

IZA DP No. 8019

**Inside the Black Box of Class Size:
Mechanisms, Behavioral Responses,
and Social Background**

Peter Fredriksson
Björn Öckert
Hessel Oosterbeek

March 2014

Inside the Black Box of Class Size: Mechanisms, Behavioral Responses, and Social Background

Peter Fredriksson

*Stockholm University,
IZA, IFAU and UCLS*

Björn Öckert

IFAU and UCLS

Hessel Oosterbeek

University of Amsterdam

Discussion Paper No. 8019
March 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

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

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

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

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

ABSTRACT

Inside the Black Box of Class Size: Mechanisms, Behavioral Responses, and Social Background*

Studies on the effect of class size on student achievement typically find that disadvantaged students benefit more from reduced class size than others. To better understand this differential impact, we analyze changes in the learning environment due to class size, and behavioral responses to class size among parents, schools, teachers and students. The variation in class size is induced by a maximum class size rule applying to upper primary schools in Sweden. We find that in response to an increase in class size: i) teachers seem to assign more responsibility to students; ii) low-income students find their teachers hard to follow when taught in full-class iii) high-income parents help their children more with homework; iv) parents are more likely to change schools; and v) other school inputs and student effort adjust very little. These findings help explain why we find that the negative effect of class size on achievement in our data is concentrated among low-income students.

JEL Classification: I21, I28, J24, C31

Keywords: class size, social background, heterogenous effects, regression discontinuity

Corresponding author:

Peter Fredriksson
Department of Economics
Stockholm University
10691 Stockholm
Sweden
E-mail: peter.fredriksson@ne.su.se

* We thank Erik Plug and Miguel Urquiola as well as seminar and workshop participants at Aarhus, Amsterdam, Bergen, CESifo, IFAU, Lund, Reus, SOFI, SOLE (Boston), Stockholm, UCLS and Växjö for useful comments. We gratefully acknowledge the financial support from the Marcus and Amalia Wallenberg Foundation and Handelsbanken.

1 Introduction

There is mounting experimental and quasi-experimental evidence that larger classes are detrimental for educational outcomes.¹ This appears to be particularly the case for students with disadvantaged backgrounds (e.g. Summers and Wolfe, 1977; Angrist and Lavy, 1999; Krueger, 1999; Heinesen, 2010). While some theoretical contributions suggest possible mechanisms, such as classroom disruption (Lazear, 2001) and responses of teachers or parents (Albornoz et al., 2011), the empirical literature has been silent about the reasons for this differential treatment effect.

This paper examines empirically why students from low-income families benefit more from small classes than other students. A better understanding of the mechanisms underlying class size effects (or their absence) for specific groups helps to identify under which circumstances class size reductions can be expected to be effective.

To structure our analysis we adopt the framework proposed by Todd and Wolpin (2003) and Pop-Eleches and Urquiola (2013). This framework divides the total policy effect of an exogenous change in class size into its direct (or *ceteris paribus*) effect and its indirect effects operating through endogenous responses of parents, schools, teachers and students. Experimental or quasi-experimental estimates of class size effects are typically interpreted as total policy effects, whereas the direct effect is associated with a parameter of the education production function. Both the direct effect and the indirect effects are potentially different for students from different social backgrounds.

In our analysis we study how the classroom environment varies with class size and how the behaviors of teachers, schools, parents and students respond to variation in class size. To examine heterogeneous effects, we present separate estimates for students and parents from above and below median income families, and for teachers and classes in schools which are above or below median income. The source of variation that we exploit stems from a maximum class size rule in Swedish upper primary schools. This maximum class size rule gives rise to a (fuzzy) regression discontinuity design. We apply this identification strategy to data covering the cohorts born in 1967, 1972, 1977 and 1982. For these cohorts we have access to rich data from questionnaires among 5% and 10% samples providing information about the behavior of teachers, schools, students and parents. We supplement this information with register data. In Fredriksson et al. (2013), we have used data from these cohorts to show that larger classes in upper primary school (age 10 to 13) are detrimental for cognitive test

¹See, for instance, Angrist and Lavy (1999), Krueger (1999), Urquiola (2006) and Fredriksson et al. (2013). There are also studies finding no effect (e.g. Hoxby, 2000) but these are fewer in number.

scores at ages 13 and 16, and for wages and earnings at age 27 to 42.

Our paper is related to educational research on the behavior of students and teachers in large and small classes², to a smaller economics literature on parental responses to class size (Bonesrønning, 2004; Datar and Mason, 2008), and to studies that examine parental responses to other education interventions (Houtenville and Conway, 2008; Das et al., 2013; Gelber and Isen, 2013; Pop-Eleches and Urquiola, 2013). However, our paper is the first to study the behavioral responses to changes in class size using credible identification. Unlike the earlier evidence, we also focus on the differential treatment effects by social background typically found in the literature. Moreover, no previous study has considered the wide range of mechanisms and behavioral responses that we inquire in this paper.

In line with previous results, we find that the effect of class size on cognitive outcomes differs by social background. A reduction of class size in upper primary school in Sweden by one student during three years, increases academic achievement of students from below median income families by 6.6 percent of a standard deviation while this effect is only 2.1 percent of a standard deviation for students from high-income families. The difference between these two effects is statistically significant. A similar pattern is observed for the impact of class size on cognitive ability.

Three results of our analysis contribute to the explanation of this pattern. First, in larger classes teachers seem to assign more responsibility to students for their own learning. While this response does not vary by social background, we believe that the effect of less teacher-student interaction is more detrimental for low-income students than for high-income students. Second, students from low-income families find their teachers hard to follow when they explain in front of a large class. Third, high-income parents respond to larger classes by helping their children more with homework.

We also document that parents respond to larger classes by moving their child to another school. This response is larger among low-income parents than among high-income parents, but not significantly so.

The paper proceeds as follows. The next section summarizes the conceptual framework. Section 3 describes the relevant institutions of the Swedish schooling system. Sections 4 and 5 describe the data and the estimation strategy. Sections 6 and 7 present and discuss the results. Section 8 summarizes and concludes.

²Many of these studies ignore endogeneity problems. Studies based on project STAR are exceptions (see Finn et al. 2003 for a survey). These studies are typically based on direct classroom observation or questionnaires to teachers. Interpreting these results is not straightforward, however, since the behaviors of the students and teachers might be affected by the fact that they are aware of taking part in an experiment.

2 Conceptual framework

Todd and Wolpin (2003) formalized the notion that family inputs in the education production function can change in response to a change in class size (or any other education intervention). This endogenous response of family inputs causes a difference between the direct effect of class size and its policy effect. Pop-Eleches and Urquiola (2013) extend this by also allowing other school inputs to change in response to the intervention that is evaluated. By extension it is clear that any other input that responds endogenously to class size variation will cause a deviation between the direct effect of class size and its total policy effect. To make explicit which effects we identify in our empirical analysis and how these are related to each other, we briefly present a framework building on these two papers.

Student achievement at the end of upper primary school (A) depends on child innate ability (μ) and actual resources invested up to that point. We distinguish three types of inputs: class size (S), other resources (R) and family inputs (F).

$$A = A(S, R, F, \mu)$$

Our objective is to examine the effects of an exogenous change in class size. For that purpose we assume the following sequence of events. First, nature determines class size. Then other resources adjust. And, finally, families adjust their inputs. This sequence of events implies that we can write $R = R(S)$ and $F = F(R, S)$.

To be concrete we have in mind the following situation. Consider a steady state where school principals, teachers, pupils, and parents have adjusted behavior to their expectations about class size. Then there is an unexpected departure of a student from the school. The departure of a single student causes the school to surpass a class size threshold, and class size increases exogenously. The increase in class size may cause the school principal to increase remedial training and teachers may resort to teaching in full class as opposed to tutoring. Parents observe all of these adjustments and change their behavior depending on whether they think the net effect will be positive or negative for pupil achievement.

The total effect of a change in class size on achievement is then given by the following expression:

$$\frac{dA}{dS} = \frac{\partial A}{\partial S} + \frac{\partial A}{\partial R} \frac{\partial R}{\partial S} + \frac{\partial A}{\partial F} \left(\frac{\partial F}{\partial S} + \frac{\partial F}{\partial R} \frac{\partial R}{\partial S} \right) \quad (1)$$

In our analysis we will be able to identify the total effect of class size on achievement (dA/dS), the effect of class size on a broad range of other school inputs ($\partial R/\partial S$) and the

effect of class size on indicators of family inputs ($dF/dS = \partial F/\partial S + (\partial F/\partial R)(\partial R/\partial S)$).³ It is much harder to provide evidence on the “production function parameters”: $\partial A/\partial S$, $\partial A/\partial R$, and $\partial A/\partial F$. This would require an experiment where all other inputs are fixed, which seems unfeasible. Nevertheless, we report evidence on aspects of the learning environment which is informative of the direct effect of class size, $\partial A/\partial S$.

Since the primary focus of our analysis is to explain why class size effects vary with social background, we estimate equations separately for different social groups. For outcomes that are measured at the student level we distinguish between students from families with income below and above median family income. For outcomes measured at the class or school level we distinguish between below and above median income schools.

3 Institutional background⁴

Here we describe the institutional setting pertaining to the cohorts we are studying (the cohorts born 1967-1982). During the relevant time period, earmarked central government grants determined the amount of resources invested in Swedish compulsory schools and allocation of pupils to schools was basically determined by residence. Compulsory schooling was (and still is) 9 years. The compulsory school period was divided into three stages: lower primary school (age 7-10, grades 1-3), upper primary school (age 10-13, grades 4-6) and lower secondary school (age 13-16, grades 7 to 9).

The compulsory school system had several organizational layers. The primary unit was the school. Schools were aggregated to school districts (note that these school districts are very different from U.S. school districts).⁵ School districts typically had one lower secondary school and at least one primary school. The catchment area of a school district was determined by a maximum traveling distance to the lower secondary school. The recommendations concerning maximum traveling distances were stricter for younger pupils, and therefore there were typically more primary schools than lower secondary schools in the school district. There was at least one school district in a municipality.

The municipalities formally ran compulsory schools. But central government funding

³Since we cannot separately identify the building blocks of dF/dS , reversing the sequence of events and letting other school inputs react to parental inputs, does not change anything of substance.

⁴This section follows Fredriksson et al. (2013) closely; see also Du Rietz et al. (1987).

⁵We use the term “school district” for want of a better word. The literal translation from Swedish would be “principal’s district” (*Rektorsområde*). The prime responsibility of the school district was to allocate teachers over classes within district. Unlike U.S. school districts, they cannot raise funding on their own and there is no school board. In the Swedish context, the municipality is the closest analogy to U.S. school districts.

and regulations constrained the municipalities substantially. The municipalities could top-up on resources given by the central government, but they could not employ additional teachers. The central government introduced county school boards in 1958 to allocate central funding to the municipalities. In addition, the county school boards inspected local schools.⁶

Maximum class size rules have existed in Sweden in various forms since 1920. Maximum class sizes were lowered in 1962, when the compulsory school law stipulated that the maximum class size was 25 at the lower primary level and 30 at the upper primary and lower secondary levels.⁷

We focus on class size in upper primary school, i.e., grades 4 to 6. More precisely, the main independent variable in our analyses is the average of the class sizes students experience in grades 4, 5 and 6. The reason for this focus is that our primary source of data are surveys that were administered when the students were in 6th grade.

The maximum class size rule at the upper primary level stipulated that classes were formed in multiples of 30; 30 students in a grade level in a school yielded one class, while 31 students in a grade level in a school yielded two classes, and so on.⁸ We will use this rule for identification in a (fuzzy) regression discontinuity (RD) design. This method has been applied in several previous studies to estimate the causal effect of class size.⁹

Implementing the RD design must be done with care, however. The compulsory school law from 1962 opened up for adjustment of school catchment areas *within* school district such that empty class rooms would be filled. In that process, the county school boards were instructed to take the “needs” of the pupil population into account. Thus, it is likely that the school catchment areas are adjusted within school districts to favor disadvantaged pupils. In a companion paper we show that such sorting takes place, rendering the RD design at the school level invalid.¹⁰ Because of these problems, we implement the RD design at the school district level rather than at the school level. The virtue of the school district level is that pupils were assigned to a school district given their residential address, and that district boundaries were not adjusted in response to enrollment levels. Here we focus on districts containing one

⁶In the late 1970s, Sweden was divided into 24 counties and around 280 municipalities.

⁷The fine details of the rule were changed in 1978. Prior to 1978, the rule was formulated in terms of maximum class size. From 1978 onwards, a resource grant (the so called base resource) governed the number of teachers per grade level in a school. The discontinuity points were not changed.

⁸There have always been special rules in small schools. In such areas, the rules pertained to total enrollment in 2 or 3 grade levels.

⁹The seminal paper is Angrist and Lavy (1999).

¹⁰In Fredriksson et al. (2012) we show that there is bunching around the cut-offs when school enrollment is the forcing variable. In particular it is more likely that schools are found just below than just above the cut-offs. Moreover, expected class size according to the rule predicts parental education; more children with well-educated parents are found just below the kink when school enrollment is the forcing variable.

upper primary school, which we refer to as one-school districts.¹¹ We provide evidence that the RD design at the school district level is valid in Subsection 6.1.

4 Data

The primary data source for our study is the so-called ETF-project which is run by the Department of Education at Göteborg University; see Härnquist (2000) for a detailed description. This data source contains measures of school performance of students in the final year of upper primary school for roughly a 10 percent sample of the cohorts born in 1967, 1972, and 1982, and a 5 percent sample for the cohort born in 1977. The ETF-project also contains answers from parents, pupils and teachers to questions about issues related to the school. These questionnaires were distributed when the pupils were in 6th grade (age 13). Based on this information we constructed variables that measure: elements of the learning environment (e.g. whether it is easy to understand the teacher); other school inputs (e.g. incidence of remedial training); student effort (e.g. absenteeism); and parental responses (e.g. help with homework).¹²

The data contain several measures of pupil performance at age 13. We collapse these into academic achievement and cognitive ability. Academic achievement is based on achievement tests in Swedish and mathematics. Cognitive ability is based on traditional “IQ-type” tests with logical and verbal elements.

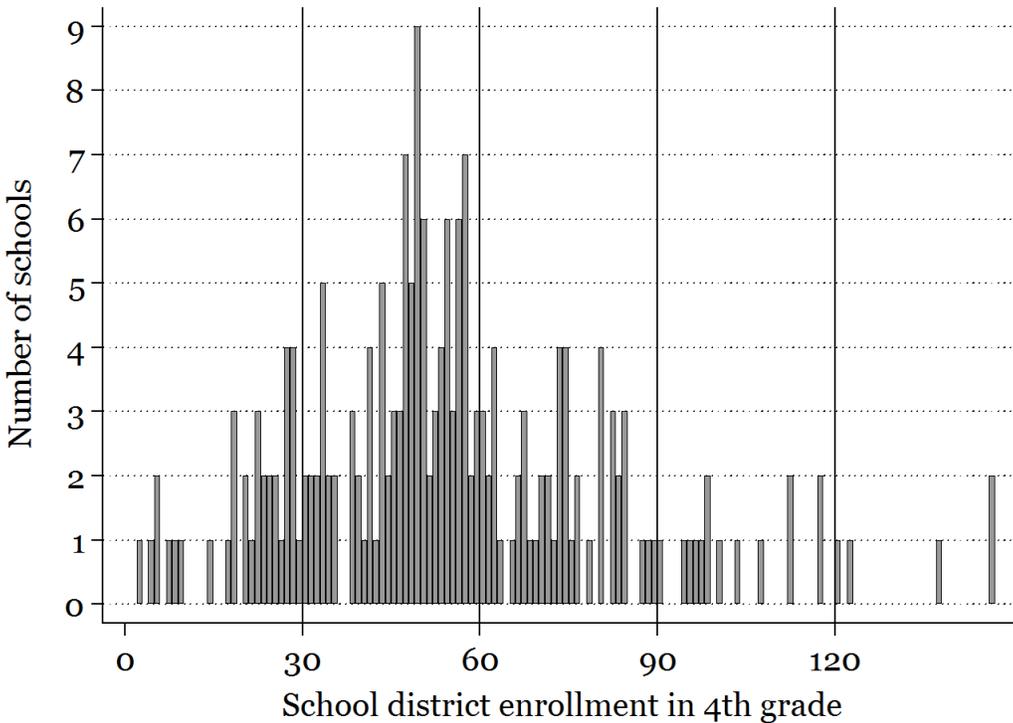
For all cohorts, a two-stage sampling procedure was used. In the first stage, 30 out of the 280 municipalities were systematically selected; the selection criteria were based on population size and political majority. In the second stage, classes were randomly sampled within municipality. This sampling procedure implies that it is unlikely that comparisons across municipalities for a given cohort are valid, but comparisons within municipalities are likely valid. For this reason all analyses condition on municipality-by-cohort fixed effects.

To the ETF-data we have matched information from the Class Register, the Educational Register, the Income Tax Register, and the Teacher Register, all maintained by Statistics Sweden. From the Class register we get information on actual class size and school enrollment by grade (the assignment variable). Using links between parents and children we get infor-

¹¹In Fredriksson et al. (2013) we compared the results for one-school districts with the results for all school districts (which includes districts with more than one school). The magnitudes of the coefficient estimates are basically the same, but the estimates for all school districts are less precise. The reason for this is that the maximum class size rule has less bite in districts with more than one school.

¹²See Appendix A5 for a detailed description of the creation of these variables. Table A6 shows which questions were included in the questionnaires for the respective cohorts.

Figure 1. Distribution of enrollment in grade 4 in one-school districts



mation on parental income and education from the Income Tax Register and the Educational Register. We also use the registers to construct other outcome variables including teacher education and teacher experience.¹³

Figure 1 shows the distribution of enrollment in grade 4 across one-school districts. In total there are 205 districts with one school in our data.¹⁴ Therefore, the distribution is “thin” at certain points. It is especially thin around the 4th threshold (enrollment of 120), which implies that this threshold does not contribute to the identification of the class size effects. We therefore restrict the sample to districts with school 4th grade enrollment of at most 120 students, and are thus exploiting the variation in class size induced by the thresholds at 30, 60 and 90. The fact that there are few school districts in our data restricts what we can do with the data. For example it is not meaningful to estimate separate effects of class size at the different thresholds. Given the amount of information we have in the data, we must pool information across the different thresholds. Section 5 outlines our approach.

¹³See again Appendix A5 for a detailed description of the creation of these variables.
¹⁴Since some of the outcome variables are not available for all cohorts, the number of districts×cohorts is sometimes smaller than 205.

Figure 2. Distribution of class size in grades 4-6

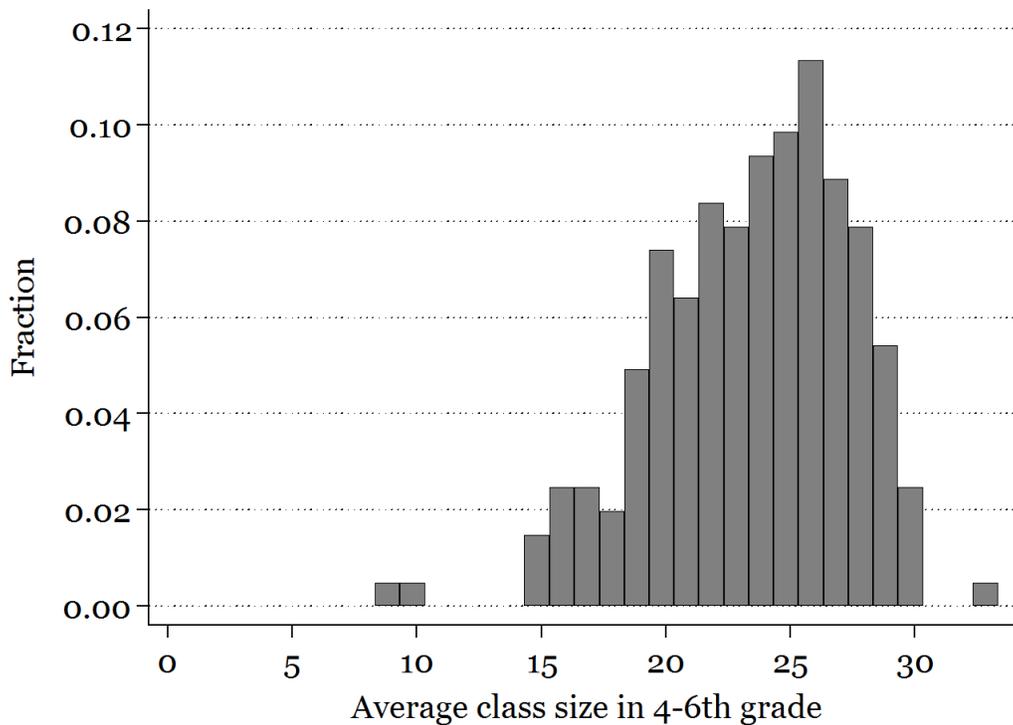


Table A1 in the Appendix reports descriptive statistics of all variables, for all individuals together and broken down by parental income. In square brackets below each variable the table also reports: (i) the number of students; the number of districts \times cohorts; and the number of enrollment counts. Following Lee and Card (2008), we cluster the standard errors in the regression discontinuity analyses by enrollment.

Figure 2 shows the distribution of actual class size in grade 4. Very few classes are below the minimum of 15 and the official maximum of 30.

5 Estimation strategy

The maximum class size rule generates thresholds at each multiple of 30 and thereby multiple discontinuities. In Fredriksson et al. (2013) we pooled the different thresholds using the analogy of aggregating several small scale experiments. Another, and more efficient, way of pooling information from different thresholds is to define a separate instrument at each threshold (see Angrist and Lavy (1999) and the web-appendix of Fredriksson et al. (2013)). With enrollment segment fixed effects (but no other covariates) this produces a minimum

variance linear combination of Wald estimates from each threshold. This is the approach we take here. In the appendix we demonstrate that the two approaches deliver very similar results; see Table A4. With separate instruments at each threshold, the estimates are more precise, however.

The RD-design also involves choices of bandwidths and of the control function for the assignment variable. To formalize our estimation approach we introduce the following notation: y_{id} is the outcome for student i in district d , e_d is district enrollment in the fourth grade, CS_d is average class size in grades 4 to 6 in the district, and Z_{td} is a dummy variable equal to 1 if 4th grade enrollment in the district is above the t -th threshold, with $t = \{1, 2, 3\}$. The equations that we estimate are (omitting further control variables):

$$y_{id} = \beta CS_d + f(e_d) + \alpha_\tau + \varepsilon_{id} \quad (2)$$

and

$$CS_d = \gamma_1 Z_{1d} + \gamma_2 Z_{2d} + \gamma_3 Z_{3d} + g(e_d) + \delta_\tau + v_d \quad (3)$$

where β is the effect of interest, γ_1 , γ_2 and γ_3 are the first stage effects, $f(e_d)$ and $g(e_d)$ are functions of enrollment, α_τ and δ_τ are fixed effects for five different enrollment segments (1-15; 16-45; 46-75; 76-105; 106-120), and ε_{id} and v_d are the error terms. Due to the inclusion of the segment fixed effects the instrumental variables (Z_{td}) effectively contrast observations at most 15 students above a threshold with observations at most 15 students below a threshold. Observations on the first (1-15) and last (106-120) segments only contribute to the estimation of the enrollment controls.

The preferred specification of the control function should capture any direct effect of school enrollment on the outcome. Because of the relatively small numbers of observations available for most of the outcome measures, however, we are restricted in the flexibility of the control functions. In our main specification we control for enrollment linearly, and let the slopes vary across the three thresholds. Table A3 shows what happens to the results with alternative ways of controlling for enrollment. Overall, the results are fairly robust to alternative ways of controlling for the assignment variable.¹⁵

¹⁵Table A3 column (3) presents the results with linear control for school district enrollment interacted with threshold. The specifications in columns (1) and (2) of Table A3 are too restrictive in our view since the enrollment control function is not interacted with threshold. In general, it seems that the 2nd order polynomial control function provides too much flexibility relative to the amount of variation that we have in the data. When we move to a 2nd order polynomial, the F-statistic for instrument relevance drops substantially (although the instruments are never weak in the conventional sense); see columns (4) and (6). The relevant contender to the specification in column (3) is the specification in column (5). In column (5), the enrollment controls are also allowed to vary by segment. Notice that results are similar with this specification. Importantly, the specifications

In all our analyses the endogenous variable is the average of the class sizes a student experiences in grades 4, 5 and 6, while the instruments are derived from enrollment in grade 4. There are two reasons for this. The first reason is that enrollment levels in 5th and 6th grade are potentially endogenous to class size in 4th grade. Therefore, we cannot validly treat enrollment in 5th and 6th grade as exogenous. Enrollment in 4th grade can arguably be treated as exogenous since 3rd (lower primary school) and 4th grade (upper primary school) belong to different stages of compulsory school. The transition between lower primary and upper primary school often implies a change of school, and class size rules are different in lower primary and upper primary school. Given that enrollment in 5th and 6th grade are potentially endogenous we have no instruments for class size in grades 5 and 6. The second reason is that class sizes in grades 4, 5, and 6 are highly correlated. The correlation between class size in grades 4 and 5 is 0.79 and the correlation between class size in grades 4 and 6 is 0.57. Attributing all effects only to class size in grade 4 would not be correct. By focusing on the average of the class sizes in grades 4, 5 and 6, the instrumental variables estimates reflect the effects of an increase of class size by one student during three years. This interpretation is clear-cut when the outcome variable is school performance measured in grade 6, which measures the accumulated skills and knowledge of the student. This is not the case for other outcomes we consider such as “help with homework” or finding the teacher “easy to understand”, which are also measured in grade 6. Here it may be more natural to relate them to class size in grade 6. However, this only amounts to rescaling the estimates; the precision of the estimates is driven by the precision of the reduced form. We prefer to rescale the estimates using the same factor to ease interpretation.

6 Balancing, first stage, and achievement

This section starts by presenting the results from balancing tests and of the first stage (Subsection 6.1). It then continues with the estimates of the impact of class size on cognitive outcomes at age 13 (Subsection 6.2). In Section 7 we move on to examining the mechanisms and the behavioral responses to class size variation.

in columns (3) and (5) have very similar impacts across the distribution of parental income.

6.1 *Balancing and first stage*

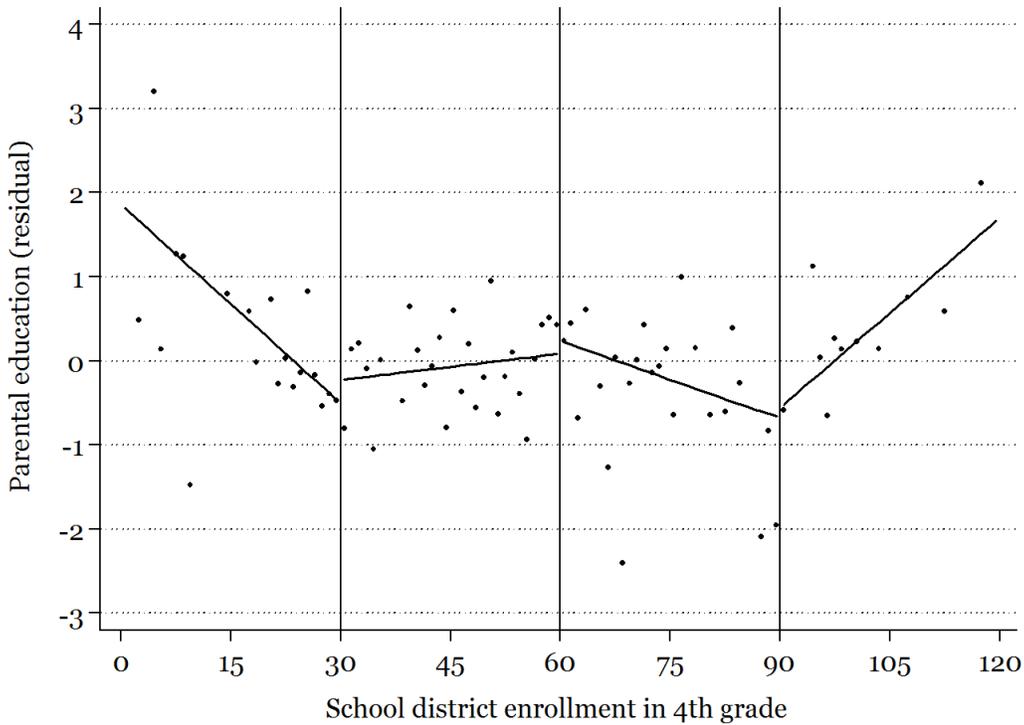
A threat to the validity of the RD design is bunching on one side of the cut-offs, since that indicates that the assignment variable is manipulated; see McCrary (2008).¹⁶ Figure 1 shows the density of school district enrollment. Visual inspection reveals no suspect discontinuities in the distribution of the assignment variable. The McCrary (2008) density test confirms this: we cannot reject the hypothesis that the density is smooth at the discontinuity.¹⁷

A more direct way to assess whether the instrument is valid is to examine if pre-determined characteristics are balanced across observations above and below the thresholds. Figure 3 shows that this is the case for parental education (where parental education is defined as the maximum years of schooling for the two parents). There is no significant jump in parental education at any of the thresholds. Analogous plots for other covariates show very similar pictures.

¹⁶Urquiola and Verhoogen (2009) document an extreme example of bunching in the context of a maximum class size rule in Chile.

¹⁷To implement the test we pooled thresholds and used a bin size of 1 student and a bandwidth of 5 students. The estimated log difference in the height of the density is 0.19 with a standard error of 0.57.

Figure 3. Parental education by enrollment in grade 4



The figure shows residual parental education, after controlling for fixed effects for enrollment segments and municipality-by-cohort fixed effects, by enrollment in grade 4. The data pertain to one-school districts for cohorts born 1967, 1972, 1977 and 1982. The regression lines were fitted to individual data. Discontinuities at thresholds: 0.322 (0.480), 0.130 (0.513), 0.175 (0.722).

Table 1 addresses the question of the balancing of pre-determined covariates more formally. Column (1) shows that the baseline covariates are highly relevant predictors of academic achievement at age 13. For instance, children who have more educated mothers score higher on the cognitive test (a year of education is associated with an increase in test scores of 0.058 standard deviation units). Column (2) shows that baseline covariates also predict, albeit to a lesser extent, one of the indicators of the learning environment. We strongly reject the hypothesis that the pre-determined covariates do not predict these two outcome; the p-values of the F-tests equal 0.0000 and 0.0005. The next three columns in Table 1 show the results of regressing each of the three instrumental variables on all baseline covariates. Only one out of 27 coefficients is significantly different from zero.¹⁸ The next to last row reports the

¹⁸We cluster the standard errors by enrollment counts. In all, there are 84 clusters. However, when studying separate thresholds as in Table 1, the number of clusters contributing to identification is much fewer. In particular, there are only 18 clusters around the third threshold, which may produce too small standard errors. If we

result of F-tests of the hypothesis that all the coefficients on baseline covariates are jointly zero. The message of these tests is that pre-determined characteristics are unrelated to the instruments.¹⁹

We have also performed another balancing test. We have asked whether surpassing one of the thresholds in upper primary school is related to class size in lower primary school; class size in lower primary school is predetermined relative to class size in upper primary school. The answer to this question is no. In other words we cannot reject the hypothesis that class size in lower primary school is unrelated to the instruments (p-value = 0.646).

Figure 4 shows the first stage graphically. There are clear, and statistically significant, jumps at all three thresholds. Districts that have just surpassed one of the thresholds have classes that are systematically smaller than districts just below the thresholds. When enrollment surpasses the first threshold, class size drops by 10 students, At the second, as well as the third, threshold class size falls by 5 students.²⁰

6.2 Cognitive outcomes

We start by presenting estimates of the effect of class size on student performance measured at age 13 at the end of upper primary school. This is the short-term total policy effect of class size on student performance, captured by the term dA/dS in equation (1).²¹ Figure 5 gives a graphical illustration, and shows that at each threshold academic achievement jumps up when a threshold is surpassed.

Table 2 presents the estimation results. For both outcome measures the effects of class size is significantly negative. A one student increase in class size during three years lowers academic achievement by 4.4 percent of a standard deviation and cognitive ability by 2.7 percent of a standard deviation. Splitting the sample by parental income, reveals that the negative effects are concentrated among students from low-income families.²² For students from high-income families, the estimates have the expected negative sign but are not statistically significant. The larger negative effect of class size on the academic achievement of students from low-income families is consistent with results from other studies, including

instead use robust standard errors, all estimates in columns (3) to (5) become insignificant.

¹⁹Tables for low-income students and high-income students separately, look very similar.

²⁰Table A2 shows estimates of the first stage for various specifications of the control function.

²¹These estimates are comparable but not the same as results reported in Fredriksson et al. (2013). Differences are due to inclusion of small school districts, a slightly different empirical strategy and a split of the sample into below and above median income parents.

²²Splitting the sample by above and below median parental education instead of income gives very similar results, throughout.

Table 1. Balancing of covariates

	Academic achievement (1)	Easy to understand (2)	Above 1st threshold (3)	Above 2nd threshold (4)	Above 3rd threshold (5)	p-value (6)
Female	0.0474 (0.0409)	0.0093 (0.0140)	0.0013 (0.0025)	-0.0023 (0.0034)	0.0038*** (0.0014)	0.1143
Month of birth	-0.0236*** (0.0038)	-0.0035* (0.0019)	-0.0008 (0.0005)	0.0005 (0.0005)	-0.0002 (0.0003)	0.3867
Immigrant	-0.2349*** (0.0633)	-0.0329 (0.0299)	0.0025 (0.0075)	0.0112 (0.0167)	-0.0032 (0.0053)	0.9826
Mother's years of education	0.0583*** (0.0072)	0.0058* (0.0030)	-0.0002 (0.0008)	0.0009 (0.0010)	0.0002 (0.0005)	0.9720
Father's years of education	0.0462*** (0.0062)	-0.0003 (0.0027)	0.0004 (0.0007)	-0.0015 (0.0012)	0.0003 (0.0006)	0.9623
Parental income (SEK 100,000s)	0.0383*** (0.0084)	0.0076 (0.0049)	0.0006 (0.0014)	0.0020 (0.0018)	-0.0012 (0.0010)	0.7475
Mother's age at birth	0.0107*** (0.0028)	0.0017 (0.0016)	-0.0001 (0.0003)	0.0004 (0.0006)	-0.0005 (0.0003)	0.6039
Number of siblings	-0.0478*** (0.0129)	-0.0138** (0.0052)	-0.0005 (0.0012)	-0.0003 (0.0025)	-0.0013 (0.0011)	0.7310
Parents separated, age 9	-0.1349*** (0.0325)	-0.0271* (0.0159)	0.0007 (0.0043)	-0.0019 (0.0065)	-0.0055 (0.0045)	0.7330
p-value of F-test	0.0000	0.0005	0.9393	0.3928	0.2421	
N	4,707	2,719	6,009	6,009	6,009	

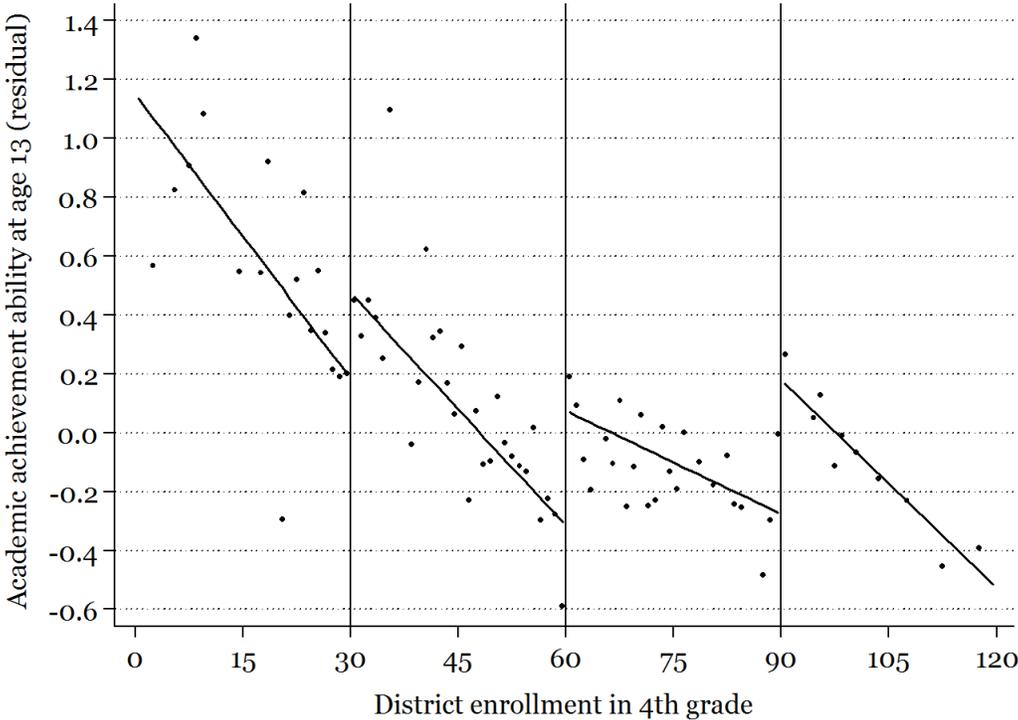
Note: The estimates are based on representative samples of individuals born in 1967, 1972, 1977 or 1982 in one-school districts. Academic achievement at age 13 is standardized. Columns (1) – (5) report results of OLS regressions on the variables listed in the rows. These regressions also include the following control variables: fixed effects for enrollment segment, linear controls for school district enrollment interacted with threshold, and municipality-by-cohort fixed effects. Above 1st/2nd/3rd threshold are indicators equaling unity if school district enrollment in 4th grade exceeds the 1st/2nd/3rd threshold of the class size rule. Independent variables are pre-determined parent and student characteristics. The p-value reported at the bottom of the columns is for an F-test of the joint significance of the variables listed in the table. Each row of column (6) reports a p-value from separate OLS regressions of the pre-determined variable (listed in the corresponding row) on the instruments, and the same set of control variables as in columns (1)–(5). The p-value is for a t-test of the significance of the class size instrument. Standard errors adjusted for clustering by enrollment count (84 clusters) are in parentheses. ***/**/*=#the estimates are significantly different from zero at the 1/5/10 percent level, respectively.

Figure 4. Class size by enrollment in grade 4



Notes: The figure shows residual class size, after controlling for municipality by cohort fixed effects, by 1-student bins. The regression lines were fitted to individual data. Discontinuities at thresholds: -9.984 (0.925), -5.137 (1.055) and -4.712 (1.414).

Figure 5. Academic achievement by enrollment in grade 4



Notes: The figure shows residual academic achievement at age 13, after controlling for fixed effects for enrollment segments, municipality-by-cohort fixed effects and baseline covariates, by enrollment in grade 4. The data pertain to one-school districts for cohorts born 1967, 1972, 1977 and 1982. The regression lines were fitted to individual data. Discontinuities at thresholds: 0.295 (0.149), 0.397 (0.121) and 0.448 (0.156).

Table 2. IV-estimates of class on school performance

Outcome [# individuals]	Income parents			Difference (2) and (3) (4)
	All (1)	low (2)	high (3)	
Academic achievement, age 13 [N=4,707]	-0.0441*** (0.0122)	-0.0658*** (0.0174)	-0.0206 (0.0148)	-0.0453* (0.0229)
Cognitive ability, age 13 [N=5,197]	-0.0268** (0.0116)	-0.0553*** (0.0151)	-0.0029 (0.0140)	-0.0524** (0.0206)

Note: The estimates are based on representative samples of individuals born in 1967, 1972, 1977 or 1982 in one-school districts. The outcomes are standardized. Average class size in grades 4-6 is instrumented by indicators for being above 1st, 2nd, or 3rd threshold of the class size rule. All models include the following controls for school district enrollment in grade 4: fixed effects for enrollment segment; linear controls for enrollment which are interacted with threshold. In addition all models include the following baseline controls: municipality-by-cohort fixed effects, gender, dummy variables for month of birth, dummy variables for mother's and father's educational attainment, parental income, mother's age at child's birth, indicators for being a first or second generation Nordic immigrant, indicators for being a first or second generation non-Nordic immigrant, an indicator for having separated parents, and the number of siblings. Standard errors adjusted for clustering by enrollment count (84 clusters) are in parentheses. ***/**/* = the estimates are significantly different from zero at the 1/5/10 percent level, respectively.

Angrist and Lavy (1999) and Krueger (1999). This is the differential treatment effect we try to understand.

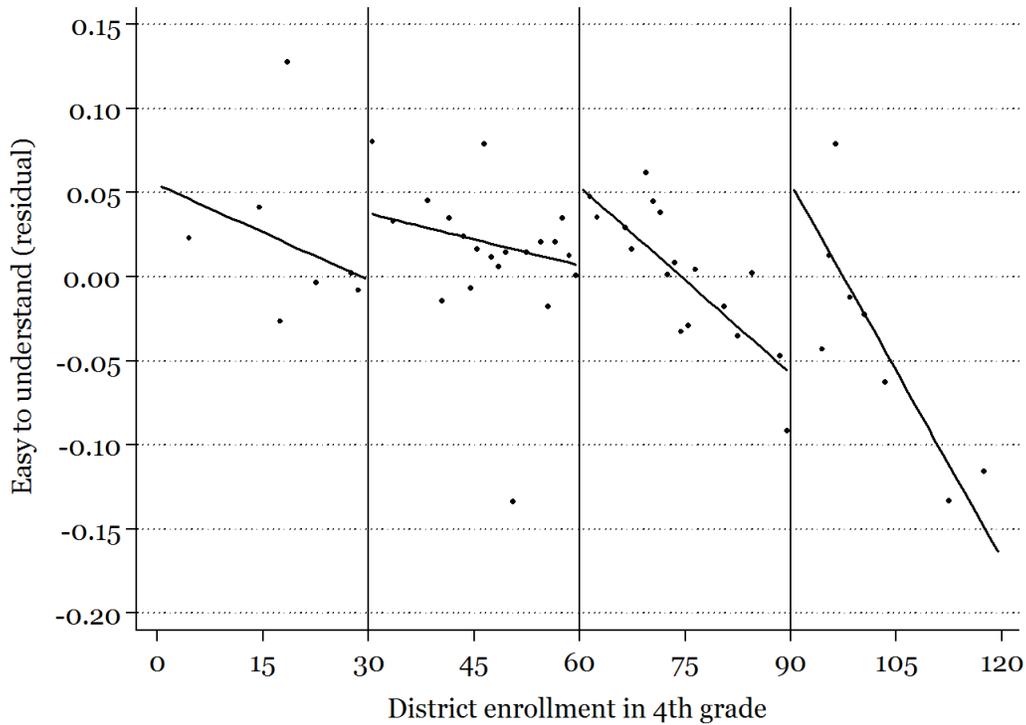
7 Mechanisms and behavioral responses

This section presents our main results. In four subsections we present the impacts of class size on the learning environment (7.1), other school inputs (7.2), student effort (7.3) and parental responses (7.4). To try to understand the difference in total policy effects between low-income and high-income pupils, we present results separately for these two groups.

7.1 Learning environment

This subsection examines how variation in class size affects the learning environment as perceived by students and teachers. This informs us about what actually changes in the classroom when the number of students changes. In terms of the conceptual framework of Section 2, we interpret these outcomes as relating to the factors that determine the sign and magnitude of the direct effect $\partial A/\partial S$.

Figure 6. Easy to understand by enrollment in grade 4



Notes: The figure shows residual “easy to understand”, after controlling for fixed effects for enrollment segments, municipality-by-cohort fixed effects and baseline covariates, by enrollment in grade 4. The data pertain to one-school districts for cohorts born 1967, 1972, 1977 and 1982. The regression lines were fitted to individual data. Discontinuity at thresholds: 0.040 (0.045), 0.045 (0.027), and 0.111 (0.077).

By way of illustration, Figure 6 shows the relation between one indicator of the learning environment (how easy students find it to understand their teacher in front of the class) and enrollment. It shows that at each threshold it is easier to understand the teacher when a threshold is surpassed.

Table 3 reports IV estimates of the effect of class size on teachers’ views on exams, homework, and pupil responsibility (unfortunately we do not have information on what teachers actually do in the classroom). The results indicate that teachers in larger classes assign more responsibility for learning to the pupils; there appears to be no shift, however, on the emphasis placed on exams and homework. We also examine whether teacher perceptions vary by socio-economic characteristics of the school. Columns (2) and (3) show that results are very similar across income groups, which is confirmed by the insignificant differences in column (4). Independently of the characteristics of the school, teachers assign more responsibility

for learning to the pupils when they teach in a larger class.

Although differential teacher responses do not account for differential policy effects, we still believe that teacher responses are important for understanding the differential policy effect. The fact that students appear to be left on their own to a greater extent in a larger class may be particularly detrimental for low-income students. In other words, it seems likely that the effect of less student-teacher interaction is larger (in absolute value) for low-income students.

Table 4 reports results regarding pupil perceptions about aspects of the learning environment. Students are not more or less distracted (think about other things) when placed in a large class; see first row. They do, however, find it harder to understand full-class teaching in larger classes. This effect is entirely due to students from low-income families; a 5-student increase in class size reduces the probability that a low-income student finds the teacher easy to understand by 7 percentage points. This should be compared to the sample mean of around 85 percent of the students finding it easy to understand when the teacher explains in front of class. The perceptions of students from high-income families are unaffected by class size. Low-income students also appear to ask teachers for help less frequently in larger classes; this effect is, however, not significant at the 10% level.²³

Our interpretation of these findings is that teachers in larger classes use less intensive teaching methods (assign more responsibility for learning to the students), probably implying more frontal teaching and less personalized instruction. As a result of an increase in class size, students from low-income families find it harder to understand when taught in full class. Students from high-income families, apparently do not experience a change in the learning environment when the class size changes. This suggests that the direct effect of class size on achievement is larger (in absolute size) for students from low-income families than for students from high-income families.

7.2 *Other school inputs*

This subsection explores whether variation in class size has an impact on other school inputs. In terms of the conceptual framework of Section 2, this addresses the sign and magnitude of the term $\partial R/\partial S$ in equation (1) and the channels through which it operates.²⁴ Results are presented in Table 5.

²³Notice, that we do not have a lot of variation to exploit for this outcome. 96% of students answer that they ask for help when they do not understand. This share does not vary by family income.

²⁴In Fredriksson et al. (2013) we looked at this issue from the perspective of the exclusion restriction and presented corresponding reduced form estimates.

Table 3. IV-estimates of class size on learning environment: teacher perceptions

Outcome [# individuals]	All		Schools with family income		Difference (2) and (3) (4)
	(1)	(2)	below median	above median	
Teacher thinks that exams are important for pupils' learning [N=1,847]	0.0023 (0.0172)	0.0019 (0.0139)	0.0023 (0.0109)	-0.0003 (0.0177)	
Teacher thinks that regular homework is important for pupils' learning [N=1,847]	-0.0081 (0.0101)	-0.0040 (0.0104)	-0.0119 (0.0110)	0.0079 (0.0151)	
Teacher thinks it is important that pupils take responsibility for their learning [N=1,848]	0.0155** (0.0072)	0.0169** (0.0074)	0.0144** (0.0065)	0.0026 (0.0098)	

Note: The estimates are based on representative samples of individuals born in 1967, 1972, 1977 or 1982 in one-school districts. Average class size in grades 4-6 is instrumented by indicators for being above 1st, 2nd, or 3rd threshold of the class size rule. All models include the following controls for school district enrollment in grade 4: fixed effects for enrollment segment; linear controls for enrollment which are interacted with threshold. In addition all models include the following baseline controls: municipality-by-cohort fixed effects, gender, dummy variables for month of birth, dummy variables for mother's and father's educational attainment, parental income, mother's age at child's birth, indicators for being a first or second generation Nordic immigrant, indicators for being a first or second generation non-Nordic immigrant, an indicator for having separated parents, and the number of siblings. Standard errors adjusted for clustering by enrollment count (84 clusters) are in parentheses. ***/**/*=the estimates are significantly different from zero at the 1/5/10 percent level, respectively.

Table 4. IV-estimates of class size on learning environment: pupil perceptions

Outcome [# individuals]	Income parents			Difference (2) and (3) (4)
	All (1)	low (2)	high (3)	
Pupil thinks about other things in school [N=2,758]	-0.0076 (0.0049)	-0.0124 (0.0076)	-0.0008 (0.0085)	-0.0116 (0.0114)
Pupil thinks it is easy to understand when teacher explains in front of class [N=2,719]	-0.0055*** (0.0019)	-0.0138*** (0.0042)	0.0025 (0.0040)	-0.0163*** (0.0058)
Pupil asks the teacher for help if (s)he do not understand [N=2,802]	-0.0015 (0.0015)	-0.0043 (0.0029)	-0.0003 (0.0030)	-0.0040 (0.0041)

Note: The estimates are based on representative samples of individuals born in 1967, 1972, 1977 or 1982 in one-school districts. Average class size in grades 4-6 is instrumented by indicators for being above 1st, 2nd, or 3rd threshold of the class size rule. All models include the following controls for school district enrollment in grade 4: fixed effects for enrollment segment; linear controls for enrollment which are interacted with threshold. In addition all models include the following baseline controls: municipality-by-cohort fixed effects, gender, dummy variables for month of birth, dummy variables for mother's and father's educational attainment, parental income, mother's age at child's birth, indicators for being a first or second generation Nordic immigrant, indicators for being a first or second generation non-Nordic immigrant, an indicator for having separated parents, and the number of siblings. Standard errors adjusted for clustering by enrollment count (84 clusters) are in parentheses. ***/**/*=the estimates are significantly different from zero at the 1/5/10 percent level, respectively.

Table 5. IV-estimates of class size on other school inputs

School input [# individuals]	Schools with parental income			Difference (2) and (3) (4)
	All (1)	below median (2)	above median (3)	
Remedial training [N=4,370]	0.0038 (0.0071)	-0.0006 (0.0080)	0.0077 (0.0080)	-0.0083 (0.0113)
Age integrated class [N=6,009]	-0.0179** (0.0076)	-0.0189** (0.0085)	-0.0169** (0.0069)	-0.0020 (0.0110)
<u>Dissimilarity index</u>				
Parental income peers [N=6,009]	-0.0019 (0.0027)	-0.0017 (0.0028)	-0.0022 (0.0025)	0.0005 (0.0038)
Parental education peers [N=6,009]	-0.0047 (0.0032)	-0.0049 (0.0035)	-0.0044 (0.0028)	-0.0005 (0.0045)
Teacher experience [N=5,901]	0.0228 (0.1057)	0.0311 (0.1087)	0.0097 (0.1009)	0.0213 (0.1483)
Teacher education [N=5,901]	0.0039 (0.0028)	0.0049 (0.003)	0.0030 (0.0026)	0.0019 (0.0040)

Note: The estimates are based on representative samples of individuals born in 1967, 1972, 1977 or 1982 in one-school districts. Average class size in grades 4-6 is instrumented by indicators for being above 1st, 2nd, or 3rd threshold of the class size rule. All models include the following controls for school district enrollment in grade 4: fixed effects for enrollment segment; linear controls for enrollment which are interacted with threshold. In addition all models include the following baseline controls: municipality-by-cohort fixed effects, gender, dummy variables for month of birth, dummy variables for mother's and father's educational attainment, parental income, mother's age at child's birth, indicators for being a first or second generation Nordic immigrant, indicators for being a first or second generation non-Nordic immigrant, an indicator for having separated parents, and the number of siblings. Standard errors adjusted for clustering by enrollment count (84 clusters) are in parentheses. ***/**/*=the estimates are significantly different from zero at the 1/5/10 percent level, respectively.

The first row examines whether the probability of being assigned to remedial training is affected by class size. Schools may attempt to compensate for larger classes by more use of remedial training. We find no evidence of this. This is also the case when we estimate the model separately for schools where average parental income is above and below the median of average parental income (columns (2) and (3)). The next row indicates that class size reduces the likelihood of being placed in an age-integrated class. The reason for this is probably that schools are induced to save resources by pooling grades when surpassing a threshold. The presence of this effect is independent of average parental income in the school. The evidence on the effect of age-integrated classes on school performance is mixed. Using an RD-type of design, Leuven and Rønning (2012) find that age-integrated classes boost performance for students in lower secondary education in Norway. In contrast, using an instrumental variable approach, Sims (2008) finds that a higher fraction of students in combination classes negatively affects performance for 2nd and 3rd graders in California.

The maximum class size rule causes variation in class size together with variation in the number of classes. Some potential responses of schools are caused by the combination of these two changes. They are therefore not applicable to variation in class size per se but to variation in class size due to the specific source of variation exploited here.

The opening of an extra classroom does not only reduce average class size, but the increase in the number of classrooms may also give more scope for tracking of students across classrooms. With two classes instead of one at the grade level, schools could for instance place all students whose prior performance is above the median in one classroom and all others in another class. In the context of primary schools in Kenya, Duflo et al. (2011) find this form of tracking to be beneficial for all students. Since we have no information on students' prior performance we cannot directly test whether schools engage in ability tracking. Instead we assess whether students in smaller classes are more homogenous in terms of their social background. We constructed two dissimilarity indices (Duncan and Duncan, 1955) which relate class composition to school composition. One index is based on parents' education, the other on parents' income. Using these measures tests for ability tracking insofar as ability and social background are correlated.²⁵ The negative signs of the IV-estimates of class size on the dissimilarity indexes of income and education of peers' parents in Table 5 are consistent with increased segregation by social background, when a school surpasses a threshold, but the estimates are far from being statistically significant. The estimates are very similar for the two groups of schools distinguished on the basis of average parental income.

²⁵Ability tracking on the basis of prior performance may not be feasible for Swedish upper primary schools because there is no test for students in lower primary schools that is comparable across schools.

Schools with more classes have to hire more teachers. This may affect the average quality of teachers in a school. The last two rows of Table 5 report IV-estimates of the effects of class size on the average experience and average education level of teachers in a school. This shows that class size is unrelated to both measures of teacher quality.²⁶ This is also true in the subsamples of schools differentiated by average parental income.²⁷

All in all, the results in this subsection indicate that the adjustments of other school inputs to class size variation are minor. The only significant change is that larger classes are less often age-integrated. But this response does not vary by average parental income in the school. Moreover, we have seen no evidence suggesting that the effect of being placed in an age-integrated class varies by parental income. We conclude that responses in terms of school inputs do not contribute to the differential impact of class size by family background.

7.3 *Pupil effort*

The conceptual framework in Section 2 has family input as an argument in the education production function. In this subsection we examine whether class size affects pupil effort; in the next subsection we consider parental responses. In terms of our conceptual model, these estimates shed light on the sign and magnitude of dF/dS .

Table 6 shows IV-estimates of the effects of class size on the three measures of pupil effort that we have constructed. The first column presents estimates for all observations together, while the other columns present estimates by parental income (and the difference therein). The overall message of the results is that we find no effects of class size on any of the three measures of student effort: On average, there are no effects on students' persistence, absenteeism and time spent on homework. The only significant result is that students from high-income families score lower on the persistence measure when they are placed in a large class. The effect is small, however, and borderline significant. It seems that the negative effect of class size on achievement is not due to a reduction in pupil effort (at least for the measures of effort we have at our disposal). Moreover, there is no evidence of a differential response along the family income distribution.

²⁶In Fredriksson et al. (2013) we also asked whether class size variation in upper primary school (grades 4-6) affected class sizes in lower secondary school (grades 7-9). The answer to this question was no.

²⁷We also estimated the effect of class size on teacher mobility (across school districts or out of the teaching profession) and find no significant effects (not reported). This concurs with the absence of an effect of class size on teacher experience.

Table 6. IV-estimates of class size on pupil effort

Outcome [# individuals]	All (1)	Income parents low (2)	high (3)	Difference (2) and (3) (4)
Pupil give up if (s)he gets a difficult task in school [N=4,701]	0.0020 (0.0023)	-0.0014 (0.0036)	0.0046* (0.0025)	-0.0060 (0.0044)
Absenteeism (hrs) [N=3,163]	0.1959 (0.4070)	-0.3010 (0.5217)	0.8274 (0.6743)	-1.1284 (0.8525)
Student does homework (hrs) [N=5,149]	0.0071 (0.0183)	0.0078 (0.0230)	-0.0007 (0.0231)	0.0085 (0.0326)

Note: The estimates are based on representative samples of individuals born in 1967, 1972, 1977 or 1982 in one-school districts. Average class size in grades 4-6 is instrumented by indicators for being above 1st, 2nd, or 3rd threshold of the class size rule. All models include the following controls for school district enrollment in grade 4: fixed effects for enrollment segment; linear controls for enrollment which are interacted with threshold. In addition all models include the following baseline controls: municipality-by-cohort fixed effects, gender, dummy variables for month of birth, dummy variables for mother's and father's educational attainment, parental income, mother's age at child's birth, indicators for being a first or second generation Nordic immigrant, indicators for being a first or second generation non-Nordic immigrant, an indicator for having separated parents, and the number of siblings. Standard errors adjusted for clustering by enrollment count (84 clusters) are in parentheses. ***/**/*=the estimates are significantly different from zero at the 1/5/10 percent level, respectively.

7.4 Parental responses

We look at two dimensions of parental response: parental help with homework and the probability of moving the child to another school. Table 7 reports the results. These show that high-income parents are more likely to help their child with homework when it is placed in a large class, while low-income parents do not respond at all along this dimension. The difference across the two groups is statistically significant. Notice that Tables 3 and 4 show that an increase in class size has no effect on the views of teachers vis-a-vis homework and the amount of homework that students do. Therefore, the response among high-income parents should be interpreted as an increase in the propensity to help with homework for a given amount of homework.

With respect to the mobility response, we find that parents are more likely to move to another school when their child is placed in a large class. The response magnitude is larger among low-income parents, but the difference across the two income groups is not significant.²⁸

A literal interpretation, based on the magnitudes of the responses across the two groups, is that parents respond according to comparative advantage. High-income parents may be better able to help their child with homework but it may be more costly for them to move. Low-income parents may be incapable to help their children with homework but may have lower moving costs, for instance because they rent instead of own their home.

The final row of Table 7 shows results from an IV-regression where the dependent variable is the difference between actual class size and expected class size (the class size for non-movers) and where mobility is instrumented by the class size rule. The estimate answers the question: By how much does mobility (caused by surpassing a threshold) contribute to reducing class size? The identifying assumption (apart from random assignment of the instrument) is that the class size rule affects the difference between actual and expected class size only through its effect on mobility, which is (almost) true by construction. The results show that the movers are indeed moving to a school with smaller class size. Again there is no statistical difference across the parental income distribution.

It is also interesting to note that class size is reduced in excess of what would be expected from a random choice of destination school. The nature of the RD design implies that the

²⁸The differential mobility response across the family income distribution is also evident in STAR. We have looked at the responses among the population randomized to small and regular sized classes in Kindergarten. As a proxy for mobility we used attrition, i.e. whether a student's test score is recorded in grade 3. The class size reduction in STAR reduces attrition among students on free lunch by 5 percentage points (t-ratio 2.43). Among pupils not eligible for free lunch there was a reduction by 0.8 percentage points (with a t-ratio 0.47).

origin school has maximum class size, i.e. 30. This implies that we would also expect a decline in class size with a random choice of destination school. With a random choice of destination school (within municipality) class size would be reduced by 4.4 students which is less than the 7.4 students documented in Table 7.

To summarize this section, we have found that parental assistance may contribute to making the policy effect of a change in class size smaller in absolute magnitude for high-income students than for low-income students. High-income parents respond to an increase in class size by helping out more with homework; low-income parents do not respond at all along this dimension. It is more difficult to take a stance on the mobility response. The mobility response is not statistically different across the two groups. But it is twice as large in the low-income group. It is not clear that moving to another school is to the benefit of the student. The move comes with a reduction in class size, but it also comes with changes of school, teachers, peers and neighborhood. The net effect of these changes may be harmful for the school performance of the child. Indeed, using data from Texas, Hanushek et al. (2004) find that the disruption associated with moving is negative for student achievement.

8 Conclusion

A recently emerging literature uses credible identification to assess how other inputs in the human capital production function respond to a specific education intervention (e.g. Pop-Eleches and Urquiola, 2013; Das et al., 2013; Gelber and Isen, 2013). This paper contributes to this literature by: (i) assessing such responses in the context of class size reduction; and (ii) examining whether differential responses by social background can account for the heterogeneous impact of class size on achievement by family background.

We first show that class size has a negative effect on school performance, especially for students from low-income families. This reduction in performance is related to changes in the *learning environment*. As a result of an increase in class size, teachers respond that pupils should take greater responsibility for their own learning. A larger class necessarily means less pupil-teacher interaction. Teachers may also shift towards full class instruction, as a response to a class size increase. Interestingly, we show that an increase in class size implies that low-income students find their teachers harder to follow when taught in full class. Students from high-income families appear not to experience a change in learning environment resulting from a change in class size.

Parental responses to class size vary with income. High-income parents respond to larger classes by helping their children more with homework. We also document that parents re-

Table 7. IV-estimates of class size on parental responses

Outcome [# individuals]	Income parents			Difference (2) and (3) (4)
	All (1)	low (2)	high (3)	
Parents help child with homework [N=5,107]	0.0086** (0.0043)	-0.0001 (0.0053)	0.0156*** (0.0046)	-0.0157** (0.0070)
Move to other school district [N=4,370]	0.0112*** (0.0033)	0.0151*** (0.0039)	0.0071 (0.0051)	0.0080 (0.0065)
Change in class size for movers [N=4,370]	-7.3569** (2.9874)	-7.6672** (3.0597)	-4.6179* (2.5686)	-3.0493 (4.1840)

Note: The estimates are based on representative samples of individuals born in 1967, 1972, 1977 or 1982 in one-school districts. Average class size in grades 4-6 is instrumented by indicators for being above 1st, 2nd, or 3rd threshold of the class size rule. All models include the following controls for school district enrollment in grade 4: fixed effects for enrollment segment; linear controls for enrollment which are interacted with threshold. In addition all models include the following baseline controls: municipality-by-cohort fixed effects, gender, dummy variables for month of birth, dummy variables for mother's and father's educational attainment, parental income, mother's age at child's birth, indicators for being a first or second generation Nordic immigrant, indicators for being a first or second generation non-Nordic immigrant, an indicator for having separated parents, and the number of siblings. Standard errors adjusted for clustering by enrollment count (84 clusters) are in parentheses. ***/**/*=the estimates are significantly different from zero at the 1/5/10 percent level, respectively.

spond to larger classes by moving to another school. This response is larger among low-income parents than among high-income parents, but not significantly so. The parental responses we document imply that the total policy effect is different from the direct effect of reducing class size.

We have also looked at whether increases in class size affect *other school inputs* and *student effort*. The overall picture is that class size has negligible impacts on all these outcomes, this is true for all students together but also for low-income and high-income students separately.

We have identified two forces that may explain why achievement among low-income students is more susceptible to class size variation than achievement among high-income students. First, low-income pupils experience a deterioration of the learning environment when class size is increased. Second, high-income parents are more likely (and presumably more apt) to compensate for an increase in class size. By opening the black box of class size reductions we have thus gained a better understanding of the distributional impact of class size variation.

References

- Albornoz, F., Berlinski, S., and Cabrales, A. (2011). Motivation, resources and the organization of the school system. Unpublished working paper.
- Angrist, J. D. and Pischke, J. V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, 114(2):533–575.
- Bonesrønning, H. (2004). The determinants of parental effort in education production: Do parents respond to changes in class size? *Economics of Education Review*, 23:1–9.
- Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K., and Sundararaman, V. (2013). School inputs, household substitution, and test scores. *American Economic Journal: Applied Economics*, 5:29–57.
- Datar, A. and Mason, B. (2008). Do reductions in class size "crowd out" parental investment in education? *Economics of Education Review*, 27:712–723.
- Du Rietz, L., Lundgren, U. P., and Wennås, O. (1987). Ansvarsfördelning och styrning på skolområdet. DsU 1987:1, Stockholm: Ministry of Education.

- Duflo, E., Dupas, P., and Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, 101(5):1739–1774.
- Duncan, O. D. and Duncan, B. (1955). A methodological analysis of segregation indices. *American Sociological Review*, 20:210–217.
- Finn, J. D., Pannozzo, G. M., and Achilles, C. M. (2003). The "why's" of class size: Student behavior in small classes. *Review of Educational Research*, 73(3):321–368.
- Fredriksson, P., Öckert, B., and Oosterbeek, H. (2012). The devil is in the (institutional) detail: Sorting and the RD design in a public school system. Unpublished manuscript, Stockholm University.
- Fredriksson, P., Öckert, B., and Oosterbeek, H. (2013). Long-term effects of class size. *Quarterly Journal of Economics*, 128(1):249–285.
- Gelber, A. and Isen, A. (2013). Children's schooling and parents' behavior: Evidence from the Head Start Impact Study. *Journal of Public Economics*, 101:25–38.
- Hanushek, E. A., Kain, J. F., and G, S. (2004). Disruption versus tiebout improvement: the costs and benefits of switching schools. *Journal of Public Economics*, 88:1721–1746.
- Heinesen, E. (2010). Estimating class-size effects using within-school variation in subject-specific classes. *Economic Journal*, 120:737–760.
- Houtenville, A. J. and Conway, K. S. (2008). Parental effort, school resources, and student achievement. *Journal of Human Resources*, XLIII:437–453.
- Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *Quarterly Journal of Economics*, 115(4):1239–1285.
- Härnquist, K. (2000). Evaluation through follow-up. In Jansson, C., editor, *Seven Swedish Longitudinal Studies in the Behavioral Sciences*. Forskningsrådsnämnden, Stockholm.
- Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, 114(2):497–532.
- Lazear, E. (2001). Educational production. *Quarterly Journal of Economics*, 116(3):777–803.

- Lee, D. S. and Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142:655–674.
- Leuven, E. and Rønning, M. (2012). Classroom grade composition and pupil achievement. Working Paper.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142:698–714.
- Pop-Eleches, C. and Urquiola, M. (2013). Going to a better school: Effects and behavioral responses. *American Economic Review*, 103:1289–1324.
- Sims, D. (2008). A strategic response to class size reduction: Combination classes and student achievement in California. *Journal of Policy Analysis and Management*, 27:457–478.
- Summers, A. A. and Wolfe, B. L. (1977). Do schools make a difference? *American Economic Review*, 67:639–652.
- Todd, P. and Wolpin, K. I. (2003). Towards a unified approach for modeling the production function for cognitive achievement. Features:F3–F33.
- Urquiola, M. (2006). Identifying class size effects in developing countries: Evidence from rural Bolivia. *Review of Economics and Statistics*, 88(1):171–176.
- Urquiola, M. and Verhoogen, E. (2009). Class-size caps, sorting, and the regression-discontinuity design. *American Economic Review*, 99:179–215.

Appendix

Table A1. Descriptive statistics

Variable [#students; #districts×cohorts; #clusters]	All	Income parents	
		Low	High
Academic achievement, age 13 (SD) [4712; 187; 80]	0.033 (1.002)	-0.171 (1.027)	0.210 (0.945)
Cognitive ability, age 13 (SD) [5201; 197; 82]	0.005 (1.023)	-0.223 (1.040)	0.201 (0.965)
Teacher thinks that exams are important for pupils' learning (rank 0/1) [1852; 93; 56]	0.503 (0.278)	0.508 (0.281)	0.499 (0.277)
Teacher thinks that regular homework is important for pupils' learning (rank 0/1) [1853; 93; 56]	0.473 (0.281)	0.470 (0.286)	0.475 (0.277)
Teacher thinks that it is important that pupils take responsibility for their own learning (rank 0/1) [1853; 93; 56]	0.516 (0.184)	0.514 (0.184)	0.517 (0.184)
Pupil thinks about other things in school [2758; 83; 56]	0.607 (0.488)	0.602 (0.490)	0.612 (0.488)
Pupil thinks it is easy to understand when the teacher explains in front of the class [2719; 83; 52]	0.873 (0.333)	0.855 (0.352)	0.888 (0.315)
Pupil asks the teacher for help if (s)he do not understand [2802; 83; 52]	0.956 (0.204)	0.956 (0.206)	0.957 (0.202)
Remedial training [4389; 162; 75]	0.191 (0.393)	0.231 (0.421)	0.155 (0.362)
Age integrated class [6028; 200; 82]	0.042 (0.177)	0.047 (0.187)	0.038 (0.167)
Parental income of peers [6028; 200; 82]	0.048 (0.058)	0.050 (0.061)	0.047 (0.055)
Parental education of peers [6028; 200; 82]	0.058 (0.080)	0.058 (0.079)	0.058 (0.081)

Table A1: continued

Variable [#students; #districts×cohorts; #clusters]	All	Income parents	
		Low	High
Teacher experience (years) [5920; 197; 82]	10.767 (4.290)	10.759 (4.265)	10.774 (4.312)
Teacher education [5920; 200; 82]	0.958 (0.074)	0.958 (0.076)	0.959 (0.072)
Pupil gives up if (s)he gets a difficult task in school (rank 0/1) [4706; 178; 78]	0.494 (0.235)	0.502 (0.241)	0.486 (0.230)
Pupil absent (hours) [3163; 83; 52]	25.747 (33.453)	26.603 (33.761)	24.998 (33.173)
Pupil does homework (hours) [5154; 197; 82]	2.735 (1.669)	2.718 (1.716)	2.749 (1.628)
Parents help child with homework [5112; 197; 82]	0.798 (0.401)	0.772 (0.420)	0.821 (0.383)
Parents move school districts [4389; 162; 75]	0.089 (0.285)	0.109 (0.311)	0.071 (0.258)
Class size grades 4-6	24.116 (3.642)	23.822 (3.718)	24.380 (3.552)

Note: The table show means (standard deviations) of outcome variables and are based on representative samples of individuals born in 1967, 1972, 1977 or 1982 in one-school districts.

Table A2. First-stage estimates, different enrollment controls

	(1)	(2)	(3)	(4)	(5)	(6)
Above 1st threshold	-6.455*** (1.369)	-8.372*** (1.471)	-10.042*** (0.900)	-9.200*** (1.490)	-9.383*** (1.256)	-9.569*** (1.963)
Above 2nd threshold	-6.106*** (0.889)	-5.821** (0.841)	-5.197*** (1.063)	-8.107*** (1.177)	-7.008*** (0.958)	-7.843*** (1.446)
Above 3rd threshold	-6.060*** (1.015)	-3.646*** (1.517)	-4.660*** (1.284)	-2.679 (2.620)	-3.618 (2.245)	-0.369 (3.317)
F-test for instruments	18.31	20.48	47.89	39.49	40.82	18.16
<u>Enrollment controls</u>						
Polynomials:						
- 1st order	√		√		√	
- 2nd order		√		√		√
Interacted with thresholds			√	√	√	√
Interacted with segments					√	√
Number of districts × cohorts	200	200	200	200	200	200

Note: The estimates are based on representative samples of individuals born in 1967, 1972, 1977 or 1982 in one-school districts. All models include fixed effects for school district enrollment in grade 4, municipality-by-cohort fixed effects, gender, dummy variables for month of birth, dummy variables for mother's and father's educational attainment, parental income, mother's age at child's birth, indicators for being a first or second generation Nordic immigrant, indicators for being a first or second generation non-Nordic immigrant, an indicator for having separated parents, and the number of siblings. Standard errors adjusted for clustering by enrollment count (84 clusters) are in parentheses. ***/**/*=the estimates are significantly different from zero at the 1/5/10 percent level, respectively.

Table A3. IV estimates of class size, different enrollment controls

	(1)	(2)	(3)	(4)	(5)	(6)
		Academic achievement (n=4,707)				
Class size grades 4-6	-0.0592*** (0.0154)	-0.0516*** (0.0134)	-0.0441*** (0.0122)	-0.0419*** (0.0119)	-0.0415*** (0.0114)	-0.0607*** (0.0270)
Class size grades 4-6		Easy to understand when the teacher explains before the class (n=2,719)				
	-0.0077*** (0.0026)	-0.0046** (0.0020)	-0.0055*** (0.0019)	-0.0094*** (0.0030)	-0.0107*** (0.0026)	-0.0044 (0.0068)
Class size grades 4-6		Pupil does not give up if (s)he gets a difficult task in school (n=4,701)				
	-0.4783 (0.3243)	-0.4082 (0.3169)	-0.2019 (0.2253)	-0.1437 (0.2141)	-0.2643 (0.2129)	-0.1854 (0.3320)
Class size grades 4-6		Parents help child with homework (n=5,107)				
	0.0063 (0.0046)	0.0097** (0.0043)	0.0086** (0.0043)	0.0029 (0.0041)	0.0057 (0.0039)	0.0037 (0.0052)
<u>Enrollment controls</u>						
Polynomials:						
- 1st order	√		√		√	√
- 2nd order		√		√		√
Interacted with thresholds			√	√	√	√
Interacted with segments					√	√
Number of districts × cohorts	200	200	200	200	200	200

Note: The estimates are based on representative samples of individuals born in 1967, 1972, 1977 or 1982 in one-school districts. Average class size in grades 4-6 is instrumented by indicators for being above 1st, 2nd, or 3rd threshold of the class size rule. All models include fixed effects for school district enrollment in grade 4, municipality-by-cohort fixed effects, gender, dummy variables for month of birth, dummy variables for mother's and father's educational attainment, parental income, mother's age at child's birth, indicators for being a first or second generation Nordic immigrant, indicators for being a first or second generation non-Nordic immigrant, an indicator for having separated parents, and the number of siblings. Standard errors adjusted for clustering by enrollment counts (84 clusters) are in parentheses. ***/**/*=the estimates are significantly different from zero at the 1/5/10 percent level, respectively.

Table A4. IV-estimates of class size on different outcomes

Outcome [# individuals]	Current spec (1)	QJE-spec (2)
Cognitive ability, age 13 [N=5,197; N=5,116]	-0.0268** (0.0116)	-0.0330** (0.0146)
Non-cognitive ability, age 13 [N=4,740; N=4,681]	-0.0177 (0.0113)	-0.0265** (0.0118)
Academic achievement, age 16 [N=5,377; N=5,318]	-0.0265*** (0.0085)	-0.0233*** (0.0101)
Years of schooling, ages 27-42 [N=5,669; N=5,588]	-0.0318 (0.0234)	-0.0545* (0.0256)
P(Bachelor's degree), ages 27-42 [N=5,669; N=5,588]	-0.0059 (0.0042)	-0.0076* (0.0043)
Earnings, ages 27-42 [N=6,009; 5,920]	-0.0149** (0.0057)	-0.0117* (0.0061)
P(earnings>0), ages 27-42 [N=6,009; N=5,920]	-0.0035 (0.0022)	-0.0016 (0.0024)
ln(Wage), ages 27-42 [N=3,227; N=3,185]	-0.0065*** (0.0022)	-0.0063* (0.0033)
Number of districts×cohorts	200	191

Note: The estimates are based on representative samples of individuals born in 1967, 1972, 1977 or 1982 in one-school districts. All ability measures are standardized. The educational outcomes are measured in 2009, while the labor market outcomes have been averaged over the 2007-2009 period. Earnings effects (and their standard errors) are divided by average earnings level to facilitate interpretation. The ln(wage) estimates are restricted to wage-earners. Average class size in grades 4-6 is instrumented by indicators for being above 1st, 2nd, or 3rd threshold of the class size rule in column (1). In column (2), average class size in grades 4-6 is instrumented by an indicator for being above any threshold of the class size rule. All models include the following controls for school district enrollment in grade 4: fixed effects for enrollment segment; linear controls for enrollment which are interacted with threshold. In column (2), the controls for enrollment are also interacted with segment. In addition all models include the following baseline controls: municipality-by-cohort fixed effects, gender, dummy variables for month of birth, dummy variables for mother's and father's educational attainment, parental income, mother's age at child's birth, indicators for being a first or second generation Nordic immigrant, indicators for being a first or second generation non-Nordic immigrant, an indicator for having separated parents, and the number of siblings. Standard errors adjusted for clustering by enrollment count (84 clusters) are in parentheses. ***/**/*=the estimates are significantly different from zero at the 1/5/10 percent level, respectively.

Table A5. Construction of outcome variables

Table	Outcome	Measure
2	Academic achievement, age 13 Cognitive ability, age 13	Standardized average achievement in Swedish and mathematics Standardized average verbal and logical abilities
3	Teacher thinks that exams are important for pupils' learning Teacher thinks that regular homework is important for pupils' learning Teacher thinks that it is important that pupils take responsibility for their own learning	Percentile ranked teacher response divided by 100 Percentile ranked teacher response divided by 100 Percentile ranked teacher response divided by 100
4	Pupil think about other things in school Pupil think it is easy to understand when the teacher explains in front of the class Pupil ask the teacher for help if (s)he do not understand	Dummy variable for student response Dummy variable for student response Dummy variable for student response
5	Remedial training Age integrated class Parental income of peers Parental education of peers Teacher experience Teacher education	Dummy variable for receiving remedial training Share of pupils in integrated classes in school Dissimilarity index of between-class sorting of peers by parental income Dissimilarity index of between-class sorting of peers by parental education Average experience among teachers in the school Share of teachers in the school with college education
6	Share of teachers who change school districts Share of teachers who leave occupation	Share of teachers who change school districts Share of teachers who leave occupation
7	Pupil give up if (s)he gets a difficult task in school Pupil absent (hours) Pupil do homework (hours)	Percentile ranked student response divided by 100 Number of hours absent from school in grade 6 Student response on the number of hours of homework The values for the 1967 cohort have been predicted from response on frequency of homework
8	Parents help child with homework P(Move from school district)	Dummy variable for student response on receiving help with homework Dummy variable for moving from school district

Note: This table shows which outcome variables appear in which table in the main text and how these are measured.

Table A6. Data coverage for different cohorts

	1967	1972	1977	1982
Academic achievement, age 13	✓	✓	✓	✓
Cognitive ability, age 13	✓	✓	✓	✓
Teacher thinks that exams are important for pupils' learning				✓
Teacher thinks that regular homework is important for pupils' learning				✓
Teacher thinks that it is important that pupils take responsibility for their own learning				✓
Pupil thinks about other things in school	✓	✓		
Pupil thinks it is easy to understand when the teacher explains in front of the class	✓	✓		
Pupil asks the teacher for help if (s)he do not understand	✓	✓		
Remedial training		✓	✓	✓
Age integrated class	✓	✓	✓	✓
Parental income of peers	✓	✓	✓	✓
Parental education of peers	✓	✓	✓	✓
Teacher experience	✓	✓	✓	✓
Teacher education	✓	✓	✓	✓