjects, it may be that different students profit differently from extra instructional time, as suggested by the descriptive statistics in Table 1. In order to control whether students with different academic skill levels have different returns on instructional time, we divided the sample into students in upper-level school types, students in intermediate-level school types and students in basic-level school types.

Results in Table 2 show that one additional hour of instruction has, in fact, a significantly higher effect for pupils in more demanding types of school.[8] This result holds independently of the data source for instruction time. We can think of two possible explanations for this: First, pupils in lower-level tracks might indeed have lower learning capabilities. Second, the inputs in the different tracks might be different. This difference may be the case if, for example, teachers are less qualified or the discipline is worse in the lower tracks. This finding would be in line with results from Rivkin and Schiman (2015), which show that returns to additional time depend on the quality of learning environment and on the quality of teachers and schools.[9]

The PISA data set provides some information in the form of indexes for teacher/student relations, discipline climate, student-related factors affecting school climate and teacher-related factors affecting school climate (for a detailed description of the indexes, see

[8] Lavy (2015, p. F410) presents the estimated effects of hours of instructional time by tracking status, comparing systems that have no tracking at all with systems that track by class or within class. This is not exactly the same as in our analyses, where we compare the effects between track levels.

[9] There are several studies that show direct relation between classroom climate and student achievement, see Benner, Graham and Mistry (2008) as well as Arens, Morin and Watermann (2015).

OECD, 2010). These indexes could be used to analyze different potential inputs and learning environments among different school tracks.

Table A2 in the appendix presents the different indices by school track. Two of the measures, student-related factors affecting school climate and teacher-related factors affecting school climate, were constructed using questions from the school questionnaire. The rest were constructed using questions from the student questionnaire. There is some indication that the learning environment and other inputs are worse in basic-level school tracks. All values of the indices are significantly worse in lower-level school tracks, with the exception of student/teacher relations, whereas the intermediate school tracks are better than the lower-level tracks but below the values of the upper-level tracks. The values of disciplinary climate and student-related factors affecting school climate, especially, are significantly lower in the lower-level tracks. This might be one explanation for why the efficiency of one hour of instruction is lower in the basic-level tracks. Possibly, students' disruption in class, noise and disorder, and the fact that the teacher has to wait a long time for students to quiet down decreases the actual time that pupils spend learning. This delay, added to the increased difficulty of concentrating, reduces what students might be able to learn in the given instructional time (Lazear, 2001). Our empirical analyses show that interactions of these indices with instruction time are mostly statistically non-significant, and in the cases where we find significances, they do not change the coefficient of the hours of instruction (see table A3 in the appendix).

A related issue is that tracks may differ in terms of teaching styles. Teachers in lower tracks need to prevent some pupils from failing to reach even basic competencies by the end of compulsory school. Therefore, teachers have an incentive to focus on the weakest pupils by reducing the intensity and speed of instruction. In the highest track, however, the goal is to make pupils ready for academic baccalaureate schools, and therefore the focus is on excellence and on the brightest pupils. If some pupils are not able to follow the increased speed of instruction, they can always be re-directed to the intermediate track, where teaching style and speed focus on the median pupil.

6.3 Impact of instruction time on the variance of student performance within schools

It is easy to understand why school and class climate potentially affect the effectiveness of instruction time. Another source that could potentially explain the differences in effectiveness of instruction hours between tracks, however, could be differences between students in time needed to learn. Students are sorted into the different tracks due to their different potentials or abilities; therefore, it is possible that students in the highest tracks can make more productive use of an hour of instruction than less talented students sorted into the lowest tracks can – all other things being equal.

If time needed to learn is the explanation for differences in the effectiveness of instructional time between different ability tracks, then we can also assume that the effectiveness of instructional time also differs within tracks and schools. In this case, additional instructional time would not only increase the average performance but also increase the performance gaps between students in the same class or school. Of course, this assumption depends on how extra time would be used by teachers. If teachers used the extra hours to compensate the weaker students and to help them catch up with the better students instead of teaching additional or new content, extra time would produce more homogenous results.

We investigate the last issue by regressing heterogeneity in school performance on number of hours. Once again, we exploit the within-school variation in performance heterogeneity and instructional time among the three different subjects. This allows us to identify the effect by keeping constant all unobservable school characteristics, such as those that might be correlated with the number of hours. In order for this to work, we need to assume that the effect of one additional hour on the score gap between best and worst students is the same for all subjects.

Insert Table 3 here

Table 3 shows that one extra hour of instruction increases the score variance by almost 150 points when using PISA information and almost 110 when using official cantonal hours. Increasing hours of instruction does not, however, affect the skewness of the results within a school, which could be expected if additional time were used specifically for a particular part of the ability distribution of students, either by only focusing on the weakest or by using the extra time for additional content for the gifted students. The considerable impact on the within-school variance and the non-effect on the skewness of the results are therefore hints that teachers gravitate toward the median students in their teaching.

A potential problem might be the existence of simultaneity in determining hours of instruction and score heterogeneity, if cantons or schools with more heterogeneity increase the instructional time in order to reduce the score gap.[10] We do not think that this would be an issue in this case. Cantonal authorities regulate the number of hours, and therefore schools do not have much freedom. At the same time, cantonal authorities cannot easily change the number of hours already set. On the one hand, there are the

[10] As the PISA test is based on a school sample and not a class sample, the sample of bigger schools might comprise students of different tracks. In order to test whether this would bias our results, we also ran our regressions with mono-track schools only. The number of observations is considerably smaller, but the coefficients are almost identical.

teachers' unions who advocate teachers' interests, and on the other hand, there is the financial aspect of increasing expenditures. We therefore conclude that increasing the number of hours of instructional time increments performance heterogeneity.

6.4 Robustness checks

It is possible that the relationship between PISA score and hours of instruction may not be linear. In order to investigate potential non-linearities, we estimate a model including an hours of instruction square term that allows us to analyze whether there are diminishing returns to additional instruction time. Second, we estimate another model including hours of instruction as categorical variables. The distribution of students by number of hours of instruction and by subjects is presented in Table 4. The three categorical variables for hours of instruction were created as follows: one binary variable equal to one if the time of instruction is up to 3.5 hours, the second binary variable equals one for students that report between 3.5 to 4.5 hours, and the third one equals one for students who report more than 4.5 hours.

Insert Table 4 here

The intervals were constructed in this way because, except in the case of science, the number of hours does not vary considerably. For reading and math, the instruction time is between 3 and 5 hours.

Overall, the number of hours of instruction reported by students in the PISA data set and the number of mandatory hours set by the cantonal authorities are similar. However, there are large differences by subject.

Insert Table 5 here

Results of the regressions, including hours of instruction square and hours of instruction in categorical form, are presented in Table 5. The coefficient of hours of instruction square in Specifications 1 and 2 are negative, but the significance depends on the data used; therefore there is no clear indication of diminishing returns to instruction time. Moreover, the non-linearity appears beyond the number of hours of interest at approximately 7-8 hours per week. Specifications 3 and 4 include binary variables for instruction time. The results using official data and PISA data do not differ significantly.

The results using PISA and cantonal data would suggest that returns to more than 4.5 hours are considerably smaller than returns to between 3.5 and 4.5 hours.

7. Conclusions

Summarizing our results, we would like to highlight three important findings:

First, our replication of Lavy's (2015) use of subject variation in instruction time to assess the causal impact of instruction time on student test scores exhibits remarkably similar results in the context of an analysis for one country. Controlling for additional information on extra in or out of school hours spent for remedial or enrichment purposes and private tutoring does not affect the result significantly. However, using official, prescribed school hours instead of self-reported data lowers the coefficients of instruction time significantly. Although we find a difference between the two measures of instruction time, we cannot say whether one is better than the other. Self-reported data are obviously prone to measurement errors, but prescribed hours also might not adequately depict the reality in schools.

Second, in all of our specifications, we identify a significant impact of additional instructional time on learning outcomes measured with PISA test scores, but the effectiveness of an additional hour of instruction is only between thirty to forty percent of the impact that we would expect from an average hour of instruction. In other words, variations of instructional time should be considered carefully by educational authorities, as the marginal gains from more instruction time might be to low compared to alternative uses of time and to the financial resources needed for additional instruction time.

Third, the heterogeneity of the effectiveness of instructional time between different ability tracks and the impact of additional instruction time on the within-school variance of student test scores puts the individual student time needed to learn at the centre of attention. John Carroll, when revisiting the literature and findings that had emerged after he developed what later became known as the Carroll Model, stressed the same idea:

> *"The model's emphasis on aptitude as a determinant of time needed for learning suggests that increased efforts be placed on predicting student potentialities and designing instruction appropriate to those potentialities, if ideals of equal opportunity to learn are to be achieved within a diversity of educational objectives."* J.B. Carroll, 1989, p. 26

Interestingly, in discussions with educational practitioners or policy makers, the desire for additional instruction time is often motivated by the argument that more time is needed to close the gap between the performance levels of individual students. Our results support this view, insofar as we find indications that less able students would indeed need more instruction time to achieve similar results to those of more able students. However, our additional finding that more instruction time increases the variance of test results, rather than closing the gap, reveals that apparently, the additional instruction time in schools is not used to compensate the weak. Therefore, the use of instruction time in schools has to be reconsidered if the achievement of more equal results is sought.

References

Allensworth E. and T. Nomi (2009). Double-dose algebra as an alternative strategy to remediation: Effects on students' academic outcomes, *Journal of Research on Educational Effectiveness*, March, 111-148.

Angelone, D. and Moser, U. (2012). More hours do not necessarily pay off. The effect of learning time on student performance at different school types in Switzerland. In M. Prenzel, M. Kobarg, K. Schöps & S. Rönnebeck (eds.), Research on PISA. Research Outcomes of the PISA Research Conference 2009 (pp. 129-144). Dordrecht: Springer.

Arens, A.K., Morin, A.J.S., and Watermann, R. (2015). Relations between classroom disciplinary problems and student motivation: Achievement as a potential mediator?, *Learning and Instruction*, 39, 184-193.

Benner, A. D., Graham, S., and Mistry, R. S. (2008). Discerning direct and mediated effects of ecological structures and processes on adolescents' educational outcomes, *Developmental Psychology*, (3), 840–854.

Berliner, D.C. (1990). What's all the fuss about instructional time?, In: M. Ben-Peretz & R. Bromme (Eds): The Nature of Time in School. Theoretical Concepts, Practitioners Perceptions. New York: Columbia University Teachers College Press: 3-35.

Betts, J.R. (2011). The economics of tracking in education. In: E. A. Hanushek, S. J. Machin, & L. Woessmann (eds.), Handbook of the Economics of Education (Bd. 3, S. 341–381). Amsterdam: Elsevier.

Carroll, J.B. (1989). The Carroll Model: A 25 year-retrospective and prospective view, *Educational Researcher*, 18(1), 26-31.

Chamberlain, G. (1982). Multivariate regression models for panel data, *Journal of Econometrics*, 18, (1), 5-46

EDK [Swiss Conference of Cantonal Ministers of Education] (2008). Stundentafeln der Volksschule. Bern: Schweizerische Konferenz der kantonalen Erziehungsdirektoren.

Goldstein, H. (2004). International comparisons of student attainment: some issues arising from the PISA study, *Assessment in Education*, 11(3), 319-330.

Grogger, J. (1996). Does school quality explain the recent black/white wage trend?, *Journal of Labor Economics*, vol. 14, 231-253.

Hanushek, E.A. (2015). Time in education: introduction, *The Economic Journal*, Volume 125, Issue 588, F394-396.

Hanushek, E.A. and Woessmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries, *The Economic Journal*, 116(150), 63-76.

Hanushek, E.A. and Woessmann, L. (2011). The economics of international differences in educational achievement. In: E. A. Hanushek, S. J. Machin, & L. Woessmann (Eds.), Handbook of the Economics of Education (Bd. 3, pp. 89–200). Amsterdam: Elsevier.

Konsortium PISA.ch (2011). PISA 2009: Regionale und kantonale Ergebnisse. Bern und Neuchâtel: BBT/EDK und Konsortium PISA.ch

Lavy, V. (2012). "Expanding school resources and increasing time on task: effects of a policy experiment in Israel on student academic achievement and behavior", NBER Working Paper Number w18369.

Lavy, V. (2015). Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries, *The Economic Journal*, Volume 125, Issue 588, F397–F424.

Lazear, E.P. (2001). Educational production, *The Quarterly Journal of Economics*, 116 (3) (August 1): 777–803. doi:10.1162/00335530152466232.

Mandel, P., and B. Suessmuth (2011). Total instructional time exposure and student achievement: An extreme bounds analysis based on German state-level variation. CESifo Working Paper Series. CESifo Group Munich. http://ideas.repec.org/p/ces/ ceswps/_3580.html.

Metzler, J. and Woessmann L. (2012). The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation, *Journal of Development Economics,* 99, 486-496.

OECD (2010). PISA 2009 Results: What makes a schools successful? - Resources, policies and practices (Volume IV). Paris: OECD.

OECD (2015). Education at a glance, Paris: OECD.

OECD (2016). Student learning time: A literature review, OECD Education Working Paper No. 127, Paris: OECD.

Pischke, J.-S. (2007). The impact of length of the school year on student performance and earnings: evidence from the German short school years, *The Economic Journal*, Vol. 117(4), 1216-1242.

Rivkin, S.G., & Schiman, J.C. (2015). Instruction time, classroom quality, and academic achievement, *The Economic Journal*, Vol. 125, Issue 588, F425–F448.

Robert, D. and Symons, J. (2003). Do peer groups matter? Peer group versus schooling effects on academic attainment, *Economica*, 70, 31–53.

SCCRE (2014). Swiss Education Report 2014, Aarau: Swiss Coordination Centre for Research in Education.

Scheerens, J. (2014). Effectiveness of time investments in education. Insights from a review and meta-analysis, SpringerBriefs in Education, DOI 10.1007/978-3-319-00924-7.

Skirbekk, V. (2006). Does school duration affect student performance? Findings from Canton-based variation in Swiss educational length, *Swiss Journal of Economics and Statistics*, 142(1), 115-145.

Woessmann, L. (2003). Schooling resources, educational institutions and student performance: The international evidence, *Oxford Bulletin of Economics and Statistics*, 65, 117-170.

Woessmann, L. (2010). Institutional determinants of school efficiency and equity: German states as a microcosm for OECD countries, *Journal of Economics and Statistics* (Jahrbücher für Nationalökonomie und Statistik), 230, 234-270.

Tables

Table 1: Weekly hours of instruction by subject and data set

	Self reported hours (PISA Data)				
	Mean	Std. Error	Lower Track	Middle track	Upper track
All subjects	3.40	0.005	3.45	3.44	3.35
Reading	3.85	0.01	3.84	3.81	3.90
Math	3.72	0.004	3.74	3.74	3.68
Science	2.55	0.01	2.58	2.69	2.41
Mean Score	528	0.43	460	526	573
Score reading	511	0.71	441	508	559
Score math	545	0.74	477	544	591
Score science	528	0.73	460	524	569
N	11433				
	Official hours (Council of Cantonal Ministers)				
	Mean	Std. Error	Lower Track	Middle track	Upper track
All subjects	3.40	0.01	3.41	3.42	3.37
Reading	3.90	0.01	3.85	3.87	3.98
Math	3.73	0.004	3.66	3.74	3.76
Science	2.48	0.01	2.53	2.58	2.31
N	11433				

Notes: In order to standardized the number of hours per weeks in different schools we multiplied the minutes per week times the number of weeks in the canton. Afterwards we divide this by 38, which is the mean number of weeks in Switzerland.

Table 2: Effect of hours of instruction on PISA test scores

	Whole sample	Low skills track	Middle skills track	High skills track
	Self reported hours (PISA Data)			
Hours of instruction	0.058	0.038	0.060	0.080
	(0.007)	(0.006)	(0.004)	(0.005)
	Official hours (Council of Cantonal Ministers)			
Hours of instruction	0.046	0.034	0.043	0.062
	(0.005)	(0.004)	(0.004)	(0.004)
n	11433	2870	4103	4460
N	32411	7851	11637	12923

Notes: Standard errors in parentheses are clustered at the school level. The regressions control for subject and additional hours of instruction taken, which include enrichment, remedial and private tutoring lessons. Scores are standardized using the mean and standard deviations for Switzerland of 518 and 92 respectively.

Table 3: Effect of hours of instruction on the heterogeneity in PISA test scores

	Self reported (PISA)		Official hours	
	Variance	Skewness	Variance	Skewness
Hours of instruction	149.21	0.01	108.28	0.01
	(36.80)	(0.01)	(27.94)	(0.01)
N	321		321	

Notes: Standard errors in parentheses. The regressions include subject fixed effects.

Table 4: Proportion of students by subject number of hours

	Self reported (PISA Data)		
	<3.5 Hours	3.5-4.5 Hours	More than 4.5 hours
All subjects	0.48	0.44	0.08
Reading	0.30	0.52	0.18
Math	0.29	0.69	0.02
Science	0.83	0.12	0.05
N	11433		
	Official hours		
	<3.5 Hours	3.5-4.5 Hours	More than 4.5 hours
All subjects	0.42	0.44	0.14
Reading	0.24	0.47	0.29
Math	0.23	0.76	0.01
Science	0.78	0.09	0.13
N	11433		

Table 5: Effect of hours of instruction on PISA test scores, controlling for non-linearities

	Self reported (PISA Data)		Official hours	
	(1)	(2)	(1)	(2)
Hours of instruction	0.110		0.142	
	(0.033)		(0.023)	
Hours of instruction squared	-0.008		-0.013	
	(0.005)		(0.003)	
3.5-4.5 Hours		0.093		0.106
		(0.015)		(0.017)
>4.5 Hours		0.089		0.105
		(0.021)		(0.014)
n	11433			
N	32411			

Notes: Standard errors in parentheses are clustered at the school level. The reference category for the hours of instruction is less than 3.5 hours. Scores are standardized using the mean and standard deviations for Switzerland of 518 and 92 respectively. n= 12010

Appendix

Table A1: Test for equality of effects following Metzler and Woessmann, (2012)

	Reading	Math	Science
Hours of instruction reading	0.009	-0.003	-0.101
	(0.063)	(0.064)	(0.061)
Hours of instruction math	0.034	0.080	0.074
	(0.051)	(0.063)	(0.057)
Hours of instruction science	0.030	0.035	0.034
	(0.028)	(0.032)	(0.030)
Tests			
	Chi square	P value	
$\beta_{reading}= \beta_{math}= \beta_{science}$	0.57	0.75	
$\delta_{read}=\delta_{math}$	3.70	0.06	
$\delta_{read}= \delta_{science}$	0.46	0.50	
$\delta_{science}=\delta_{math}$	0.01	0.94	

Notes: Standard errors in parentheses are clustered at the school level.
The regressions were estimated using PISA data on hours of instruction. The regressions include controls for individual background characteristics such as gender, parental education, immigration status, language spoken at home, and for school characteristics.

Table A2: Input indexes by school track

	Low skills track	Middle skills track	High skills track
Attitudes towards school	-0.04	-0.02	0.04
	(0.01)	(0.01)	(0.01)
Teacher/student relations	0.11	0.18	0.17
	(0.01)	(0.01)	(0.01)
Disciplinary climate	-0.04	0.10	0.16
	(0.01)	(0.01)	(0.01)
Teacher related factors affecting school climate	0.13	0.12	0.24
	(0.01)	(0.01)	(0.01)
Student related factors affecting school climate	0.02	0.09	0.23
	(0.01)	(0.01)	(0.01)

Notes: Standard errors in parentheses.
All indexes have a mean of 0 and a standard deviation of 1. Higher values indicate a better performance in the index.

Table A3: Effect of hours of instruction on PISA test scores controlling for school climate

	Self reported hours	Official hours
Hours of instruction	0.058	0.046
	(0.007)	(0.005)
Attitudes towards school*hours of instruction	0.007	0.005
	(0.002)	(0.001)
n	11433	

Notes: Standard errors in parentheses are clustered at the school level. The regressions control for subject and additional hours of instruction taken, which include enrich, remedial and private tutoring lessons. Scores are standardized using the mean and standard deviations for Switzerland of 518 and 92 respectively. Attitudes towards school is an index that includes answers to the following questions: School has done little to prepare me for adult life when I leave school , School has been a waste of time, School has helped give me confidence to make decisions, School has taught me things which could be useful in a job.