Labour Economics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Increased instruction hours and the widening gap in student performance

Mathias Huebener[a,b,*], Susanne Kuger[c], Jan Marcus[a,d]

[a] DIW Berlin, Germany
[b] Freie Universität Berlin, Germany
[c] DIPF Frankfurt, Germany
[d] Universität Hamburg, Germany

ARTICLE INFO

JEL classification:
I24
I28
J24

Keywords:
Instruction time
Student achievement
PISA
G8-high school reform
Curriculum
Quantile difference-in-differences
Recentered influence function

ABSTRACT

Do increased instruction hours improve the performance of all students? Using PISA scores of students in ninth grade, we analyse the effect of a German education reform that increased weekly instruction hours by two hours (6.5 percent) over almost five years. In the additional time, students are taught new learning content. On average, the reform improves student performance. However, treatment effects are small and differ across the student performance distribution. Low-performing students benefit less than high-performing students. We argue that the content of additional instruction time is an important determinant explaining this pattern. The findings demonstrate that increases in instruction hours can widen the gap between low- and high-performing students.

1. Introduction

Increasing the time that students spend in the classroom has moved into the policy focus in OECD countries. In the UK and the US, it is a central element of education policy agendas (OECD, 2016a). Policymakers raise two main arguments for increasing school instruction time: First, more instruction time could improve overall student performance by providing more learning opportunities. Second, it could help narrow performance gaps between low- and high-performing students by compensating for lacking resources or supervision outside school (OECD, 2016b). Despite the high hopes of policymakers and the high costs of instruction time as a school input factor, the question of whether spending more time in the classroom can effectively improve student performance has received surprisingly little research attention (Patall et al., 2010; Lavy, 2015; OECD, 2016b). Even less is known about how additional classroom time should be spend and how the effects differ between low- and high-performing students.

In this paper, we study the impact of an increase in weekly instruction time on student performance induced by a large education reform in German academic track schools. The so-called G8-reform reduced the length of academic track schooling by one year, while increasing instruction hours in the remaining school years such that students will have covered a similar curriculum when they graduate from school in one year less. We focus on the performance of students in ninth grade, when they are typically 15 years old. These students are only affected by the additional instruction hours, but not yet by the reduced length of schooling. An important feature of the increased instruction time is that it covered more learning content. The reform serves as a natural experiment to estimate the effect of spending two additional instruction hours per school week (+ 6.5 percent) in the classroom from grade 5 to grade 9, i.e. between the ages of 11 and 15. The additional instruction time totals to about 350 hours. Our analyses rely on data from the Programme for International Student Assessment (PISA), pooled across five waves from 2000 through 2012. The reform was implemented with regional and temporal variations in only one school track, which we exploit in linear difference-in-differences models to estimate average treatment effects, as well as in non-linear difference-in-differences models to estimate quantile treatment effects.

Estimates of the average treatment effects suggest that the reform increased PISA test scores of ninth graders in reading, mathematics, and science by 5 to 6 percent of an international standard deviation. The estimated quantile treatment effects reveal that the bottom of the student performance distribution shows almost no effects, while treatment effects increase further up in the performance distribution. This widening gap between low- and high-performing students is most pronounced in mathematics and science. Our findings are robust to

* Corresponding author at: DIW Berlin, Department Education and Family, Mohrenstraße 58, 10117 Berlin, Germany.
E-mail address: mhuebener@diw.de (M. Huebener).

http://dx.doi.org/10.1016/j.labeco.2017.04.007
Received 1 November 2016; Received in revised form 31 March 2017; Accepted 27 April 2017

various model specifications. Different placebo regressions support the main identification assumption.

This study contributes to the previous literature in three important aspects: First, we study a policy experiment in which additional classroom time was devoted to additional learning content rather than the same content. This is a highly relevant policy experiment, as policymakers are typically referring to more instruction hours covering more learning content when they discuss increases in instruction time to improve student performance. Second, many previous studies rely on small and short-lived exogenous changes in instruction time to estimate the effects on student performance (e.g. Marcotte, 2007; Sims, 2008; Fitzpatrick et al., 2011; Herrmann and Rockoff, 2012; Goodman, 2014; Carlsson et al., 2015; Aucejo and Romano, 2016). Only a few studies generate insights from considerable, policy-induced increases in instruction time, and they are often accompanied by changes in other school input factors or the peer environment (Bellei, 2009; Lavy, 2012; Cortes and Goodman, 2014; Taylor, 2014; Cortes et al., 2015). Our study exploits a policy reform within the same school environment and peer environment that led to a substantial and lasting increase in instruction hours from a level close to the OECD average (OECD, 2015). Third, the previous literature mostly focuses on average treatment effects of instruction time. Differential effects by student ability received less attention (exceptions are Bellei, 2009; Carlsson et al., 2015; Cattaneo et al., 2016), but they are very relevant from a policy perspective. Increases in instruction time with additional learning content may have different effects on students depending on their capabilities of understanding and processing new learning content. We estimate such effects across the performance distribution and address this gap in the literature.

We conclude that (i) additional instruction time improves average student performance; (ii) the effect sizes are rather small given the substantial increase in instruction time; and (iii) the student performance distribution widens, especially in mathematics and science. That the increased instruction time is spent on new learning content seems to be crucial for explaining why effect sizes are small on average, and why they increase as one moves up the performance distribution. Students' existing set of skills may be important in transforming instructional input into student performance: Lower-performing students might need more time than better-performing students to process new learning content. When policymakers decide about additional classroom time, they should be aware of the potential to widen gaps in student performance when new learning content is added to the curriculum.

Previous studies on the G8-reform mainly analyse the joint effect of fewer years of schooling and additional weekly instruction hours (see Huebener and Marcus (2015) and Thomsen (2015) for overviews of these studies). Dahmann (2017) is an exception: She analyses the G8-reform effect on fluid and crystallised intelligence. Comparing students at age 17 (with different levels of instruction time) in survey data, she finds positive reform effects on crystallised intelligence of boys, but not for girls. At the end of academic track schooling, after treated students attended one year less of schooling, she finds no reform effects. In our study, we focus on a different set of outcomes and look at the effects of additional instruction time on student performance at age 15 in the three PISA domains of reading, mathematics, and science. The domain-specific effects are important because policymakers have an interest in learning about effective ways to improve student competencies in certain domains. Another distinct feature of our study is that we show differential effects across the performance distribution. Two further working papers also examine the effects of the G8-reform in PISA data (Andrietti, 2016; Andrietti and Su, 2016). The work has been developed independently and at the same time. The combined statistical findings of both these working papers are similar to our core findings. Additionally, we conducted extensive archival research on official timetable regulations, as decreed by the education ministries of each federal state, allowing us to determine the exact, subject-specific instruction hour increase induced by the G8-reform, which is not provided in previous work on the reform. Furthermore, we examine numerous other channels, in addition to instruction time, through which the reform may impact student performance. In addition, we draw on another, large data set of teachers and study reform adjustments in the teacher body.

The remainder of this study is organised as follows. Section 2 reviews the related literature. Section 3 describes the institutional setting and the school reform from which we derive our findings. Section 4 introduces the data and outlines the empirical approach. We report the main findings in Section 5, and check the sensitivity of the findings and potential reform channels in Section 6. Section 7 concludes.

2. Related literature

Understanding the effectiveness of school input factors in increasing student performance is important for policymakers allocating resources. The effectiveness of instruction time in increasing student performance has received little attention, even though classroom time is an omnipresent, easy-to-manage, but also costly input factor in education systems (Patall et al., 2010; Lavy, 2015; OECD, 2016b).

The challenges involved in identifying the causal effects of instruction time on student performance may be one reason. Some studies correlating student performance with instruction time in cross-sectional data find at most small positive, but not robust, relationships (Card and Krueger, 1992; Grogger, 1996; Lee and Barro, 2001; Woessmann, 2003). Yet, observed cross-country correlations might be confounded by other features of education systems. In individual-level data, it is students' endogenous selection into more or less instruction time that poses challenges for the identification of causal effects. Lower-performing students might receive additional instruction hours in order to revise and understand the classroom content. Better-performing students might select additional courses in subjects they like the most. With the availability of better data sources in education research (Machin, 2014), new approaches can be applied to address this challenge.

To address endogeneity problems, two approaches dominate the literature on this topic. The first looks at within-student variation in subject-specific instruction time. For instance, Lavy (2015), Rivkin and Schiman (2015), and Cattaneo et al. (2016) use cross-subject variations in instruction time and control for time-invariant, student-specific characteristics in student-fixed effects models. In contrast to previous correlation analyses, these studies find a strong positive effect of instruction hours on student achievements. Despite the advantages of this econometric approach, it assumes that only classroom time in a certain subject affects the performance of students in the respective subject, i.e. spillovers between subjects do not exist. As these studies typically relate the current level of instruction hours to student performance, little is known about both the effect of instruction hours in earlier grade levels on current performance and about the learning content of additional time in school.

The second approach exploits quasi-experimental settings to learn about causal effects of instruction time on student performance. Marcotte (2007), Marcotte and Hemelt (2008) and Goodman (2014) use variation in winter weather that affected school instruction time prior to centralised exams. Sims (2008), Fitzpatrick et al. (2011) and Carlsson et al. (2015) use school day variations induced by quasi-random assignments of school start dates or assessment dates. Herrmann and Rockoff (2012) and Aucejo and Romano (2016) identify the effects with random variations in student and teacher absence days. These quasi-experimental studies find mostly beneficial impacts of more instruction time. Although the content of the additional classroom time is not explicitly stated, one can think of these studies as identifying the effects of spending varying amounts of time on a fixed