On the importance of school-based inputs in the production of student achievement: Evidence in a recent Scandinavian context

Anne Brink Nandrup[1]
Department of Economics and Business Economics, Aarhus University,
email: *anandrup@econ.au.dk*

Summary: This paper summarises and interprets findings from the recent Scandinavian and international literature on the effects of school inputs in the general framework of a production function for cognitive achievement. Although the literature previously has been somewhat inconclusive, evidence using improved methodology and research designs largely supports an emerging consensus that direct school-based inputs may be adjusted to improve student achievement. Simultaneously, a growing body of research finds that parents respond to changes in school inputs, which may crowd-out the effect of additional school resources.

Keywords: Knowledge production function; Instruction time; Class size; Teacher aide; Parental response; School resources

JEL: D1; H52; I20; I28

1. Introduction

The determinants of student achievement have long been a priority of – not only scholars within the social sciences – but also the public in general. Demonstrating that variation in per student spending was unrelated to variation in student

1. This version: June 2017. This paper is produced in connection with the semi-annual meeting of Nationaløkonomisk Forening at Koldingfjord, January 2016. The author gratefully acknowledges helpful comments from and discussions with Marianne Simonsen and Morten Krægpøth. All errors are my own.

achievement in the 1960s, the Coleman Report (Coleman et al. 1966) sparked an extensive multi-disciplinary literature concerned with the link between school inputs and student achievement. A literature that was long characterised by a somewhat surprising lack of consensus over the impact of increased school resources in the production of student outcomes (Hanushek 2003, Krueger 2003; Todd and Wolpin 2003, 2007). Due to improved methodology and research designs in recent years, however, the literature has progressed substantially. Importantly, generous educational investments and unique data records imply that evidence from the Scandinavian countries is at the forefront of this research.

Across the OECD countries, investments in schooling differ considerably. Annual expenditures per student in primary school (grades 1-6) range from USD 2,577 (PPP-adjusted) in Turkey to USD 20,020 in Luxembourg. Such differences may in part reflect differing priorities with respect to primary education, but also differences in productivity of school inputs, e.g. class size and instruction time, across regimes. Not surprisingly, the Scandinavian countries, spending USD 10,000+ annually per student, rank in the upper part of the expenditure distribution. Despite sharp reductions in the expenditures on educational institutions, Denmark top the list of OECD countries spending the largest share of GDP (5 percent) on nontertiary education (OECD 2015). The rest of the Nordics follow suit. Nevertheless, the performance of Danish students on the Programme for International Student Assessment (PISA) studies continues to disappoint. For example, Danish and Norwegian students performed around the OECD average in math, reading and science (OECD 2014, Egelund 2013).

Economics scholars often describe the process of learning as a knowledge production function (Todd and Wolpin 2003, 2007). The analogy between the process of acquiring cognitive skills in school-aged children and the production process of a firm is advantageous in the sense that it provides a conceptual framework for evaluating inputs along with straightforward interpretations of their effects. In this framework, different school inputs are assumed to have varying productivities when producing knowledge. Some inputs are assumed to directly affect the cognitive skill production while others may be of a more implicit nature.

In a modern welfare state, we believe that attending school provides students with a positive amount of cognitive skills. This must be given; otherwise we would not require children to attend school in the first place (surely, there are cheaper ways of 'storing' children during work hours). Nonetheless, when the literature disagrees about the effect of a factor that directly increases a child's exposure to learning opportunities in school, it may, among other things, be caused by offsetting behaviour by parents and school administrators. Therefore, the production function ideally includes parental inputs (investments) and multiple dimensions of school resources. Such inputs are not often observed, however, and empirically estimated effects of inputs are therefore interpreted as total policy effects only, i.e. the *ceteris paribus* effect of changing a given input plus the response of all other inputs to that change (Todd and Wolpin 2003). Although the total policy effects are often of practical interest, they reveal little about the nature of the production function itself. Further, if parental responses differ by type, e.g. socioeconomic group, policies based on empirical evidence may cause unintended distortions in the achievement distribution of students.

The remainder of this survey is organised as follows. Section 2 sketches out a model of the knowledge production function, while Section 3 briefly comments on school-based inputs. In particular, I will distinguish between first-order (direct) school-based inputs and inputs of a more implicit nature in the production of cognitive skills, of which this paper focuses on the first. Section 4 summarises recent empirical evidence primarily in a Scandinavian context complemented by significant papers from the international literature. Section 5 is devoted to the adjustments of parental investments in students, and Section 6 concludes.

2. A simple model

Fredriksson et al. (2016) construct a stylized model for the relationship between school quality, parental inputs and student achievement.[2] Simplified, let the cognitive achievement (A) of the student be given by:

$$A = A \cdot (S, P, \mu).$$

where S denotes public investment in terms of school inputs, P denotes parental investments in the production of cognitive achievement and μ is the endowed student ability. Generally, student achievement is assumed to be (weakly) increasing in each argument (i.e. $A_j \geq 0, j = S, P, \mu$), but with decreasing returns to school and parental inputs ($A_{SS} < 0$ and $A_{PP} < 0$). Although not strictly necessary, these restrictions seem plausible within reason.[3] No restrictions are placed on the cross partials, i.e. whether public and private investments are complements or substitutes in the achievement production of students.

Parental inputs are assumed to be determined by the family's permanent resources, W, the child's initial ability and the level of school inputs: $P = P\ (W, S, \mu)$, where parental input decisions are assumed to be made subsequent to

2. The basic model is a simplication of the detailed framework presented in Todd and Wolpin (2003). Although imperfectly measuring learning, knowledge is generally proxied by test or exam achievements (Jackson et al. 2016). Fredriksson et al. (2016) augment the framework to explicitly model the response to school inputs for unconstrained and constrained parents.
3. One may imagine a scheme in which e.g. instruction time is extended unduly such that students achievement actually suffers because of exhaustion not only in the additional but also the regular lessons. Or alternatively, where the initial level of school inputs is so low that dynamic complementarities cause productivity to be increasing in additional school resources.

the realisation of school investments. To see why parental inputs depend on the level of school inputs, think of the parents as choosing a desired level of school inputs (e.g. at the time of enrolling the student into the preferred school). If S then deviates from the desired level, parents may respond by adjusting their own investment to either augment or offset the difference. Lastly, the endowed ability of the student need not be uni-dimensional and include only cognitive skills. Production of cognitive achievement is likely to also depend on noncognitive skills (Carneiro and Heckman 2003), for example self-discipline and motivation.

Because parents respond to school inputs, the empirically estimated effects generally differ from the *ceteris paribus* effect of a school input, i.e.

$$\frac{dA}{dS} = \frac{\partial A}{\partial S} + \frac{\partial A}{\partial P}\frac{\partial P}{\partial S} \neq \frac{\partial A}{\partial S}' \tag{1}$$

and the sign of $\frac{\partial A}{\partial P}\frac{\partial P}{\partial S}$ depends on whether parental investments act to offset ($P_S < 0$) or augment ($P_S > 0$) changes in school inputs, as we assume student achievement is increasing in parental investment (see, for example, Park and Holloway 2017). In this framework, there are two reasons why students from disadvantaged backgrounds would be more sensitive to changes in school inputs. Firstly, if their parents have less scope of adjusting inputs (lower W) and public and parental investments are substitutes in the production of achievement.[4] Secondly, if endowed ability and school inputs are substitutes in the achievement production ($A_{\mu S} < 0$) and students from disadvantaged backgrounds generally have a lower μ. Intuitively, this happens if extra resources primarily are spent on (or aiming at) low-performing students, e.g. increased instruction time is spent on the regular curriculum but at a slower pace.

A key challenge with applying the production function approach empirically is to account for the endogeneity of school inputs. Credible identification strategies are hard to come by and it is di-cult to separate school quality from the quality of students themselves without exogenous variation in school inputs. In his survey, Hanushek (2003) pessimistically observes that the documented relationship between school spending and student performance is weak. He requests solid evidence based on well-defined evaluation methods before additional school resources can be expected to affect student achievement.

The following sections present the school inputs considered in this paper and interpret the results of the empirical literature in this simple framework.

4. Strictly speaking, the outside option of parents is likely to increase in W. For example, Albornoz et al. (*forthcoming*) construct a framework in which an improvement in parental outside options increases their opportunity costs of time investments in the student, thus, reducing parental investments. However, such an adjustment is likely to be 'localised' compared to level of parental investments on the full range of W.

3. First-order school inputs

Schools have a wide range of instruments at their disposal (S). Some may be subject to government-imposed restrictions, such as class size caps (e.g. in Denmark and Sweden) and minimum instruction requirements. Others may largely depend on the school management, for example pedagogical foci and composition of teacher expertise. One may think of the cognitive skill accumulation in students as a process that occurs when the student is exposed to a learning opportunity, but not all school inputs may affect this exposure directly. As school inputs are vast in numbers, I have chosen to limit the scope of this survey to select inputs that are (relatively) easily manipulable and are thought to be directly linked to the frequency of learning opportunities.

We may expect that learning is predominantly generated by interactions between a student and a teacher (learning opportunities). Simplified, the school can directly increase the probability of these interactions by either increasing the length of the school day or by increasing student exposure to teachers during the school day. This is what I will refer to as first-order school inputs in the achievement production process. Increasing instruction time increases the time spent with the teacher in the classroom, thus, increasing the opportunities of teacher-student interactions. Along the other dimension, class size is a muchdebated input. Reducing the size of classrooms increases the opportunity of teacher-student interactions because there are fewer students to attend to during a lesson. Introducing additional teacher resources to the classroom produces the same outcome. Class size reductions are, however, likely to increase learning opportunities by another dimension as well. Lazear's (2001) disruption model predicts that the probability of disruptive behaviour, that impedes the learning of all others, will increase non-linearly in the number of students in the room.

Other factors, such as student and teacher absenteeism, affect the prevalence of student-teacher interactions as well. However, these will not be considered here. Although frequently studied and of great importance, absenteeism is not very precisely manipulated by schools or policy makers. Likewise, school inputs such as school size, general levels of school spending, peer characteristics, and managerial and pedagogical foci is beyond the scope of this survey. Lastly, the quality of the teacher may increase learning opportunities if better-qualified teachers are more e-cient when teaching students, thus being able to teach more students in the same amount of time. An extensive literature considers teacher quality (see e.g. Hanushek 2003), but I will only briefly discuss this quality measure in relation to increasing school resources (see Section 4).

While increasing instruction time necessarily increases the opportunity of learning, increased learning opportunities need not map into improved cognitive achievement of students. Factors such as depleted self-discipline (fatigue effects) on the part of the student may result in little or no gains of the extra lessons.

Furthermore, since schools in most countries are not subject to a welldefined maximization objective on achievement alone (unlike profit-maximizing firms), it is not obvious that all teachers will make use of such opportunities e-ciently. Lastly, parents may offset the increased time spent in school by reducing investment in their children at home, e.g. by less restrictive rules for homework.

4. Empitical evidence

The importance of understanding the production process of cognitive achievement is reflected by the sheer size of the literature. To limit the scope of this survey, however, I focus on pivotal contributions to the literature and recent evidence in Scandinavian or comparable settings.

Given the size of their public sectors in general, it is not surprising that the Scandinavian countries rank high in terms of per student school expenditures. Table 1 presents the annual per student school spending and the PISA scores for a select range of countries. Besides the Nordic countries, Table 1 includes two countries: (South) Korea and the US. Ranking five on the PISA math test, Korea is the highest scoring country for which data on school spending is available.[5] US is included as a reference country, as much pivotal research within the economics of education is based on US data. School expenditures as well as PISA scores (age 15) are measured in 2012. Among the Nordic countries, Norway spends the most resources per student in both primary and lower secondary school. The Finnish primary school is relatively cheap in terms of per student expenditures, but this is offset by a higher level of lower secondary school spending. Finland scores significantly above the OECD average in all subjects, whereas the Swedish performance is well below average.[6] Denmark scores around the average in reading and science, but significantly above average in math. Although per student expenditures in Denmark are relatively high in an international context, the Danish performance on the 2012 PISA was on par with most other Nordic countries.

5. Shanghai-China, Singapore, Hong Kong-China and Chinese Taipei top the list.
6. The Swedish performance on the PISA has been deteriorating for years (see Egelund 2013). An OECD (2015) assessment suggests that lack of classroom discipline and low student commitment are among the key drivers.

Table 1. Primary and secondary school spending per student in 2012 by 2012 PISA scores, selected countries.

	Expenditures		PISA score		
Contry	Primary school	Lower secondary school	Math	Reading	Science
Korea	7,395	7,008	554	536	538
Japan	8,595	9,976	536	538	547
Finland	8,316	12,909	519	524	545
Denmark	10,953	11,460	500	496	498
Iceland	10,003	10,706	493	483	478
Norway	12,728	13,373	489	504	495
United States	11,030	11,856	481	498	497
Sweden	10,312	10,966	478	483	485
OECD average	8,247	9,966	494	483	501

Notes. Expenditures reported in PPP-adjusted USD. Source: OECD (2014, 2015).

In terms of general spending, Jackson et al. (2016) successfully link improved long-term outcomes for students to the implementation of US school finance reforms in the 1970s-1990s, particularly, for children from low-income families. Without explicitly examining the mechanisms, the authors associate the general spending increases with lower student-to-teacher ratios and longer school years. The remainder of this section considers recent empirical evidence on the importance of these school inputs (also characterised as first-order in Section 3) for the production of cognitive achievement: Instruction time, class size and additional teacher resources.

4.1. Instruction time

Generally speaking, instruction time may be changed along two margins: the length of the school day or the number of days in a school year. Because of fatigue effects and offsetting spare time activities (academic summer camps, after-school tutoring etc.), the effect of a change by the same number of hours need not be the same along both dimensions. To simplify interpretations of input effects, the education literature operates with difierent levels of learning time (Patall et al. 2010). *Allocated school time* refers to the number of hours that the students are required to attend school including breaks etc. Then, *allocated class time* is the time that students are required to be in class, which is split into *instruction time* (time devoted to instruction) and *noninstruction time* (classroom time without instruction, e.g. for activities such as classroom management). Instruction time may further be narrowed to *engaged time*, which excludes time where the student is inattentive. Finally, the time where the student actually learns is referred to as *academic learning* *time*. Although engaged time and academic learning time will be more closely related to academic achievement, it is generally unobserved and likely varies across individuals and so instruction time is the preferred level of analysis. Importantly, the interpretation of estimated effect sizes hinges on which time measure is observed.[7]

The student-teacher interactions thought to predominantly generate learning in Section 3 are somewhat overlapping with the concept of academic learning time: increasing instruction time increases the opportunity of student-teacher interactions as well as the opportunity of academic learning time but does not guarantee it. Fatigue and boredom (e.g. arising from depleted self-discipline) may cause decreased effort during the extra lessons (Duckworth and Seligman 2005, Baumeister et al. 2007), which may at least in part offset any beneficial effect of increased instruction time.

Much like the monetary spending in primary and lower secondary school, substantial variation exists in the annual supplied hours of instruction time. For policy makers, instruction time is considered one of the most tangible tools. It is easily observed and can be adjusted without large fixed costs as opposed to for example class size changes that may require additional building capacities etc. As such, instruction time has been a key factor in many recent school reforms, for example, the 2001 No Child Left Behind Act in the US and the 2014 Danish school reform. Figure 1 illustrates the most recent distribution of compulsory instruction time for children aged 10 across the OECD countries. With 1000 hours per year, Denmark is placed in the top-markedly above Norway and Sweden with around 750 hours per year each. Following the school reform implemented in August 2014, compulsory instruction time in the Danish primary and lower secondary schools has increased drastically. Prior to the reform Denmark ranked below the OECD average with 803 hours per year. The sharp increase reects a new instructional regime where subjects promoting practical and vocational skills (compared to the pre-reform regime) and compulsory study cafes for homework among others are prioritised.

In their review of the research prior to 2009, Patall et al. (2010) echo the conclusions of Hanushek (2003) and rather vaguely note that extending school time may be beneficial to student learning, but that the existing evidence is based on weak research designs that hinder strong causal inference.

7. Allocated class time outside of instruction time or even allocated school time outside of class may very well affect student achievement in their own respect, however, the channels through which these effects operate are less clear.

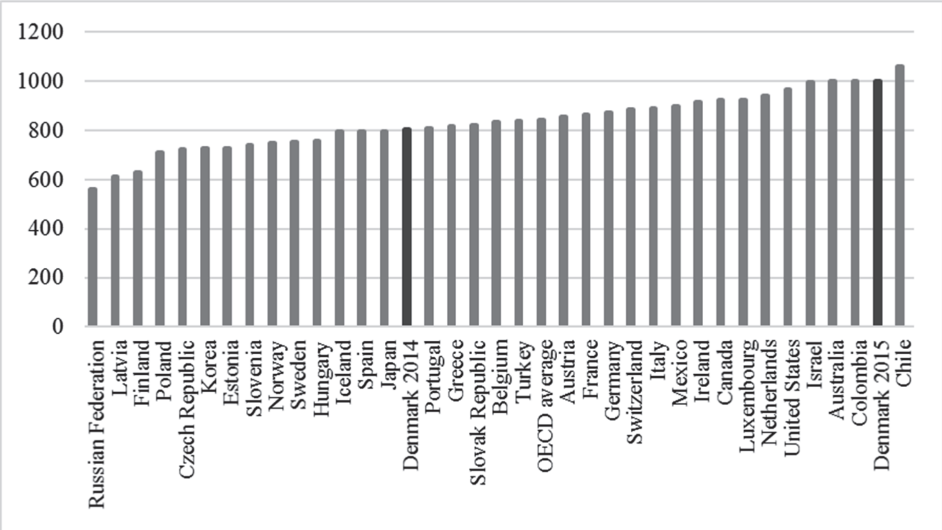


Figure 1. Yearly compulsory instruction time at age 10, 2015. Source: OECD (2014, 2015)

Exploiting within-student or within-school variation in subject-specific instruction time has been a popular strategy to address endogeneity issues in the instruction time literature.[8] Most notably, Lavy (2015) and Rivkin and Schiman (2015) document positive effects on student achievement on the 2006 and 2009 PISA studies, respectively. Lavy (2015) restricts his main sample to 22 OECD countries and estimates effect sizes of 0:06 standard deviations (SD) in the distribution of 2006 PISA scores for a one-hour weekly increase, while Rivkin and Schiman (2015) include all 72 countries for a point estimate of 0.02 SD. This is in line with the much smaller point estimate for developing countries presented in Lavy (2015). Lavy (2015) find little evidence to suggest heterogeneity across genders, but disadvantaged students in terms of low parental education and second generation immigrants benefit more from increasing instruction time. Both studies suggest that increased instruction time is most productive when schools have greater autonomy and better classroom environments.

Few studies evaluate the effect of policy-induced changes in instruction time. Lavy (2012) exploits variation in instruction time for Isareli primary schools following a reform in 2004 that redefined the basis of allocation for distributed funds to schools. The main results are remarkably similar to the within-student estimates. Interestingly, there is no evidence of either spillover or crowdingout effects across subjects: instruction time changes affect only same-subject achievement. Positive

8. For causal interpretation, this identification strategy relies on the assumption of no spillover effects between subject. I will return to these below.

spillover effects are most likely to occur when skills in one subject are advantageous for acquiring skills in other subjects. Proficiency in reading is for example often assumed to be a prerequisite for accumulating skills in other subjects. This relationship need not hold for all primary school subjects, though, and further the timing of skill accumulation is expected to be of importance. Negative spillover effects on other subjects are a real concern if acquiring cognitive skills requires self-discipline to sacrifice short-term pleasures and engage in the lessons (Duckworth and Seligman 2005). Assuming that selfdiscipline is a scarce resource that can be depleted, increasing instruction time in one subject may 'crowd out' learning in other subjects – particularly in settings such as the Israeli with high levels of instruction time (see Figure 1). Huebener et al. (2017) interpret their results in favour of positive spillover effects in their evaluation of a German reform increasing weekly instruction time for ninth graders on the academic track by two hours across five years. The increase did not affect instruction time in language arts, yet the authors document positive effects on the PISA reading scores of the same magnitude as the effects on math and science scores (0.05-0.06 SD in the international distribution). Further, the reform widened the achievement gap as low-performing students benefited less than high performers.

In a Danish context, Jensen (2013) uses the 2003 national harmonisation of minimum yearly instruction hours for broad subject groups in the ninth grade (i.e. humanities, natural sciences and practical subjects). While the average instruction time in Danish and math only changed slightly following the reform, the variation in hours was markedly reduced. Results suggest that a one-hour increase in the yearly instruction time in math corresponds to a 0:002 SD increase in math achievement (grade 9 written exit exam) whereas no significant effects are found on the written, essay-based outcome in Danish. All studies, however, rely on the (strong) identifying assumption that the changes in school resources following these reforms are not correlated with other unobserved factors of schools, students and municipalities.

Andersen et al. (2016a) evaluate a large-scale randomized experiment on increasing instruction time along two dimensions: a general increase in Danish lessons and an increase by the same weekly hours but with a detailed teaching program developed by experts.8 Increasing instruction time in Danish (unconditionally) by 3 hours per week for 16 weeks, corresponding to a 25 percent increase in annual instruction time in Danish, increased student achievement in reading by 0.15 SD. Students of both genders as well as students with and without parents with qualifying degrees benefited equally from the intervention – only students of non-Western origin did not seem to improve. The detailed teaching program in the other treatment arm benefited girls only but by the same order of magnitude as the high-discretion intervention (0:15 SD). Strict teaching guidelines have previously been found to increase student outcomes for developing countries where teachers to a large extent are considered low-skilled, see Ganimian and Murnane (2014) for a recent and comprehensive review. The evidence in Andersen et al. (2016a) therefore confirms the hypothesis that teachers in Denmark are very capable of developing and preparing material to suit the needs of their students themselves. Indeed, Ganimian and Murnane (2014) find evidence that tailoring additional instruction time to students' individual needs is what matters for academic achievement – particularly for low-achieving students.

The impact of increased instruction time likely depends on a second factor affecting the quality of student-teacher interactions as well: Namely, how motivated the students are to learning. Among students, the stock of self-discipline is typically lower for boys and disadvantaged students (Duckworth and Seligman 2005; Raver 2004), which counteracts the prediction from Section 2 that students from disadvantaged backgrounds benefit more from increasing school resources. To study the potential side-effects arising from depleted self-discipline, Andersen et al. (2016a) obtain additional information about student's self-reported behavioural difficulties. In line with the self-discipline literature, a general increase in instruction time significantly increases behavioural problems for boys. There is marginal evidence of an increased dificulty score for students with at least one parent with a qualifying education. These findings are somewhat surprising as, in terms of achievement gains, these groups benefit equally from the extra lessons compared to girls and students with low-educated parents. Lavy (2012) also includes measures of school and social satisfaction as well as violence in school in his analyses, however, these are unaffected by weekly instruction time.

Improvements in recent methodology and research designs have strengthened the evidence on the relationship between instruction time and student achievement. Recent studies mainly report positive achievement effects from increasing instruction time with relative small behavioural repercussions. Although Lavy (2012) finds suggestive evidence of the opposite (a marginal increase in productivity with weekly hours), the literature assumes that the marginal returns to instruction time are decreasing. Moreover, the marginal returns likely differ across groups of students: the specific relationship is a trade-off between the effect of depleted self-control and the benefit from an increase in studentteacher interactions. For this reason, increasing instruction time may not benefit students from disadvantaged backgrounds to the extent that additional school resources along other dimensions may. The trade-off is further expected to be influenced by the level of instruction time as well as other features such as teacher and classroom quality.

4.2. Class size

Class size reductions are a straightforward way of increasing the exposure of students to the classroom teacher without increasing allocated class time. While constituting large budget expenses, they are welcomed by nearly all school stakeholders (Hoxby 2000). Parents may prefer reduced class sizes because they allow teachers to devote more instructional resources to each child. In addition, teachers may exert less effort when teaching the same material to a smaller class. Teachers' unions may favour smaller classes because they are thought to improve the work environment of teachers as well as increase the demand for teachers. School administrators may like class size reductions as they increase their leverage. Finally, politicians up for election favour smaller classes for all of the above reasons.

In this framework, class size reductions constitute an increase in schoolsupplied inputs. Greater public investments in terms of smaller classes translate into an increased probability of the individual student interacting with the teacher – the teachers simply have more time per student in a given lesson. In addition, fewer students means less risk of disruptions (Lazear 2001).

Danish primary and lower secondary schools (*grundskoler*) are subject to a government-imposed class size cap of 28 students, however, private schools traditionally operate under more restrictive (self-imposed) rules. Although frequently debated, the cap size has not been reduced since 1975. As illustrated in Figure 2, the average class size in Danish public schools has been increasing in recent years; in 2016, public school classrooms held on average 1:5 students more compared to 2009. Private schools have also been facing an upward trend in class sizes but from a lower initial level. This likely reflect reduced government subsidies to private schools (rolled back in 2017) in addition to a somewhat unanticipated increase in demand for private schools following the much-criticised school reform in 2014.

Among economics scholars, class size reductions have been a delicate topic as well. Most notably, two papers published in the same year highlight the incongruity in the literature. In his survey, Hanushek (2003) finds that the literature on class size effects yields inconsistent results that are quite sensitive to specification and estimation approaches. Krueger (2003) surveys the same literature and argues that the pessimistic conclusion in Hanushek (2003) arises as he is oversampling results from studies using small samples and misspecified models. Empirically, class size effects are dificult to study because the majority of the variation in class size is driven by choices (Hoxby 2000) and theory predicts that there is negative selection into smaller classes (Lazear 2001).

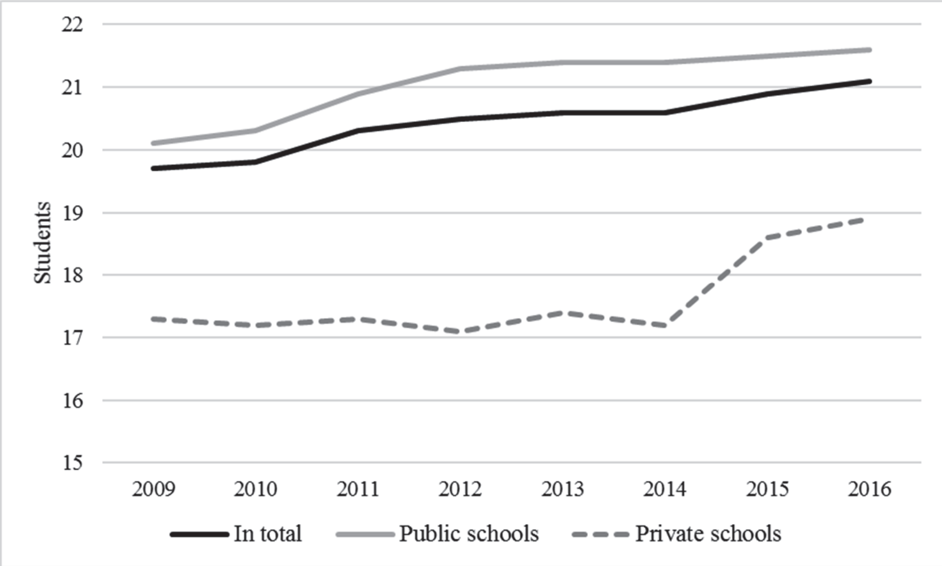


Figure 2. Average class size by school type, 2009-2016. Source: Statistics Denmark

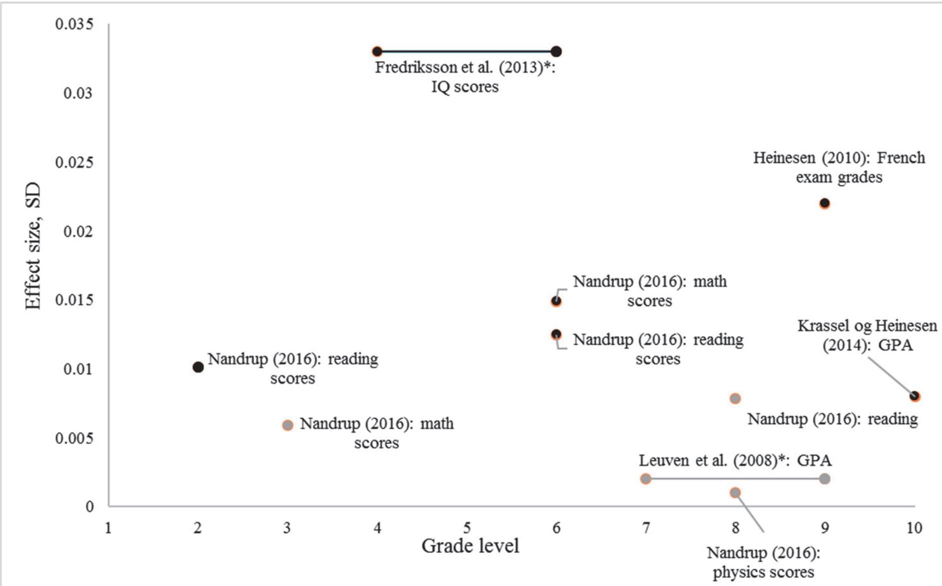
It is unjust to consider empirical estimates of class size effects without referring to the Tennessee Student Teacher Achievement Ratio (STAR) experiment from the late 1980's. As the first large-scale resource-based randomized trial, it included a treatment in which class sizes were on average decreased by seven students in kindergarten through grade 3 resulting in increased student achievement by 0.2 SD (e.g. Finn and Achilles 1999). The experiment has since been criticised for noncompliance and attrition (e.g. Hanushek 2003). Krueger (1999) estimates that the effects are driven by the first year in a small class, albeit with a larger effect for minority students and student from low-income families.

In their seminal paper, Angrist and Lavy (1999) use a government-imposed class size cap to generate exogenous variation in the endogenous class size variable. This identification strategy has since been widely replicated, and most studies in a Scandinavian context are based on this. Fredriksson et al. (2013) consider the effect of increasing average class size in grades 4-6 by one student on a nationally representative Swedish sample from 1967-1982. Here, cognitive and academic achievement scores in grades 6 and 9 significantly decrease by 0:02-0:03 SD. The authors find few examples of heterogeneous effects by student characteristics.[9] Applying the same strategy to more recent Norwegian data, Leuven et al. (2008) find no significant effects on increasing average class size in grades 7-9. Using within school

9. Fredriksson et al. (2016) reexamine the data. By means of a slightly difierent specification, the authors find evidence suggesting that results are driven by student with below-median parental income.

variation in class size for French electives, Heinesen (2010) estimates that increasing class size in grade 9 French classes significantly decreases French examination marks by 0:03 SD. This effect is significantly smaller for girls compared to boys.

The international class size literature suggests that effect sizes are decreasing across grades (e.g. Krueger 1999, Rivkin et al. 2005), which may to some extent explain the mixed evidence in the literature. Ex ante, one may expect the effect to differ for several reasons (Nandrup 2016). Peer-tutoring or group work may be more effective in older grades, because younger student may depend more on adult supervision and help (Blatchford and Mortimore 1994). Reminiscent of Lazear's (2001) disruption model, older students may be increasingly selfdisciplined and as such less likely to participate in disruptive behaviour. Mischel and Mischel (1983) suggest that older children are more effective in creating a favourable environment for self-control. Nandrup (2016) identifies negative short-run effects of around 0:01 SD of a one-student increment in class size in grades 2 and 6 of public Danish schools. No effects are found in grade 8, which is consistent with diminishing class size effects in lower secondary school, although it should be noted that the instrument has less bite on the older grade levels. Krassel and Heinesen (2014), still on Danish data, find significant effects although modest in size in grade 10. Both studies find no evidence that larger classes more adversely affect disadvantaged students. Figure 3 summarizes the empirical estimates on Scandinavian class size effects on student achievement.



*) Estimates are of average class sizes in the displayed years, academic achievement is measured in the last year.

Figure 3. Estimated short-term class size effects of a one-student decrease on academic achievement Black bullets indicate statistical significance on the 5% level.

Although the effects of school-based inputs on cognitive skills are of the main interest here, significant negative effects of larger classes have been found on noncognitive skills in terms of e.g. motivation, effort and absenteeism as well (Fredriksson et al. 2013, Dee and West 2011). In term of wellbeing and mental health, however, Jakobsson et al. (2013) fail to recover significant effects for Swedish ninth graders.

Based on weak empirical findings in the general literature, Carneiro and Heckman (2003) note that effects on test scores are likely to be only temporary. Fredriksson et al. (2013) contrast this by arguing that imputational methods previously used to back out effects of class size on later life outcomes are highly unreliable and present evidence that the short-run effects on test scores may in fact be highly persistent. The authors find that decreasing average class size in grades 4 through 6 by seven students (by linear extrapolation) decreases adult wages by 4.4 percent, however, whether this is driven mainly by positive effects on cognitive or noncognitive skills (or a third factor) remains unsaid. In Denmark, the results on the medium-run are less convincing. Heinesen and Browning (2007) find modest and marginally significant negative effects of larger classes in grade 8 on years of education measured at age 25-32 while no consistent evidence is found for having completed a secondary education. In line with Nandrup (2016), this may in part reflect that class size has a diminishing effect across grades – in particular, essentially a zero effect in grade 8. Heinesen and Browning (2007) employ a somewhat arbitrarily chosen class size cap of 24 students, arguing that municipalities may operate under a lower class size cap than the government-imposed rule of 28 students, which may also affect the results.

If information on teacher resources in the classroom is unavailable, as is the case with several of the above studies, and cognitive achievement are indeed produced by student-teacher interactions, then the estimated class size effects may be downward biased assuming that additional teacher resources are mainly assigned to larger classrooms. Heinesen and Browning (2007) obtain information on the total number of teaching hours on the grade level but find that their results are unaffected by whether class size or students per teacher hour are used as the main regressor.

In a Scandinavian context, the empirical literature does suggest beneficial short-term effects on academic performance of reducing class size. The estimated effects are, however, modest compared to those of the STAR experiment, and they seem to diminish across grade levels. Although statistically significant in primary school, these modest effect sizes suggest that public investments along the class size dimension is a relatively expensive way of improving student achievement. This potentially reflects that teachers are unable to adjust instruction to take full advantage of the smaller classrooms. If this is true, efficiency gains may be achieved by providing teachers with flexible instructional regimes aimed at teaching classrooms of varying sizes. Section 5 considers a competing hypothesis that parental investment may act as substitutes to school inputs and effectively offset the effects of class size. Contrary to general belief, the empirical evidence on heterogeneous effects across student characteristics is relatively weak in a Scandinavian context suggesting that disadvantaged students are not particularly sensitive to school investments along the class size dimension. Potential explanations of this finding include, but are not limited to, allocation of extra resources and teachers being particularly observant of such students when class sizes are big because of the risk that they may be left behind.

4.3. Additional teacher resources

Introducing a second teacher or a teacher aide to the classroom is a more flexible way of lowering the student-teacher ratio compared to reducing class sizes. Much like adjusting instruction time, teacher resources can be added without large fixed costs.

The use of teacher aides have become more widespread in many Western countries, but the literature evaluating such initiatives is primarily concerned with disadvantaged students and classrooms (Cook et al. 2011). A few studies are concerned with identifying the causal effect of extra teacher resources on student achievement in the general population of students.

In addition to reducing class sizes, the Tennessee STAR experiment included a second treatment arm in which classrooms were allocated a full time teacher aide. Evaluations of the experiment fails to identify significant effects on student achievement of the additional teaching resource (e.g. Finn and Achilles 1999, Krueger 1999), and Krueger (1999) speculates that the treatment effects may be attenuated because of widespread use of part time teacher aides in the control classrooms.

Andersen et al. (2016b) evaluate a large-scale randomized trial that allocated extra teacher resources to Danish 6th grade classrooms. Two kinds of teacher resources were allocated in the trial; a less expensive teaching assistant without a teaching degree and a more expensive co-teacher with a teaching degree. Both interventions had significantly positive effects on student achievement in reading. No significant effects were found for math. The effect of having additional teaching resources in the form of a teaching assistant was slightly larger (although not significantly so) compared to the effect of a co-teacher. However, the less expensive assistant could be afforded in the classroom for more hours a week compared to the more expensive co-teacher. Dividing the intervention effects by the number of lessons spent with the class, the two treatments each had an effect on reading achievement of 0.007 SD per lesson spent. The authors find evidence that effect sizes are generally larger for students from disadvantaged backgrounds and that the teaching assistants mainly benefited students with low-educated parents while the co-teacher mainly benefited students with higheducated parents. Further, they hypothesize that the lack of effect on student achievement in math is driven by a greater emphasis on (and, thus, allocation to) reading proficiency.

On the other hand, Leuven et al. (2007) exploit variation in teacher resources caused by an allocation rule assigning additional resources to schools with a large share of minorities in the Netherlands. The authors conclude that the additional funds obtained from the allocation rule were indeed spent on acquiring extra teachers (as well as on extra payments for current teachers), but even so the estimated effects on student achievement are negative and on some occasions significantly so. This may not be too surprising when considering that the student-teacher ratios in the affected schools were already below 14 because of a rather generous main funding scheme, thus, schools likely had dificulty allocating the additional funds in an effective manner. Moreover, a temporary increase in the wages of the current personnel may not be expected to improve teacher performance.

The solid evidence presented in Andersen et al. (2016b) suggests that additional teaching resources may indeed be beneficial for students on average.

However, evidence from Leuven et al. (2007) underpins the importance of incentive schemes and a priori well-defined purposes rather than aimlessly increasing resources. Moreover, there is reason to believe that teaching assistants without professional degrees provide a valuable supplement to the regular teachers. The fact that co-teachers obtaining a teaching degree (therefore, assumed of higher quality) are not more effective than assistants without formal degrees combined with the heterogeneity of the treatment eects presented in Andersen et al. (2016b), indicate that the productivity of the same set of teaching skills is diminishing – particularly for the disadvantaged group of students. To some extent, two qualified teachers may provide 'more of the same', whereas an assistant may offer a different classroom approach. In the latter case, there is also not the issue of 'who's the boss?' when planning and teaching lessons.

5. The nature of parental involvement

Section 2 hypothesises that parental investments in a child potentially respond to the amount of school investments in the child. In practice, the sign of the adjustment ($\frac{\partial A}{\partial P}\frac{\partial P}{\partial S}$ from (1)) may differ depending on the type of school input or even the type of parental investment considered. For example, the findings in Datar and Mason (2007) suggest that parent-child interactions are substitutes to investments in class size when producing achievement, while privately financed activities are complements.

Profound knowledge of the nature of parental investments is crucial when making policy decisions based on empirical evidence. If they act as perfect substitutes, we should expect parents to exactly offset every additional DKK invested in the schools. The presence of positive total policy effects of school inputs in Section 4 does not suggest that parents are willing or able to offset public investments completely. However, the substitution effect is likely to differ by parental ability, which may unintentionally distort the beneficial effects of increasing school resources.

In our simple framework of the achievement production function, withdrawing public funds is equivalent to fewer opportunities of student-teacher interactions. Equivalently, parental investment is then composed by parent-child interactions.[10] Withdrawing public funds could thereby cause high-ability parents to offset the decrease in school resources while low-ability parents may be unable or unwilling to pay the costs of the effort (effort may come at a high cost to low-ability parents if their interactions are of very low quality, thus, requiring relatively more). Consequently, students from disadvantaged homes will benefit more from public investments, as there is little crowding out of private investments – at the corner they simply receive no parental inputs that are beneficial in the production of achievement. On the other hand, high-ability parents may face higher opportunity costs of interacting with the child (see also the discussion in footnote 3).

The credible literature on parental response to school inputs is sparse. Bonesrønning (2004) finds evidence in favour of a compensatory allocation from class size investments on a composite measure of parental involvement for Norwegian lower secondary schools. The credibility of the identification strategy, however, is weakened as the author fails to control for parents sorting into schools of different sizes (Fredriksson et al. 2016). Evidence from the Chicago randomized lottery determining high-school admission is ambiguous in the sense that being admitted to a high-quality school both decreases the probability of parents helping regularly with homework while it increases the likelihood that parents discuss school-related issues at home (Cullen et al. 2006). The effects are similar in magnitude (10 percentage points).

Exploiting the same data and roughly the same identification strategy as in Fredriksson et al. (2013), Fredriksson et al. (2016) find that subject to a class size increase teachers increasingly believe that students should take responsibility for their own learning. This is interpreted in favour of fewer student-teacher interactions. In line with the above, parents then respond by helping the children with their homework to a greater extent – this effects is, however, entirely driven by parents with above-median income. At the same time, the probability of moving to another school district (and effectively transferring to another school) increases by 1.6 percentage points, which is primarily driven by belowmedian income parents who may be unable to offset the increase in class size by helping with homework (parent-child interactions). Also, the moving costs are very likely lower for these families.[11] The interpretation of these estimates is somewhat complicated by the endogenous variable measuring average district class size in grades 4-6 while the outcome allows students to transfer schools in any of the grades 4, 5 and 6.

10. Or private tutoring paid for by parents. Privately financed tutoring is, however, not very common in the context of Danish school-aged children.
11. Johansen et al. (2015) are unable to replicate these results on recent Danish data, potentially because of differences in institutional setting and instructional regimes across years and countries. For example, the estimated class size effects on student achievement are also larger in the Swedish setting (see Figure 3).

Student achievement among school-aged children may in part determine later-life outcomes (Todd and Wolpin 2003). Still, it is not likely to be the only child outcome to enter the utility function of the parents. Parents may care equally (or even more) about the well-being of their children. Though the recent literature on class size effects in a Danish context (Section 4) suggest negative effects on student achievement on average from a class size increase – at least on the younger grade levels – these already incorporate possible beneficial peer and network effects. Krassel et al. (2016) discuss that social benefits beyond achievement effects such as the odds of finding 'good social peers' (friends) *ceteris paribus* increase with class size, thus, parental preferences for class size may be ambiguous for different ranges. Indeed, the authors find that for moderate class sizes (17-20 students) parents' willingness to pay (WTP) for a class size reduction is not significantly different from zero. The WTP is negative for very small classes (<17 students) and positive above 21 students.

Lavy (2012) present evidence that a one-hour weekly increase in instruction time in math and English in Israeli schools is associated with a significant increase in homework time of 5 minutes a week. Assuming that student effort is increasing in parental investment, this may be construed as evidence that parental investment and instruction time are complements. However, the interpretation is ambiguous in the sense that increased instruction time in all probability implies more homework. Pop-Eleches and Urquiola (2013) exploit tracking according to endogenous admission thresholds into better high schools (grade 9) in Romania. School quality is proxied by peer test scores and highquality teachers are shown to sort into these schools. The authors find evidence that school quality and parental investments are substitutes; children who score just above the cutoff receive less homework-related help from their parents. It is, however, possible that the parents react to the signal that the student was able to 'pass' the test and gain access to a higher quality school, i.e. they respond to the perceived proficiency of the student rather than the quality of the school. Walsh (2010) finds a negative relationship between grade enrolment and parental involvement at the school. While the results are consistent with parental and public investments being substitutes to the extent that demand reflects school quality, he interprets the findings as parents reacting to larger schools by free-riding on the public goods (investments) provided by other parents.

Evidence from Sweden and Romania suggests that public and private investments in students are substitutes, i.e. parents will attempt to offset decreasing school inputs to the extent that they are able (Fredriksson et al. 2016, Pop-Eleches and Urquiola 2013). High- and low-ability parents may respond along different dimensions, though. Moreover, the reaction may not be monotonic as recent evidence on Danish data indicates that parents are aware of and respond to the trade-off between social and achievement benefits (Krassel et al. 2016). Lastly, researchers should note that in practice parental adjustments of inputs may not be symmetric over the loss and gain domains of school input changes, however, the empirical evidence on this topic is non-existent.

6. Conclusion

The former consensus among some scholars that increasing school-based inputs will not result in improved achievement is largely contradicted in the newer literature. Advances in methodology and research designs have provided the tools needed for credible analysis and inference. This paper summarises the findings of the recent literature with a particular focus on the first-order inputs and Scandinavian settings: Increasing resources in terms of instruction time and student-to-teacher ratios (both arising from lower class sizes and additional teacher resources) does have beneficial and measurable effects on the academic achievement of students. The literature on instruction time and additional teacher resources, however, suggest that 'more of the same' without improving the quality of student-teacher interactions will not necessarily increase learning for disadvantaged students in particular. The evidence on heterogeneous class-size effects is vague by comparison.

These are total policy effects, i.e. including the responses of parents and school administrators. Importantly, parents in many settings seem to act as substitutes in the production of student achievement by adjusting their investments to offset deviations from desired public investment level. As low-ability parents may have less scope of adjusting their inputs, parental response to public investments for different types of parents may be linked to the heterogeneous effects of increasing school resources. Parents may not, however, only care about the production of cognitive skills although the literature on this topic is sparse. One caveat to bear in mind: Studies in which school resources are increased without an a priory well-defined purpose generally fail to identify beneficial effects on student achievement. Many Scandinavian schools are public institutions with unclear maximisation objectives and relatively high spending levels. 'Unconditional' injections of additional resources may therefore not be put to efficient use in general and, in terms of student achievement, in particular. Thus, incorporating evidence from credible analyses on the effects of school-based inputs are immensely important when policy makers allocate resources to improve public schools.

References

Albornoz, F., S. Berlinski and A. Cabrales. Forthcoming in *Journal of the European Economic Association* as Motivation, resources and the organizationof the school system. doi: 10.1093/jeea/jvx001

Andersen, S., M. Humlum and A. Nandrup. 2016a. Increasing instruction time in school does increase learning. *Proceedings of the National Academy of Sciences (PNAS)* **113**(27): 7481-7484

Andersen, S., L. Beuchert, H. Nielsen and M. Thomsen. 2016b. The effect of teacher's aides in the classroom: Evidence from a randomized trial. *SSRN working paper*, Aarhus University. doi: 10.2139/ssrn.2626677

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

Baumeister, R., K. Vohs and D. Tice. 2007. The strength model of self-control. *Current Directions Psychological Science* **16**: 351-355

Blatchford, P. and P. Mortimore. 1994. The issue of class size for young children in Schools: What Can We Learn from Research? *Oxford Review of Education* **20**(4): 411428.

Bonesrønning, H. 2004. The determinants of parental effort in education production: Do parents respond to changes in class size? *Economics of Education Review* **23**(1), 1-9.

Carneiro, P. and J. Heckman. 2003. Human capital policy. *IZA Discussion Paper* No. 821

Coleman, J., E. Campbell, C. Hobson, F. McPartland, A. Mood, F. Weinfeld et al. 1966. Equality of educational opportunity. *U.S. Government Printing Office*, Washington, DC

Cook, B., K. McDu-e-Landrum, L. Oshita and S. Cook. 2011. Co-teaching and students with disabilities: A critical analysis of the empirical literature. In D. Hallahan and J. Kauffman (Eds.), *The Handbook of Special Education*: 147-159. New York: Routledge.

Cullen, J., B. Jacob and S. Levitt. 2006. The effect of school choice on participants: Evidence from randomized lotteries. *Econometrica* **74**(5): 1191-1230

Datar, A. and B. Mason. 2007. Do reductions in class size "crowd out" parental investment in education? *Economics of Education Review* **27**: 712-723

Dee, T. and M. West. 2011. The non-cognitive returns to class size *Educational Evaluation and Policy Analysis* **33**(1): 23-46

Duckworth, A. and M. Seligman. 2005. Self-discipline outdoes IQ in predicting academic performance of adolescents. Psychological Science **16**: 939-944

Egelund, N. 2013. PISA 2012 Undersøgelsen – En sammenfatning. *KORA*, Copenhagen.

Finn, J. and C. Achilles. 1999. Answers and questions about class size: A statewide experiment. *American Educational Research Journal* **27**(3): 557-577

Fredriksson, P., B. Öckert and H. Oosterbeek. 2013. Long-term effects of class size. *The Quarterly Journal of Economics* **128**(1): 249-250

Fredriksson, P., B. Öckert and H. Oosterbeek. 2015. Parental responses to public investments in children: Evidence from a maximum class size rule. *Journal of Human Ressources* **51**(4): 832-868

Ganimian, A. and R. Murnane. 2014. Improving educational outcomes in developing countries: Lessons from rigorous impact evaluations. NBER Working Paper Series No. 20284

Hanushek, E. 2003. The failure of input-based schooling policies. *The Economic Journal* **113**(February): F64-F98

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

Heinesen, E. and P. Browning. 2007. Class size, teacher hours and educational attainment. *Scandinavian Journal of Economics* **109**(2): 415-438

Hoxby, C. 2000. The effects of class size on student achievement: New evidence from population variation. *The Quarterly Journal of Economics* **115**(4): 1239-1285

Huebener, M., S. Kuger and J. Marcus. 2017 Forthcoming in as Increased instruction time and the widening gap in student performance. *Labour Economics* **47**: 15-34

Jackson, C., R. Johnson and C. Persico. 2016. The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. *The Quarterly Journal of Economics* **131**(1): 157-218

Jakobsson, N., M. Persson and M. Svensson. 2013. Class-size eects on adolescents' mental health and well-being in Swedish schools. *Education Economics* **21**(3): 248-263

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

Johansen, E., L. Mølbæk and A. Nandrup. 2015. Do parents shy away from large classes? *Unpublished manuscript*, Aarhus.
Krassel, K. and E. Heinesen. 2014. Class-size eects in secondary school. *Education Economics* **22**(4): 412-426
Krassel, K., J. Ladenburg and C. Dalsgaard. 2016. Balancing the risk of 'Lazearian' interrupters and the benefits of educational and social peers: tracing parental preferences for class size reduction. *Applied Economics Letters* **23**(7): 471-481
Krueger, A. 1999. Experimental estimates of education production functions. *The Quarterly Journal of Economics* **114**(2): 497-532
Krueger, A. 2003. Economic considerations and class size. *The Economic Journal* **113**(485): F34-F63
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* No. 18369.
Lavy, V. 2015. Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries. *The Economic Journal* **125**(November): F397-F424
Lazear, E. 2001. Educational production. *The Quarterly Journal of Economics* **116**(3): 777-803
Leuven, E., M. Lindahl, H. Oosterbeek and D. Webbink. 2007. The effect of extra funding for disadvantaged pupils on achievement. *Review of Economics and Statistics* **89**(4): 721-736.
Leuven, E., H. Oosterbeek and M. Rønning. 2008. Quasi-experimental estimates of the eect of class size on achievement in Norway. *Scandinavian Journal of Economics* **110**(4): 663-693
Mischel, H. and W. Mischel. 1983. The development of children's knowledge of self-control strategies. *Child Development* **54**: 613-619
Nandrup, A. 2016. Do class size effects differ across grades? *Education Economics* **24**: 83-95
OECD. 2014. PISA 2012 Results in Focus. PISA, *OECD Publishing*, Paris
OECD. 2015. Education at a Glance 2015: OECD Indicators. *OECD Publishing, Paris*. doi: 10.1787/eag-2015-52-en
OECD. 2015. Improving schools in Sweden: An OECD perspective. *OECD Publishing*, Paris.
Park, S. and S. Holloway. 2017. The eects of school-based parental involvement on academic achievement at the child and elementary school level: A longitudinal study. *The Journal of Educational Research* **110**(1): 1-16
Patall, E., H. Cooper and A. Allen. 2010. Extending the school day or school year: A systematic review of research (1985-2009). *Review of Educational Research* **80**(3): 401-436
Pop-Eleches, C. and M. Urquiola. 2013. Going to a better school: Effects and behavioral responses. *American Economic Review* **103**(4): 1289-1324
Raver, C. 2004. Placing emotional self-regulation in sociocultural and socioeconomic contexts. Child Development **75**: 346-353
Rivkin, S., E. Hanushek and J. Kain. 2005. Teachers, schools and academic achievement. *Econometrica* **73**(2): 417-458
Rivkin, S. and J. Schiman. 2015. Instruction time, classroom quality, and academic achievement. *The Economic Journal* **125**(November): F425-F448
Todd, P. and K. Wolpin. 2003. On the specification and estimation of the production function for cognitive achievement. The Economic Journal **113**(February): F3-F33
Todd, P. and K. Wolpin. 2007. The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital* **1**(1): 91-136
Walsh, P. 2010. Is parental involvement lower at larger schools? *Economics of Education Review* **29**(6): 959-970