

Koppensteiner, Martin

Conference Paper

Class assignment and peer effects: evidence from brazilian primary schools

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Evaluation Econometrics, No. A6-V2

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Koppensteiner, Martin (2010) : Class assignment and peer effects: evidence from brazilian primary schools, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Evaluation Econometrics, No. A6-V2, Verein für Socialpolitik, Frankfurt a. M.

This Version is available at:

<https://hdl.handle.net/10419/37426>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

CLASS ASSIGNMENT AND PEER EFFECTS: EVIDENCE FROM BRAZILIAN PRIMARY SCHOOLS

Martin Foureaux Koppensteiner*

Preliminary version – 1st August 2010
Please do not cite without permission.

Abstract

I propose a novel way of identifying peer group effects by employing a regression-discontinuity design that makes use of the assignment mechanism of students into classes in Brazilian primary schools. I estimate a significant negative effect for students close to the cut-off point from being in the older class in the size of about 0.5 standard deviations in math achievement. The paper provides evidence that differences in the variation of age and ability between classes lead to changes of teaching practices and lead to behavioural changes of students and make educational production less efficient in more heterogeneous classes.

JEL: I20, I21

Keywords: Peer effects, regression discontinuity, educational production, group heterogeneity.

* Economics Department, Queen Mary University of London
Correspondence address: Queen Mary, Queens Building, Mile End Road, London E1 4NS, UK.
Email: m.koppensteiner@qmul.ac.uk. Telephone: +44 207 882 8843.

1. INTRODUCTION

The question whether the composition of the peer group matters for the outcome of an individual member of the group has received considerable attention in many contexts of social interactions. Peer effects have been studied in the context of schools, universities, work place, neighbourhoods, and prisons among other with a wide range of outcomes being looked at.¹ Due to the nature of grouping students into schools, cohorts and classrooms and the potential of education policies in interfering in the peer group composition, peer effects in education have received extensive attention by economists.²

The identification of peer effects is nevertheless difficult due to conceptual problems as well as data limitations. Adhering to the education example, an identification strategy for peer effects needs to address potential endogenous selection of students into peer groups at the school and class level. With selection into peer groups, unobserved characteristics such as ability, parental support or student effort are likely to be correlated among peers and educational outcomes are therefore correlated within the peer group even in the absence of externalities associated with peer effects.³ In addition, the analysis needs to deal with separating peer effects from common shocks to the peer group, such as differential educational and teacher inputs and it needs to account for simultaneous determination of student and peer achievement (Manski 1993, Hanushek et al. 2003).

The attempt to estimating peer effects using cross-sectional variation in peer characteristics and outcomes is problematic, as the formation and hence the variation in the peer group composition is subject to selection.⁴ Variation in mean peer characteristics between classes or between schools in the same cohort needs to be treated with caution, as the group composition may be based on observable and unobservable characteristics of the students. Previous research has approached the selection problem by either controlling for observable characteristics, by estimating selection models, or by comparing siblings of families that move homes and are therefore subject to different

¹ Recent studies include Mas and Morretti (2009) for productivity effects on supermarket cashiers, Bandiera, Barankay and Rasul (2010) on social networks and worker productivity in farm production, Guryan, Kroft and Notowidigdo (2009) on the productivity of professional golf players, Bayer, Hjalmarsson and Pozen (2009) on the effect of juvenile offenders serving time on other's subsequent criminal behaviour to name just a few.

² Studies on peer effects in education include Hoxby (2000) for gender and race peer effects, Hanushek et al (2003) provide a framework for estimating peer effects trying to overcome omitted variables and simultaneous equation biases, Duflo, Dupas and Kremer (2008) provide evidence from a randomized experiment in Kenya, Lavy, Paserman and Schlosser (2008) on ability peer effects and potential channels, Lavy, Silva and Weinhardt (2009) on distributional effects of ability peer effects, Lavy and Schlosser (2007) on gender peer effects and their operational channels, Zimmerman (2003) and Sacerdote (2003) for peer effects in college education, Angrist and Lang (2004) for peer effects on racial integration and Ammermueller and Pischke (2009) for a cross-country comparison of peer effects at primary school level. Student tracking, school choice, busing, admission policies, class formation, repetition policies, and residential location decisions are relevant policy issues that can change the peer composition at school and classrooms (Zimmerman 2003 and Hanushek et. al 2003).

³ See Burke and Sass (2008).

⁴ The choice between (public and private) schools in Brazil depends strongly on the socio-economic status of the parents, so that between-school variation is contaminated by selection. Also, access to different quality public schools is often by neighbourhood, which again is self-selected or determined by income and socio-economic status.

schools (Hoxby 2000). These methods nevertheless prove unconvincing as they do not offer credible ways of dealing with unobservables or rely on very restrictive identifying assumptions. Randomized experiments are the first choice for overcoming the selection problem and there are recent uses of experimental research designs by Duflo, Dupas and Kremer (2008) on ability grouping and Whitmore (2005) on gender peer effects in higher education. Empirical strategies using natural experiments, such as conditional random assignment of college roommates by Zimmerman (1999) and Sacerdote (2003), or the use of idiosyncratic gender and race variation on the cohort level by Hoxby (2000) have also been proposed.⁵

There is still little experimental or quasi-experimental evidence that overcomes these important problems in the identification of peer group effects in primary or secondary education. This paper provides quasi-experimental evidence from exogenous variation in the peer group composition by using the assignment mechanism of students into classes which provides for the use of a regression discontinuity (RD) design. In the majority of the primary schools of the sample, students are allocated to classes using the relative age in the cohort as assignment criterion. Using the age rank as a continuous assignment variable, this creates a discontinuity in the allocation mechanism to a classroom (peer group) for the marginal students close to the class size cap. This is a novel approach in the identification of peer effects that helps to overcome the above outlined selection problem. Similarly to the prevalent use of RD designs in estimating treatment effects, where treatment is determined by whether the observable forcing variable exceeds a known threshold, the mechanism that assigns students to classes according to their age rank in the cohort creates treatment variation that can be ‘as good as random’ for individuals close to the class cut-off point.⁶ The proposed identification strategy differs nevertheless in some dimensions from typical uses of the RD design. Rather than estimating the effect of a predefined homogenous treatment or programme, in the present case treatment varies from school to school through the assignment to classes with varying mean characteristics. Furthermore, multiple discontinuities that are due to the school specificity of the cut-off point, allow identifying treatment effects over a range of the support of the discontinuity variable.⁷

⁵ An additional difficulty in estimating peer effects can arise from the simultaneity effect, known as reflection problem and outlined by Manski (1993). The reflection problem arises as student i 's outcomes may not only be affected by endogenous and exogenous peer effects, but may affect simultaneously outcomes of peers as well. I nevertheless do not consider the effect of student i 's outcome on peers' outcomes, so that the reflection problem is not relevant in this context.

⁶ See Lee and Lemieux (2009) for a comprehensive list of RD applications in economics.

⁷ Black, Galdo and Smith (2007) make use of multiple cut-offs for an RD design evaluating an employment programme in Kentucky.

As the identification strategy is nested within schools and the dataset allows controlling for a wide set of class-level characteristics, in particular teacher characteristics, the possibility that the results are driven by correlated effects on the school or class-level can be excluded.

Using two-stage-least squares to estimate the discontinuity in a fuzzy RD setup, I find evidence for considerable peer group effects. I estimate a negative effect of being in the relatively older class of around 0.5 of a standard deviation in math test scores. Though being challenging to separate the mechanisms through which peer effects work, I attempt to provide an understanding of the mechanisms at work by using information on differences in peer and teacher behaviour. The analysis of heterogeneous effects across schools provides additional evidence on the channels of estimated peer effect.

The remainder of this paper is organized as follows: Section 2 briefly describes the Brazilian educational system and the educational system in Minas Gerais. Section 3 describes the data. Section 4 presents the assignment mechanism of students into classes and introduces the identification strategy. Section 5 presents the main results and gives an interpretation of the peer group estimates. Section 6 presents tests for non-random sorting and for correlated effects and section 7 concludes.

2. THE EDUCATIONAL SYSTEM IN BRAZIL AND MINAS GERAIS STATE

Primary schooling is compulsory in Brazil for children from 6 years of age and consists of nine years of schooling. Children that turn six by 31st March of a given year are required to commence primary schooling.⁸ Brazil has a largely decentralized education system. Public schools are either under the administrative control of each state's Secretariat of Education (SEE) or under the control of municipal authorities. The federal administration is responsible for coordinating educational policies working in partnership with State and Municipal Education Secretariats and monitoring the comprehensive system of educational funds between the federal and state level. State schools account for more than half (55%) of all public schools and the vast majority of public schools in this analysis are in urban settings (91%). Allocation of students to public schools is based on the area of residence in such a way that parents cannot choose a particular school for their children within the system of public education. There exists a sizeable private sector engagement in the provision of primary schooling but as private institutions do charge substantial fees, access to private schools is limited to children from middle- and high-income families.⁹ Public schools are free of charge at all ages.

⁸ The Brazilian school year largely coincides with the calendar year. See also the data annex for the creation of the student age variable.

⁹ Around 10% of school children in Minas Gerais attend private schools. Source: School Census 2007.

In the 2006 wave of the Programme for International Student Assessment (PISA) Brazilian students at age 15 rank at the bottom of countries tested with a mean math test scores of 393 being considerably below the OECD average of 492, which shows that despite substantial improvements over the last ten years there remain concerns with the quality of primary and secondary schooling in Brazil (OECD 2006).

3. DATA AND DESCRIPTIVE STATISTICS

I use standardized test scores in mathematics of primary school students at 5th grade in public schools in the state of Minas Gerais, the second most populous state in the South-east of Brazil, as a measure for educational production. Educational standards in Minas Gerais are among the highest compared to other states in Brazil.¹⁰ The data used for the analysis comes from two data sets that are linked by school and class identifiers.

In 1999 the SEE has introduced the State System for the Evaluation of Public Education (SIMAVE), which includes the Programme of Evaluation of Basic Education (PROEB) and focus on the evaluation of student performance in primary and secondary school. The standardized math test score data stems from PROEB and for this study I use the wave of 2007 as it contains the most detailed information on student age, compared to previous waves of the test. The test is carried out at all public schools including state and municipal schools in the state of Minas Gerais and test scores are standardized to a mean of 500 with a standard deviation of 100. All classes of a given grade at each school are examined and participation is compulsory at the school and individual levels confirmed by a high participation rate of 93%. All pupils answer a detailed socio-economic questionnaire, which includes information on sex, month and year of birth, racial background and information on the socio-economic background of the family. Table 1 presents summary statistics of these variables. Average age of students on the test date is 11.27 years, which is about $\frac{3}{4}$ of a year above the appropriate age of this grade. This age-grade mismatch is due to a combination of late enrolment and grade repetition, mostly at third grade.¹¹

PROEB also includes headmaster and teacher questionnaires. The headmaster questionnaire includes questions on individual characteristics of the headmaster, such as age, sex and educational background and questions on school characteristics and pedagogic decisions at the school. The teacher questionnaire includes questions on individual characteristics, as well as statements of the students in class.

¹⁰ In the nation-wide school evaluation system of SAEB 2005 mean math performance of pupils from Minas Gerais is clearly above the Brazilian average, ranking on top of Brazilian states (http://www.inep.gov.br/salas/download/prova_brasil/Resultados/Saeb_resultados95_05_UF.pdf).

¹¹ More details can be found in the data annex.

The second source of data comes from the 2007 school census in Brazil, which is conducted by the National Institute for the Study and Research on Education (INEP) for the Federal Ministry of Education (MEC) and comprises detailed information on school characteristics. It compiles data from all primary schools in Brazil in cooperation with the states' secretariats of education and the municipal authorities. Summary statistics for the schools used in this analysis are presented in table A1 in the annex. The data comprises 16,031 students from 363 public primary schools. Students at these schools are overwhelmingly from deprived socio-economic family backgrounds. 47% of the families of the students at these schools are recipients of Bolsa Família, a Brazilian conditional cash transfer programme for poor and very poor families, compared with around 25% in the total population.¹²

4. EMPIRICAL STRATEGY

4.1 Assignment of students into classrooms

In this paper I combine the mechanism to divide enrolment cohorts based on a maximum class-size cap with the assignment mechanism of students into classes in primary schools to identify class peer effects on individual student performance.¹³

In the setting of public schools of Minas Gerais, when the number of students per entry cohort plus repeaters in grade one exceeds multiples of 25 students per class,¹⁴ the student cohort is to be divided into the appropriate number of classes. At exact multiples of 25 this theoretically creates a straightforward relationship between the enrolment cohort size and the number of classes with exactly 25 students.¹⁵ At enrolment cohort sizes different from exact multiples of 25, there is some flexibility of the school administration on whether to form classes of equal size.

With two (or more) classes per grade the school administration needs to make a choice on how to assign students into classes before the start of the school year. The allocation of students into classes

¹² Families are eligible for Bolsa Família, if per capita family income is not above R\$ 120 ('moderately poor')(US\$ 63 at 1st June 2007) and receive monthly R\$ 20 per child under the condition of regular school attendance and participation in vaccination campaigns. Families below a per capita income R\$ 60 ('extremely poor') receive an additional basic family allowance of R\$ 62. See <http://www.mds.gov.br/bolsafamilia/> and Lindert et al. (2007) for details.

¹³ Angrist and Lavy (1999) have first introduced the use of an RD design exploiting the discontinuity induced by a maximum class-size rule to estimate the effect of class-size on student proficiency. They use *Maimonides' maximum class-size rule* of 40 students which establishes a discontinuous relationship at multiples of 40 in the total number of students in the cohort to estimate class-size effects. Different from Angrist and Lavy I do not use the discontinuous relationship between the cohort size and class-size at the class-size cap using schools with cohort sizes close around multiples of the maximum class-size number. Instead I make use of the discontinuities in the assignment of students to either of the two classes created by the combination of a (school-specific) class-size cap and the age ranking of students to identify peer group effects on students close to the class-caps.

¹⁴ Law 16.056 of 24th April 2006 limits class size to 25 students in the initial years of fundamental education (1st-5th grade) in all public schools in Minas Gerais. Exceptions are theoretically only allowed in extenuating circumstances and during the transitory period of the introduction of the law.

http://crv.educacao.mg.gov.br/sistema_crv/banco_objetos_crv/%7B103FA0DB-B47A-4E66-A719-402B21F94D5B%7D_lei%2016056%202006.pdf

¹⁵ I provide more details on the mechanism in the annex.

could be done in a (quasi) random fashion, for example alphabetically by family name, or could follow strategic sorting into classes. In the case of public Brazilian primary schools, in which age variation within student cohorts is considerably larger compared to schools in many other countries or private schools in Brazil,¹⁶ an allocation rule that sorts students to form more homogeneous age groups is an option readily at hand for the assignment of students.¹⁷ It is often postulated that classes with smaller within-class variation in age make instruction easier for teachers and the education production process more efficient.¹⁸ As age of students at the point of enrolment in first grade can be easily observed by school administrators, differently from innate ability or other behavioural characteristics, age sorting provides a convenient way of grouping students along observable characteristics. School headmasters in Minas Gerais are free in choosing the allocation mechanism at their school and around 70% of the schools in the sample assign students to classes according to age to form more homogenous classes, with the remainder using alternative ways of allocation that lead to more heterogeneous age distributions in the classes.

I restrict the analysis in this paper on schools with two classes per grade level. Although not being representative for all schools tested in PROEB 2007, restricting the analysis to schools with two classes per grade level ensures that school administrators cannot establish ‘special’ classes that do not follow the general assignment mechanism. With larger cohorts the school administration may resort to form extra classes, in which only grade repeaters or disadvantaged students are brought together and are separated from all other students in the cohort. By restricting the analysis to schools with two classes, the school administration very unlikely can form such special classes, which reduces the risk of misinterpreting the findings on peer effects.

In the case of two classes per school with a cohort size between 26 and 50 students, the classes are not necessarily of equal size, potentially leading to an endogenous cut-off point. Older classes are in fact on average 2.58 students smaller than the younger classes, with a mean class size of 22.85 and 25.43, respectively.¹⁹ Furthermore, the class-size cap is not strictly enforced in all cases, which may also contribute to the difference in class size. 31% of all classes exceed the maximum class-size cap

¹⁶ The large age variation is due to a combination of late enrolment, temporary drop-out and grade repetition.

¹⁷ There is an extensive pedagogic literature on age, ability grouping, and academic tracking, but little work from economists on theoretical foundations or empirical analyses of the impact of the different forms of forming classes on the mean and the distribution of performance. See Robinson (2008), Adams-Byers, Squiller Whitsell and Moon (2004), and Betts and Shkolnik (1999) for some recent examples. Kremer (1997) provides an economic model of sorting.

¹⁸ See Hoxby and Weingarth (2006) for a discussion. Grouping students according to their age may in fact at least partially coincide with grouping according to ability, as ability likely is correlated with age at time of primary school enrolment. See Cascio & Schanzenbach (2007) and Angrist & Krueger (1991) for a discussion of student age and educational outcomes.

¹⁹ Lazear (2001) points out that optimal class size varies directly with student behaviour and classes with more disruptive students are often found to be smaller. See Lazear (2001) for a theoretical behavioural model on class-size choice and West and Wößmann (2006) for an empirical analysis on student sorting and endogenous class-size.

of 25 students. Only few classes (9%) nevertheless exceed the class size cap by more than 10% and 97.5% of all classes are smaller than 33 students.

The assignment mechanism that uses a smooth function in age to order students in a given cohort creates a discontinuity in class membership at the actual class-size cap of the younger class. The upper graph of figure 1 plots standardized local averages of the class rank in one month bins, where the age rank of students is expressed as distance from the cut-off point and local linear regression fits using a rectangular kernel with a bandwidth of 3 months are superimposed. The discontinuity in the average class rank at the cut-off point is evident and the size of the discontinuity in the probability of treatment conditional on the age rank being smaller than one makes the case for the fuzzy RD design.²⁰

As outlined above, the division of an enrolment cohort will not necessarily follow a strict rule that equally divides the students in the cohort over two classes. The discontinuity point at the class-cap is therefore potentially endogenous, which might cause concerns when using the cut-off point for identification in a RD design. I will test for the plausibility of the identifying assumptions later and I will show that non-random sorting around the discontinuity point or the strategic choice of the exact threshold do not pose a threat to the identification strategy employed. I use actual class size of the younger class to determine the discontinuity point between the two classes and denote the rank of the student at the school specific class-cap of the younger class as \bar{N} .²¹ Because there are as many cut-off points as there are schools, I have re-centred the data to obtain a unique discontinuity point at zero across all schools and the individual age rank of the students in the cohort is reported as distance from zero in months.

4.2 Regression discontinuity design

The identification strategy exploits the discontinuity in the assignment rule of students in schools with two classes. As outlined above the treatment assignment mechanism depends on the value of an observed and continuous variable, the age rank n of the individual student in each school in such a way that the probability to receive treatment is a discontinuous function of that variable at the class-size cap \bar{N} . Identification of the treatment effect arises from the fact that just below and above the known cut-off point, individuals are similar in observable and unobservable characteristics, but are

²⁰ With the younger class given a value of 0 and the older class a value of 1. The local linear regression lines are fitted separately on both sides of the threshold. The values are centred around zero taking into account school fixed effects, so that deviations from zero are represented.

²¹ Potentially endogenous discontinuity points have been used for regression discontinuity strategies elsewhere in the literature; see for example Card, Mas & Rothstein (2008).

members of classes with different peer groups.²² A key difference of the regression discontinuity approach to randomization is that while randomization guarantees that treatment and control group are similar with respect to characteristics other than the treatment variable, the regression-discontinuity design in this setup makes use of treatment and control group being different in the mean values of age and other student characteristics at the class level.²³

As there are assignment imperfections due to the above outlined reasons, the assignment to treatment depends for the individual close enough to the threshold on n in a stochastic manner, but in such a way that the probability of treatment has a discontinuity at \bar{N} that varies between schools. Furthermore, I pool all schools together in the analysis irrespective of the allocation mechanism chosen. This leads to the case of a fuzzy regression-discontinuity design, where the size of the discontinuity is smaller than one.

Consider a simple reduced-form model of an education production function

$$(1) \quad Y_{is} = \delta_0 + \delta_1 T_i + f(n) + \varepsilon_i$$

where Y_{is} denotes the outcome variable math test score for individual i in school s , and T_i is the treatment indicator that takes a value of 0 for individuals in the younger class and 1 for individuals in the older class, ε_i is an individual unobserved error component, ignoring at this stage any covariates one might want to include in the specification to reduce sampling variability in the estimator. Educational achievement measured in test scores, depends on a smooth function $f(\cdot)$ representing the age rank of student i in the cohort that is constructed using the age of individual students, and on being in either the younger or older class indicated by T_i . I employ two-stage least squares to estimate δ_1 , the coefficient of interest using the discontinuity at the class cap as an instrument for treatment T_i (being in the older class).

In a first stage-equation, I assume that T_i is a smooth function of age rank of students in the cohort and a dummy D_{is} for being above or below the school-specific discontinuity point \bar{N} given by the maximum class-size rule.

$$(2) \quad T_i = \gamma_1 + \gamma_2 D_{is} + f(n) + \nu$$

where ν is an error component.

²² For example with a different fraction of repeaters and male students, and potentially students with a different socio-economic family background.

²³ See van der Klaauw 2002.

For identification of the class peer effect δ_1 , a continuity assumption needs to be satisfied, such that student achievement varies continuously with the forcing variable of the age rank in the cohort, outside of its influence through treatment T_i (Lee & Lemieux 2009), such that assignment to either side of the discontinuity threshold is as good as random.

I estimate the above first- and second-stage equations by OLS and the discontinuity by 2SLS, modelling $f(n)$ as a low-order polynomial parametrically. Because of the school specific cut-off point, which varies with cohort size, rather than having just one discontinuity, there are as many discontinuities as there are schools in the dataset, which is an exceptional feature for a RD design. Since, across schools, the cut-off points are at different absolute ages, the estimates are robust to non-linearities in age (Duflo, Dupas and Kremer 2009).

Public knowledge of the allocation mechanism and the alleged advantage of treatment may invalidate the above continuity assumption crucial to the regression-discontinuity design if, because of that, the forcing variable is subject to manipulation by optimizing agents (McCrary 2008). In the present context there is potential for manipulation of the forcing variable by two sets of agents involved, the parents of the school children and school administrators. If either parents or school administrators are able to manipulate the rank of a student precisely, the ‘as good as random’ assignment may fail. To invalidate the conditions for the consistency of the regression-discontinuity design nevertheless requires precise control over the forcing variable. Starting with the parent’s case two forms of manipulation involving the forcing variable may invalidate the above assumption for the RD design. Theoretically parents could declare a different age of their child to the school administration. To place their child into a specific class at time of first grade enrolment, parents need to have knowledge of the age distribution of the other students in the entry cohort and of the cut-off point \bar{N} . With the school specific cut-off points this creates some *ex-ante* uncertainty about the allocation into classes. Even if parents were successful in placing their child in their preferred class by manipulating the declared age, this invalidates the assumption for the RD only in a case of precise control over the resulting age rank so that the student is placed exactly at the cut-off point.²⁴ Generally, though, parents need to submit an official document, for example a birth-certificate, when they register their child at school, which makes this form of manipulation very unlikely. A more relevant threat for the identification strategy arises from the potential of parents exerting pressure on school administrators to assign their child to the younger class at initial enrolment or at a later stage.

²⁴ There is nevertheless no advantage to the students of being placed exactly at the cut-off point; on the contrary, for parents to be sure of placing their child in the younger class, strong underreporting of true age is likely. Furthermore, the enrolment process at first grade involves some form of official identification, so that manipulation of the age of the child is further impeded. Also, if several parents attempt to manipulate the age rank of their child, this most likely affects the predetermined cut-off point \bar{N} as a function of n , inevitably resulting in imprecise control over the forcing variable.

The consequences of a reassignment of a student from the older class into the younger class are more severe to the identification strategy, as such misplaced students automatically rank directly at the cut-off point in the class-specific age-rank, which would automatically fulfil the *complete manipulation* case necessary for invalidating the RD assumption of continuity.²⁵

McCrary (2008) suggests a test for the failure of the random assignment assumption by inspecting for a discontinuity in the density of the forcing variable around the discontinuity point. As the forcing variable in the present case is nevertheless uniformly distributed due to its nature of a relative rank, this test will not be informative in this analysis.²⁶ If students were strategically re-allocated from the older to the younger class, average age conditional on the age rank at the cut-off point would reveal a peak; likewise if students from the younger class were to be taken to the older class, mean age of students would slope down when approaching the threshold from the right. If the selection of students were then related to performance, this would impair the validity of the RD design. By closely examining mean age of students around the threshold one can test for strategic reassignment of students around the discontinuity point. Furthermore, I test for the balancing properties of a wide range of pre-determined individual characteristics. A discontinuity in the distribution of baseline individual and family characteristics may also be an indication for manipulation, as these observable characteristics are likely related to the effort of parents for manipulation (van der Klaauw 2008).

Considering the role of school administrators in the process of allocating students to classes, another issue is important for the validity of the RD identification strategy. The discontinuity threshold \bar{N} is at least partially under the control of the school administration such that it may be shifted along the age rank of students to select students with specific characteristics into either of the two classes without breaking up the age ranking of the students. If, for a example, given a pre-selected threshold, the school administration would like to include the youngest student of the older class rather into the younger class based on some observable characteristics, the cut-off point could simply be shifted by one more rank upwards. In reality this is unlikely to happen as the allocation of students and with it the precise class-size is decided before classes start at first grade, so that the school administration has no information on ability, race or socio-economic background of the student other than administrative information such as age or sex obvious from documents necessary for enrolment, such as a birth certificate.²⁷ If the selection of the cut-off point by the school administration took into

²⁵ This is true for re-allocating students from the older class into the younger class and vice versa. Being then the oldest or youngest in the class rank, these students will automatically rank on either side next to the threshold.

²⁶ This holds true depending on cohort size and the distribution of age for age ranks relatively close to the threshold.

²⁷ Which in fact is used for the class allocation of students. As I outline in detail in the annex, there is some evidence for the endogeneity of \bar{N} , as class-size, which is under the control of the school administration of the older class is smaller compared to the younger class. Although this may systematically affect the learning environment of students in both

account any observable characteristics of students, this would lead to a jump in any pre-determined characteristics at the cut-off point. The examination of baseline covariates will therefore be an important exercise to test the validity of the RD design.

5. RESULTS

5.1. Main estimation results

In panel B of figure 1, I plot local averages of math test scores and the local linear regression lines on both sides of the cut-off point from which the clear drop in math test scores is apparent. Table 2 presents the first-stage estimates for the size of the discontinuity in mean class rank, the OLS estimates for the size of the discontinuity in test scores at the discontinuity point and the 2SLS estimates for the causal effect of crossing the cut-off point from the younger class to the older class.²⁸ All specifications include school-fixed effects that account for observed and unobserved differences between schools.²⁹ Standard errors are heteroskedasticity consistent and adjusted for clustering at the school level. Column (1) presents the estimates for the models including only a quadratic polynomial in age. Column (2) includes controls for the whole set of predetermined individual and family characteristics. The estimates of column (3) include teacher characteristics in addition to the other covariates.

The top panel of table 2 presents estimates for the first stage regressions, where the dependent variable is 1 for students being in the older class and zero otherwise. The estimates for the size of the discontinuity range between 0.451 and 0.467, giving a measure for the discontinuity in panel A of figure 1. The inclusion of controls in columns (2) and (3) leaves the first-stage estimates unaltered.

The middle panel of table 2 reports the reduced form estimates from an OLS regression with math test scores as the dependent variable on a dummy equal 1 for being to the right of the threshold. Column (1) reports the raw estimate of the discontinuity of math test scores at the cut-off point of panel A of figure 1.

The bottom panel of table 2 reports the two-stage-least squares estimates for the class peer effects using the same specifications as for the OLS estimates in panels A and B. The size of the class peer effect, without further controls, is around 0.57 of a standard deviation in math test scores and significant at the 1% level.

classes in the form of a correlated effect, this nevertheless does not necessarily violate the above continuity condition required for the consistency of estimator. I will discuss the potential bias of this correlated effect of class size on the estimates later.

²⁸ Imbens and Angrist (1994) show that the jump in outcomes identifies a local average treatment effect (LATE) at the discontinuity point.

²⁹ A great advantage of the proposed identification strategy arises from the fact that actual variation in peer group characteristics is sufficiently large in particular when compared to identification strategies that rely on idiosyncratic variation across cohorts. Using school fixed effects increases the standard errors of the estimates considerably, as more parameters need to be estimated resulting in less degrees of freedom.

Under the weak identifying assumption for the regression-discontinuity design outlined in the previous section, the results can be interpreted as the causal effect on individuals whose treatment status changes, i.e. who were to switch from the younger class to the older class as the value of n changes from just to the left of \bar{N} to just to the right of \bar{N} . Students close enough to the right of the cut-off point are very similar in their characteristics compared to the students just to the left of the cut-off point, but are faced with a different peer group composed of a larger share of repeaters, a higher proportion of male classmates, a peer group with a lower overall socio-economic family background and a class with greater heterogeneity in age. The coefficients for the RD estimates in table 2 show that, on average, marginal students experience substantially lower academic achievement under treatment.

An important question is the channel through which this negative peer effect operates. The effects could either be driven by the lower average academic performance of students in the older class captured by the lower socio-economic background of students, or by the behaviour of students as well and the response of teaching practices that differ due to the peer composition and class heterogeneity. I will discuss the potential channels through which peer effects may operate in detail in the following sections.

To acquire some knowledge about the distribution of effects across schools, I estimate school-specific discontinuities for the peer effects. As differences of mean peer variables between classes differ across schools, treatment also differs in respect to the composition of the peer class environment. Figure 2 plots the kernel density estimates of the school specific discontinuities and shows the relatively symmetric distribution of effects around a peak at about -50. The distribution of the estimated peer effects across schools follows quite closely the superimposed normal distribution, shifted to the left of zero. I will later use the fact that the regression-discontinuity design is based on school specific discontinuities to learn about the impact of different channels of peer effects.

Table 4 presents the RD estimates for wider intervals of the discontinuity sample around the cut-off point and different orders of the polynomial terms included in the regressions as first robustness checks. Rows (1) and (2) are the estimates of the RD without any further controls, rows (3) and (4) are the estimates including the full set of controls including individual, family and teacher characteristics. The estimates do not reveal any substantial sensitivity with respect to the choice of the order of the polynomial. Replacing the quadratic by a cubic term leaves the estimates virtually unchanged. Increasing the range of observations used for the estimation also does not alter the estimates for the treatment effect in any relevant way.

5.2. Exogenous peer characteristics and direct peer effects

Column (2) of table 6 reports the estimates for the mean value of peer variables for students around the cut-off point.³⁰ The first row reports the difference in peer age in the classrooms and the second row the difference in mean months repeated by students in the class. Students in the older class are on average about 8 months older, which is almost completely due to the higher share of repeaters in these classes.³¹ The remainder is due to late enrolment at first grade and temporary drop-out from school followed by reenrolment later.

Repeaters and students that enrol late at first grade often come from families of a more deprived socio-economic background,³² which causes socio-economic indicators of peer students to be systematically different between the two classes. As discernible in figures 3 to 5, the mean of some of these variables decrease (or increase) smoothly with the age rank in the cohort across the discontinuity point, e.g. the fraction of girls in the class, the fraction of white students, or the number of books or computers per household. This leads to a discontinuity in the peer values of these characteristics similar to the discontinuity of the treatment variable of figure 1 (panel B), which are presented in columns 2 and 4 of figures 3 to 5. The discontinuity is most apparent for peer age and average number of months repeated of peer students in figure 3, but can be observed for a range of socio-economic characteristics and some racial attributes. RD estimates, using the same specification as in panel C of table 2, for these pre-determined characteristics report a statistically significant discontinuity of peer characteristics for students around the cut-off point. Overall, students to the right of the cut-off point while not being different from students just to the left on the whole range of individual and parental characteristics, are in classes with a peer group that consists of fewer girls, a higher fraction of black and a lower fraction of mixed students, and students from a more deprived socio-economic background. The estimated discontinuity in math test scores may, at least partially, be the result of being in class with less able students. In his pioneering work, Manski (1993) considers effects for which achievement is a function of pre-determined peer characteristics as exogenous effects. It is often assumed that peer characteristics such as sex, race and socio-economic status are proxies for (unobserved) peer ability and that exogenous peer effects work through being grouped with less able peers. The academic achievement of marginal students might suffer because there are fewer students from whom to learn or fewer students that ask stimulating questions in class.³³

³⁰ The mean values for all other students in the class less the students around the discontinuity point in the discontinuity sample.

³¹ Calculation based on the theoretical enrolment age of students and the number of years repeated by students show that repetition accounts for about 75% of the total age-grade mismatch.

³² See for example Patrinos and Psacharopoulos (1996) and Gomes-Neto and Hanushek (1994).

³³ So that positive spillovers may lead to aggregate coefficients being greater than individual coefficients (Glaeser, Scheinkman and Sacerdote 2003). See also Carell and Hoekstra (2010).

The estimated differences in peer characteristics for marginal students are nevertheless moderate in size when expressed in standard deviations of the variables. The drop in the fraction of girls, for example, in the peer group is about 0.19 of a standard deviation of the variable sex.³⁴ Compared to peer effects estimates in the literature associated with a one standard deviation change in the student background variable, the direct effect of peers through differences in exogenous peer characteristics can hardly explain the full effect on academic achievement in the present case. Ammermueller and Pischke (2009) find that a full standard deviation change in the student background measure leads to a 0.17 standard deviation change in reading test scores of fourth graders in six European countries, an effect size considerably smaller than the one estimated here.³⁵

Manski (1993) argues that exogenous peer effects, being based on pre-determined peer characteristics, do not generate a social multiplier through feedback loops.³⁶ Under the assumption that effects in exogenous peer characteristics are linear, grouping students by these characteristics would not have a positive effect on mean achievement in the cohort, as any gain to some students from being with peers of favourable socio-economic background would be offset by losses elsewhere.³⁷ Having this in mind, it is surprising that a large fraction of school administrators choose to group students to form more homogenous groups in age. Allowing the dispersion of student characteristics to have a non-linear effect may be able to explain this fact.

5.3 Classroom heterogeneity and indirect effects of the peer group

As pointed out in the previous section, classmates of marginal students differ in terms of their mean age and the number of repeaters in the class. Besides mean age, the distribution of age in the class also differs considerably between the two classes. The standard deviation of age is by 3.5 months around 40% larger in the older classes (table 7, row 16). The standard deviation of the standardized local averages of age is also more than double the size in the older classes.³⁸ With the larger number of repeaters and multiple repeaters the older classes are considerably more heterogeneous in age compared to the younger classes.³⁹ The graphs for age and years repeated in figure 3 confirm that, on average, the distribution differs between the two classes with relatively flat local linear regression fit

³⁴ The other significantly estimated discontinuities in peer characteristics range between 0.08 (number of freezers) to 0.36 (number of books) of a standard deviation in these variables with an average of 0.16 of a standard deviation in all student background characteristics.

³⁵ Similarly, Hoxby (2000) and McEwan (2003) find effects for a standard deviation change in peer characteristics close to Ammermueller and Pischke (2009).

³⁶ Where positive spillovers lead to aggregate coefficients being greater than individual coefficients (Glaeser, Scheinkman and Sacerdote 2003).

³⁷ See Burke and Sass (2008).

³⁸ Estimating standard deviations for the centred local averages separately for the two classes, weighted by the inverse of the observations (ranks) variance. Standard deviation for the mean of the local averages of age of the younger classes is 3.566 and 8.684 for the older classes.

³⁹ Graphs A1 and A2 in the annex show the histograms for age of students in months for the two classes and confirm that younger classes are more homogenous in age than the older classes.

for the younger classes and a fit for the older classes with a positive slope. The allocation of students by relative age in the cohort creates classes with very different age distributions, relatively more homogenous younger classes and relatively more heterogeneous older classes. Even though the aim of age grouping is to reduce age heterogeneity within classes, in schools with two classes this actually results in classes being considerably different in terms of the distribution of age.

The difference in heterogeneity of age is matched by a similar difference in the distribution of test scores between the two classes. Chart B of figure 2 shows that, on average, test scores in the younger classes are not correlated with age rank, whereas in the older classes there is clear negative relationship between the age rank and math test scores, such that the standard deviation of the standardized local averages of test scores is three times larger for the older classes.⁴⁰

Heterogeneity of peer groups as a key factor in education production has not received much attention in the peer effects literature. This is probably due to identification strategies that rely on relatively small idiosyncratic variation in student cohorts or experimental random variation in the peer group, which are not based on large variation in the distribution of characteristics of the peer group. Group heterogeneity as a determinant of educational outcomes has though been addressed in the literature on tracking. More recent research on the effects of tracking, addressing the endogeneity of tracking decisions, finds that tracking may benefit equally students from lower and higher achievement tracks. Figlio and Page (2002) find that tracking may actually help low-ability students without proposing a mechanism for this effect and Zimmer (2003) presents quasi-experiment evidence that a negative direct peer effect for low-achieving students is offset by positive effects of achievement targeted instruction, both using U.S. data. Duflo, Dupas and Kremer (2010) use random assignment of schools to study the effect of tracking students on initial achievement for Kenyan primary school students. They find persistent positive effects across the achievement distribution of tracking students in a higher and a lower ability track. They attribute this effect mainly to teacher effort and the choice of target teaching level given the particular incentives for teachers in Kenyan schools and the better match of the instruction level due to reduced heterogeneity in ability in the classrooms.⁴¹ Their results are matched by the findings of Zimmer (2003) and Hoxby and Weingarth (2006) who show that students in more homogenous classes benefit from more tailored instruction. Interestingly, in an identification strategy similar to the one employed here, Duflo, Dupas and Kremer (2010) do

⁴⁰ 5.45 for the younger and 16.37 for the older classes. This negative relationship is likely due to the larger share of repeaters and students that have enrolled late.

⁴¹ They provide some evidence that teacher payoffs in Kenya are convex in student test scores, thus teachers exerting more effort in the top section than in the bottom section. There is no evidence for similar incentive payoffs in the case of public schools in Minas Gerais, as promotion is almost exclusively based on seniority and student test scores have no relevance for teacher compensation or promotion. This is confirmed by the fact that there exist no observable differences between teachers of the different classes in the following section, while convex teacher payoffs would suggest better teachers to be found with the better students.

not find evidence for a discontinuity in test scores for students close to the 50th percentile in initial test scores that separates the cohort into two classes. Their results suggest no relevant role for direct peer effects in this setting.⁴²

The strong differences in the variation of age and ability between the classes may help to explain the pronounced negative effects for students around the cut-off point. If homogeneity of age and ability of students in a class enable teachers to target the teaching level better, students benefit from being in a more homogeneous environment, independent of the level of mean age or ability. Heterogeneity, on the contrary, may lead to teachers needing to target teaching to groups of students of different ability or age separately and answer questions from students of such diverse background, which may not benefit students of the other groups.

In the following section, I present evidence on the effect of the difference in class composition and heterogeneity on teaching practices and student class behaviour.

5.4. Response of teaching practices and students to class composition

I use information from the student questionnaire in which students report on items related to teaching practices and the behaviour of their peer students in class.⁴³ Table 3 presents marginal effects from a linear probability model on each of the variables for which separate regressions have been estimated.

The estimates reveal that the perception of teaching practices by students differs substantially between the classes, whilst there exist no observable differences in teacher characteristics.⁴⁴ As the answers to the questions are of a potentially subjective nature, there is the possibility that the answers not only reflect differences in the learning environment, but may also reflect the differences in the composition of students in the classes. To check whether the estimates may also reflect the difference in student composition, specification (2) comprises controls for the student composition, including sex, race and socio-economic composition of the classes. The estimates do not differ in a relevant way when including mean student characteristics and the pairwise difference of the coefficients is not statistically significantly different from zero for any of the variables, so that there is no indication for sensitivity of reporting to the compositional differences of the classes.⁴⁵

⁴² Their regression discontinuity design is based on the difference in average peer achievement, rather than sorting on age. They do not report distributional differences in either achievement or age between the lower and upper classes.

⁴³ Students report on the frequency with which they agree on each of the items and for the estimations these frequencies have been aggregated on the class-level to report probabilities.

⁴⁴ Results from a balancing test on teacher characteristics are presented in section 6.2.

⁴⁵ I therefore focus solely on specification (1) in the interpretation of the estimates in table 5. This does not exclude the possibility that the questions reflect perceptions of students, rather than solely an objective evaluation of teaching practices, but largely excludes the possibility that subjective perceptions differ between the classes.

The first two items in table 3 do not show a significant difference between the two classes. The questions whether teachers enforce attention of students in class and whether they correct homework are more directly related to teacher attitude and motivation, rather than being the result of an adaption of teaching practices to the class composition. This is in line with the findings in table 7, that there is no systematic allocation of teachers so that teacher characteristics and attitudes do not differ between the classes.

The estimates relating to questions on teaching practices for the following five items in table 3 reveal that there exist substantial differences in the teaching practices between the classes. In the older class the probability that the teacher is available to clarify doubts and questions is about 12% lower than in the younger class. Similarly, students in the older class feel that the opportunity to express their opinion in class is substantially lower. The difference in instructor time available for individual students is consistent with the hypothesis that class heterogeneity obstructs efficiency in class room teaching. Heterogeneity in age and ability of students demand teachers to address different skill levels and apply different learning techniques, which may lead to the application of a skill-specific targeting of teaching to specific groups of students in the same class. Strong evidence for an effect on teaching practices through the impact on the distribution of instruction time is given by the results on students' perception that the class teacher helps some students more than others. The estimate for this variable reports a 25% difference between the classes and shows that teachers in the older class are forced to distribute their attention and instructional time more unequally, possibly devoting relatively more time for specific groups of students and addressing the same material targeted at different skills levels. More evidence in the same direction comes from the question on whether teachers show an equal interest in the attainment of all students. Teachers in the older class are less able to turn their attention equally towards all students, which again is in line with the hypothesis that greater heterogeneity leads to a more unequal distribution of instructional time and teacher attention. The difference in class composition also leads to teachers being less able to explain the subject matter until every student has comprehended it (-0.082). This may nevertheless also be the consequence of having a larger fraction of repeaters in the older class.⁴⁶ In support of the hypothesis that teaching is less effective due to the absorption of instructional time by heterogeneous groups of students, the proportion of the planned curriculum actually taught during the school year is by about 4% lower for the older classes (table 7, class teacher statements).

In addition to the above findings on the differences in the teaching practices information from the student questionnaire also reveal significant differences in learning environment of the classroom. Students in older classes consider their classmates more often as noisy and disruptive (0.115%). This

⁴⁶ Although the negative estimate persists when controlling for the share of repeaters and even mean performance of students (not reported here).

may be due to a combination of the direct effect of the on average more disadvantaged peers, and a consequence of the less targeted teaching practices. With a more heterogeneous student composition and skill-targeted teaching, so that teachers may need to spend more time individually on low ability students, more idle time for the remainder of students may lead to more disruptive behaviour and noise on the part of these students.⁴⁷ Together with the higher share of disadvantaged students in the older class, this may lead to a more discouraging class experience overall. The probability of students leaving class early is also substantially higher in the older classes, which is a potential result of the less encouraging learning environment, and may lead to even more disruption of teaching in these classes. The less favourable teaching environment is also confirmed by students in the older class reporting more often that their teacher needs to wait to start teaching at the beginning of class because of noise.

The less favourable teaching environment may also have an effect on teacher motivation. Students of the older class report more often (0.122) that a teacher has been absent from school. It is nevertheless difficult to quantify the effect, as even a single day of teacher absence may be reported by students in the answer. This is in line with the results of Duflo, Dupas and Kremer (2010) that teacher absence is higher when assigned to the lower track classes. Higher absence may nevertheless not be driven by the curvature of the teacher payoff function as in their interpretation. In the absence of payoffs that are convex in student test scores, as in the Brazilian case, the effect on absence of teachers can be interpreted as a response to the more deprived and demanding teaching environment. In turn, although difficult to quantify in terms of hours lost of instruction, teacher absence may also impact on achievement of students, creating feedback effects between class composition, teacher and student behaviour.

The difference in class behaviour reported by students is confirmed by information from the teacher questionnaire. Teachers in the older classes more often report disciplinary problems with their students (14%); the estimate is nevertheless only significant at the 10% level of significance (class teacher statements, table 7).

The estimates emphasise the importance of the interaction between the peer group composition and its heterogeneity in age and skills on the one hand, and the response of teaching practices and teacher and student behaviour on the other hand, for explaining peer effects in education. Similarly to the findings of Lavy, Paserman and Schlosser (2008), and Duflo, Dupas and Kremer (2010), the above results suggest that the teaching technology seems to respond to the peer composition and may be an important channel for explaining the negative peer effect for students close to the class threshold.

⁴⁷ Interestingly, students from the entire age rank in the older class report a higher level of noise and disruption, not only marginal students close to the threshold, which suggests that behavioural changes are not only due to the higher share of repeaters. The inclusion of the socio-economic controls, including the fraction of repeaters, in fact even increases the coefficient, although the difference between specifications (1) and (2) in table 3 is not statistically significant.

It is nevertheless not apparent from the analysis whether the differences in teaching practices and class behaviour are induced by the greater heterogeneity in age and achievement in the older classes, or by the higher propensity for individual students in the older class to disrupt class teaching in line with the *Bad Apple* model of peer effects (Carrell and Hoekstra 2010, Hoxby and Weingarth 2006, Lazear 2001) and to what extent there is an interaction between the two. In the following section I use the heterogeneity of treatment across schools to shed light on the role of direct and indirect peer effects.

5.5 Heterogeneous treatment across schools and channels of peer effects

The previous sections propose different potential channels through which the peer composition in this setting leads to a strong drop in academic performance of students close to the cut-off point. It remains a challenging task to distinguish between direct and indirect effects and to understand the role of class heterogeneity.

The unique setup for the RD design in this case nevertheless allows a closer examination of the relevance of different aspects of the peer composition for the estimated gap in academic achievement. The fact that there are multiple discontinuities in the data set that differ from school to school, as well as differences in the mean peer characteristics that differ between schools, can be used to get an understanding of what determines heterogeneity of effects across schools.

In a two-step procedure I first obtain the size of the discontinuity in test scores and the difference in peer characteristics between the two classes for each school. In a second step I use the estimated gaps in test scores as dependent variable and regress these by minimum-distance estimation on the set of peer differences and other school characteristics, weighting observations by the inverse of the standard errors of the test score estimates. Column (1) in table 5 reports the coefficient for a model that includes only the difference in the heterogeneity in age between the two classes as a regressor, measured as the difference of the standard deviation of age in months. A difference of one month in the standard deviation of age is accountable for 4.6% of a standard deviation of math test scores and the coefficient is significant at the 1% level. This is direct evidence on the importance of age heterogeneity of the peers for academic achievement and is supportive of a peer effects model of class heterogeneity. The estimate of the difference in absolute age between the two classes on the test score gap is very small and not statistically significant. This is important, as there is no evidence that the effects are driven by an age specific curriculum for the grade that could be responsible for the behavioural responses of students in the older class as outlined in the previous section. Also, the coefficient for the difference in mean years repeated by students in each class is not significant.⁴⁸

⁴⁸ Alternatively, using the difference of the fraction of repeaters between the classes does not yield any different result.

Although a considerable part of the differences in mean and variation of age is due to the different fraction of repeaters in the two classes, the presence of repeaters does not seem to drive the peer effects in this setting, but rather the age variation they cause in the class. There is therefore little support for a *bad apple* model of peer effects that postulates that few disruptive repeaters can negatively affect the outcome of all the other students in the class in this setting.⁴⁹ Specification (5) includes differences of the full set of peer characteristics. The coefficient for age variation increases only slightly. With the exception of the difference in the mean number of books available in the students' households, none of the other coefficients is significantly different from zero at conventional levels.⁵⁰ Neither the sex, racial nor socioeconomic composition of students seems to be relevant for explaining the differences in estimated peer effects across schools. Specification (6) includes further controls for school characteristics as described in table A1, and only the coefficient on age variation remains significant, though being somewhat smaller than in specifications (1) and (5). The coefficient on the number of books nevertheless ceases being significant when controlling for other school characteristics.

The results of this exercise leave no major role for a direct peer effect on academic achievement. Although there are differences in the mean socio-economic composition of students between the classes as discussed in section 5.2, they cannot explain the sharp drop in math achievement at the cut-off point. Only the coefficient on the difference in the mean number of books, which is likely the most direct measure of parental education shows a significant effect, that nevertheless ceases when controlling for the full set of school characteristics. There is no evidence that being with less able peers drives the negative effect on students just to the right of the cut-off point. Evidence presented here rather suggests that the difference in the variation of age and ability has a key role in explaining the estimated peer group effect.

6. TESTS FOR NON-RANDOM SORTING AND CORRELATED EFFECTS

The key identifying assumption for the estimation in the regression discontinuity design is that around the class cut-off point, assignment is as good as random so that predetermined characteristics of students are balanced on both sides of the discontinuity point. Although a discontinuity in few of the pre-determined covariates does not necessarily invalidate the identification assumptions of the

⁴⁹ Carrell and Hoekstra (2010) provide some evidence in support of a *bad apple* model in the setting of students subject to domestic violence. Lavy, Paserman and Schlosser (2008) discuss the relevance of class disruption associated with the proportion of repeaters in the classroom.

⁵⁰ The sign of the coefficient on the difference in mean number of books is negative, because the number of books in the older classes is smaller than in the younger classes on average. The coefficient of the difference of mean numbers of freezers per household is significant at the 10% level of significance.

RD design, this may at least cast some doubt on the estimation strategy or may be an indication for misspecification of the functional form. I use rich information from the student questionnaire to formally test for the balancing properties of pre-determined student characteristics across the cut-off point.

Furthermore, one would like to exclude the possibility that the estimated peer effect is driven by *correlated effects* or common shocks. If, for example, the learning environment of the classes is systematically different this could bias the estimates. This includes the potential for strategic allocation of teachers with different qualities to the different classes, the differential provision of teaching material or different class-size. I propose to test for non-random allocation of teachers to classes and of some other classroom characteristics. I also use information from the headmaster questionnaire to learn about the allocation procedure of teachers.

6.1 Student and family characteristics

Figures 3-5 provide an informative graphical analysis of the balancing properties of baseline covariates by plotting local averages for the covariates and local linear regression fits separately on both sides of the threshold.⁵¹ The top left corner graph of figure 3 presents local averages for individual age of students in months. As outlined above, the strategic allocation of students from one class to the other would inevitably result in a discontinuous representation of average age at the cut-off point. Mean age nevertheless increases smoothly across the discontinuity point, so that there is no evidence for any strategic reallocation of students between the two classes. In figure 3, the graphs in columns (1) and (3) plot the fraction of girls and the racial composition of students around the cut-off point. The local fraction of girls reduces smoothly with age rank leading to a lower average of girls in the older class. The fraction of white, Asian and indigenous students does not reveal any discontinuity at the threshold, while the fraction of mixed and black students suggests a minor jump at the cut-off point. The average number of years repeated also does not reveal a discontinuity, but different slopes of the local linear regression fits are apparent, induced by the different distribution of repeaters in the two classes. Columns (1) and (3) of figures 4 and 5 present the same graphical representation of the local linear regression fits and local averages for a wide range of predetermined socioeconomic characteristics. These variables appear well balanced on both sides of the cut-off point and there is little indication for a discontinuity in the means of these characteristics at the cut-off point. From two additional proxies for the socio-economic status of the family, the number of domestic workers employed and the fraction of families receiving Bolsa Família, only the latter suggests a small difference around the threshold (figure 4).

⁵¹ Using the same specification as for the graphs in figure 1. All graphs are conditional on school fixed effects and centred so that the mean deviation from 0 is plotted.

In a formal analysis, I estimate all predetermined characteristics of students using the same specification as for the main estimates in table 2. Table 6 reports the estimates for these variables. Only the RD estimate for the fraction of black students is significant on the 5% level.⁵² None of the other socioeconomic characteristics of student households reveals a statistically significant difference at the threshold and most estimates are small, confirming that the balancing properties of these predetermined characteristics are satisfied. The estimate for age reveals the smooth transition across the threshold, and there is no sign of a discontinuity suggestive of non-random sorting of students around the threshold.

From the RD estimates of the pre-determined covariates and the inspection of the graphs there is no indication for discontinuities and non-random sorting of students around the cut-off point that would impair the identifying assumptions of the RD design. Although the absence of discontinuities in predetermined individual and family characteristics cannot prove the balancing property of unobservables, it is reassuring to find that individuals on both sides of the cut-off are observationally equivalent.⁵³

The inclusion of these additional individual and family controls in column (3) of table 2 changes the estimates in the reduced-form regressions nevertheless to some extent. Likewise, the IV estimates of the class peer effect are around 20% smaller than without these controls, leaving some role for individual level heterogeneity in the estimation of the peer group effect.⁵⁴ Even though the composition of students is balanced to the left and the right of the threshold, the effect of being in a specific peer environment may have differential impact on specific groups of students. As I do not find any discontinuity in the predetermined characteristics at the cut-off point, the reduction in the estimated peer effect may to some extent also be due to model misspecification with the inclusion of the set of controls (Imbens & Lemieux 2008).

6.2 Teacher characteristics and class environment

Another concern for the estimation of class peer effects is that correlated effects in form of common shocks to the peer group bias the peer effect estimates.⁵⁵ Common shocks may bias the estimates of the peer effect, if the learning environment for the students in the two classes is systematically

⁵² Choosing different specifications for the RD by including either only a linear polynomial term or a cubic term makes the estimate for this variable seizing being significant, so that the single significant estimate can either be attributed to model misspecification or random chance. Any other specification for the functional form or estimating the RD without robust standard errors does not change insignificance of the estimates of any of the variables.

⁵³ Only one of 21 coefficients shows a significant discontinuity, which given the 5% level of significance could be due to random chance.

⁵⁴ Carrell and Hoekstra (2010) find differential effects of the peer characteristics by gender, Lavy, Paserman, and Schlosser (2008) find heterogeneous effects of peers across the ability distribution and Cooley (2010) finds differential impacts across racial groups.

⁵⁵ As the identification strategy is based on between class variation and I include school fixed effects, any correlated effects based on between-school variation are excluded.

different. Despite being challenging to completely rule out the existence of any differences in the learning environments, I can nonetheless assess whether there exist observable differences in the learning environment for a broad set of teacher and class characteristics.

Systematically different learning environments may be created by assigning teachers with specific qualities to either of the two classes. This may happen in a compensatory fashion, such that better teachers are allocated to weaker classes (in this case to the older classes with a higher proportion of repeaters), which would lead to an underestimation of the age peer effect. Better educated or more experienced teachers could also be allocated to the younger class to strengthen good students further, which would lead to overestimating the peer effect. This could be the case if parents of students in the younger class more successfully lobby school administrators to receive a more beneficial learning environment than parents from more disadvantaged students in the older class.⁵⁶ In their questionnaire, headmasters are asked how they generally allocate teachers to classes. The vast majority (68%) of headmasters allocates teachers in a non-systematic fashion to classes, either by a draw or by not using a specific criterion. Less than two percent of headmasters allocate more experienced teachers to better classes, and around 16% the more experienced teachers to the worse performing class. The remainder (13%) lets teachers select the classes among themselves. Assuming that more experienced teachers are higher in the teacher hierarchy and prefer to teach better performing classes (with fewer repeaters), the latter two categories may just cancel out each other. To test whether there are any systematic differences in teacher characteristics between the younger and older classes, I estimate a separate OLS regression for teacher and class characteristics on a dummy equal to one for the older class and table 7 reports the coefficients. Generally, one might like to find pre-determined mean teacher characteristics to be equal across the two classes. None of the teacher characteristics show any significant difference between the two classes, including teacher sex, age, race, experience, education and training and the estimated coefficients are mostly very small. This is reassuring, as strategic teacher allocation does not seem to play any relevant role and the allocation mechanism indicated by the answers in the headmaster questionnaire does not materialize in a measurable difference. There is no indication that more experienced or better educated teachers are assigned to the younger class, so that the results could at least partially be driven by teacher quality. Neither is there any evidence that more experienced teachers are systematically allocated in a compensatory fashion to the older classes with the greater fraction of repeaters and with more students of lower socio-economic status.⁵⁷ Including all teacher

⁵⁶ As nevertheless the allocation of class teachers to the classes is done in first grade and the class teacher stays with the class for most often the first five years of primary school, there is little scope for such action.

⁵⁷ Even though the information from the headmaster questionnaire would suggest, that, if anything, some teachers are allocated in a compensatory fashion to classes. Also, as this question refers to all grades in the school of the headmaster, it does not exclusively refer to the teachers of interest to this analysis.

characteristics in the RD estimates (table 2, column (4)) also does not change the estimate for the peer effect in any relevant way. Different from settings in which teacher wages are a function of test scores, teacher wages and promotion in public schools in Minas Gerais state are mostly determined by qualification and seniority.⁵⁸

Additional information from teacher statements about the allocation of teaching resources within the school to the classes also provides some more evidence that the estimates are not biased by common shocks to the classes creating different teaching environments. I use answers from the teacher questionnaire about the quantity of pedagogic resources available for teaching reported on the class-level by the teachers to investigate class level teaching resources. None of the variables reported by teachers, for example on the frequency of class council meetings, the quality of textbooks, the occurrence of insufficient financial resources and pedagogic resources for class teaching are significantly different between the two groups (table 7).

As outlined above, there is some concern about the class-size difference between the older and younger classes. The estimate in table 7 reveals that the number of students in the older class is on average by about 3 smaller compared to the younger class. As class-size may have an effect on student achievement this may lead to a bias in the estimation of the peer group effect. Although there is little evidence for clear positive effects of smaller classes on student achievement in the literature, there is some agreement that smaller may be beneficial (see Angrist and Lavy 1999 and Urquiola 2006); the effects reported are nevertheless relatively small and mostly refer to substantial reduction in number of students per class. In the present case the older class is on average smaller, so that this may lead to a downward bias of the true peer group effect on student outcomes. Because older class are on average smaller, there is nevertheless no threat that the estimated class peer effect may in fact (at least partially) be due to a class-size effect.

Table 7 also shows that the rate of students that do not participate in the PROEB test, due to illness or other reasons to be absent from school on the test date, differs between the two classes. The non-participation rate is about 7% higher in the older class. It is unlikely for this to have any substantial impact on the peer effect estimates, assuming that non-participants are equally drawn from the age distribution in the cohort. Even under the assumption that it is more likely for worse performing students to be absent from the test, this at most leads to underestimating the true class peer effect if students in the older class close enough to the of the cut-off point are more likely to be absent from the test.

⁵⁸ Details can be found in law No. 15 293 *Establishing the Careers of Professionals in Basic Education* in the state of Minas Gerais.

7. CONCLUSIONS

In this paper I introduce a novel way of identifying class peer effects using a continuous assignment rule of students into classes that creates a discontinuity in the membership to the peer group for students just below and above the class-size cap. I report precisely estimated negative average treatment effects of being with a more disadvantaged peer group on standardized math test scores for 5th graders in Brazilian primary schools for students close enough to the discontinuity point. Switching treatment status from the younger to the older class negatively affects math test scores for these students by around 0.5 of a standard deviation. This negative estimate can be interpreted as the effect of spillovers in the production of education that may be due to a combination of exogenous differences in the mean peer characteristics across the two classes and behavioural changes in the classroom. There is also evidence for changes in the efficiency of educational production suggested by differences in teaching practices in the two classes. To learn about the importance of the different channels for the peer effects I make use of the school specific cut-off points that lead to multiple discontinuities and the heterogeneity of treatment across schools to estimate the impact of differences in peer characteristics between classes in schools. I present evidence that suggests that the overall negative effect on marginal students to the right of the discontinuity point may to a large extent be driven by the need for adapting teaching practices to the larger variation in age. Direct peer effects based on differences in exogenous peer characteristics do not seem to play a decisive role in explaining heterogeneous effects across schools.

The estimates may need to be considered as lower bounds of the true effect of the class peer, as smaller class-size and the smaller proportion of test takers in the older class may lead to a small downward bias of the true effect.

I test the balancing properties of pre-determined student and parental characteristics by examining these for discontinuities at the cut-off point and do not find any systematic discontinuities that may violate the underlying identifying assumptions. This is particularly reassuring given the potentially endogenous class size rule, as suggested by the difference in class size between the classes. Likewise, there is no evidence for strategic behaviour of school administrators in sorting students around the threshold.

The results may have important implications for the general debate on sorting or tracking of students into classes or schools. It may be difficult to generalize the findings to tracking of students between schools, as for example the allocation of resources, the design of teaching curricula or teacher incentives may differ considerably across schools.⁵⁹ The findings that heterogeneity of the peer composition matters for the results of students across the ability distribution may nevertheless

⁵⁹ This may particularly important in school systems that track students into different types of schools, which may be of more relevance for secondary education or school systems with private school alternatives. See Jackson (2010).

provide some insight in answering the question why grouping students by age or achievement may be beneficial to all students. The results match the findings of Duflo, Dupas and Kremer (2010), who show that tracking students into classes by initial achievement benefits students at all level of the initial achievement spectrum and the results of Zimmer (2003), who finds that tracking has a positive effect even on low-achieving students in the US through more tailored instruction and can outweigh negative direct peer effects based on peer quality. The results contribute to the understanding of a potential trade-off between direct and indirect peer effects of grouping of students by age or achievement. By tracking students by ability, low achieving students may no longer benefit from the presence of high achieving peers, but benefit from the lower variation in ability. It is particularly important to consider this trade-off in educational systems where heterogeneity of students is large as it is the case in many developing countries.

The findings in the paper also contribute to the understanding of policies that aim at reducing age variation in the cohort of students. Programs that tackle late school enrolment and therefore limit age variation in cohorts of students may have a positive effect on other students in the cohort by reducing age heterogeneity in the cohort.⁶⁰ Likewise, the effect of grade retention policies may need to be considered in a peer effects framework as well. Retained students increase age variation in the cohort and may impose a negative externality on the remaining students in the class.

The results may also be important for decisions of school administrators on the allocation of teaching resources between classes. Although the school administration may not be completely aware of the effect of age grouping on marginal students, there is some evidence that the older classes are chosen to have smaller size, possibly taking into account the negative effect of the class composition. This is in line with the findings of Lazear (2001) who characterizes the choice of class-size as a function of student characteristics and shows that classes with greater propensity for interrupting ought to be smaller. There is no evidence though that the school administration allocates other educational inputs strategically between the classes to compensate for the disadvantage in the learning environment. It may be possible that more experienced teachers in combination with smaller class-size for the older classes could at least partially compensate for the disadvantage in the learning environment of students in the older classes.

⁶⁰ In the case of Brazil, the programme Bolsa Família, which pays a cash transfer conditional on school attendance (and other conditions) also contributed to the considerable reduction of late enrolment to primary school over the last decade. Glewwe and Kassouf (2008) provide some basic evidence on the effectiveness of the programme in reducing late enrolment.

REFERENCES

- Adams-Byers, Jan, Sara Squiller Whitsell, and Sidney M. Moon. 2004. "Academic and Social/Emotional Effects of Homogeneous and Heterogeneous Grouping." *Gifted Child Quarterly*, 48(7): 7-20.
- Ammermueller, Andreas, and Jörn-Steffen Pischke. 2009. "Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study." *Journal of Labor Economics*, 27(3): 315-348.
- Angrist, Joshua, and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Education and Earnings?" *Quarterly Journal of Economics*, 106(4): 979-1014.
- Angrist, Joshua, and Kevin Lang. 2004. "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." *American Economic Review*, 94(5): 1613-1634.
- Angrist, Joshua, and Victor Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Children's Academic Achievement." *Quarterly Journal of Economics*, 114(2): 533-575.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2010. "Social Incentives in the Workplace." *Review of Economic Studies*, 77(2): 417-459.
- Bayer, Patrick, Randi Hjalmarsson, and David Pozen. 2009. "Building Criminal Capital behind Bars: Peer Effects in Juvenile Corrections." *Quarterly Journal of Economics*, 124(1):105-147.
- Betts, Julian R., and Jamie L. Shkolnik. 1999. "The Effects of Ability Grouping on Student Achievement and Resource Allocation in Secondary Schools." *Economics of Education Review*, 19: 1-15.
- Black, Dan, Jose Galdo, and Jeffery Smith. 2007. "Evaluating the Bias of the Regression Discontinuity Design Using Experimental Data." *Unpublished Manuscript*.
- Burke, Mary, and Tim R. Sass. 2008. "Classroom Peer Effects and Student Achievement." *Working Papers 08-5*, Federal Reserve Bank of Boston.
- Card, David, Alexandre Mas, and Jesse Rothstein. 2008. "Tipping and the Dynamics of Segregation." *Quarterly Journal of Economics*, 123(1):177-218.
- Carell, Scott, and Mark Hoekstra. 2010. "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids." *American Economic Journal: Applied Economics*, 2(1): 211-228.
- Cascio, Elizabeth, and Diane Whitmore Schanzenbach. 2007. "First in the Class? Age and the Education Production Function." *NBER Working Paper 13663*.
- Cooley, Jane. 2010. "Desegregation and the Achievement Gap: Do Diverse Peers Help?" *Unpublished Manuscript*.
- Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2010. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review*, forthcoming.

- Figlio, David and Marianne Page. 2002. "School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality?" *Journal of Urban Economics*, 51: 497-514.
- Glewwe, Paul and Ana Lucia Kassouf. 2008. "The Impact of the Bolsa Escola/Familia Conditional Cash Transfer Program on Enrolment, Grade Promotion and Drop-out Rates in Brazil." *Unpublished Manuscript*.
- Gomes-Neto, João B. and Eric A. Hanushek. 1994. "Causes and Consequences of Grade Repetition: Evidence from Brazil." *Economic Development and Cultural Change*, 43(1): 117-148.
- Guryan, Jonathan, Kory Kroft, and Matthew J. Notowidigdo. 2009. "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments." *American Economic Journal: Applied Economics*, 1(4): 34-68.
- Hanushek, Eric, John F Kain, Jacob M Markman, and Steven G Rivkin. 2003. "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics*, 18(5): 527-544.
- Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." *NBER Working Paper 7867*.
- Hoxby, Caroline M., and Gretchen Weingarth. 2006. "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects." *Unpublished Manuscript*.
- Imbens, Guido W., and Thomas Lemieux. 2007. "Regression Discontinuity Designs: A guide to Practice." *Journal of Econometrics*, 142(2): 615-635.
- Jackson, Kirabo. 2010. "Do Students Benefit From Attending Better Schools? Evidence From Rule-based Student Assignments in Trinidad and Tobago." *The Economic Journal*. Forthcoming.
- Kremer, Michael. 1997. "How Much Does Sorting Increase Inequality," *Quarterly Journal of Economics*, 112(1): 115-139.
- Lavy, Victor, Daniele Paserman, and Analia Schlosser. 2008. "Inside the Black Box of Ability Peer Effects: Evidence from Variation in Low Achievers in the Classroom", *NBER Working Paper 14415*.
- Lavy, Victor, and Analia Schlosser. 2007. "Mechanisms and Impacts of Gender Peer Effects at School." *NBER Working paper No. W13292*.
- Lavy, Victor, Olmo Silva, and Felix Weinhardt. 2009. "The Good, the Bad and the Average: Evidence on the Scale and Nature of Ability Peer Effects in Schools." *Unpublished Manuscript*.
- Lazear, Edward P. 2001. "Educational Production." *Quarterly Journal of Economics*, 156(3): 777-803.
- Lee, David, and Thomas Lemieux. 2009. "Regression Discontinuity Designs in Economics." *NBER Working Paper No. 14723*.

- Lindert, Kathy, Anja Linder, Jason Hobbs, and Bénédicte de la Brière. 2007. "The Nuts and Bolts of Brazil's Bolsa Família Program: Implementing Conditional Cash Transfers in a Decentralized Context." *Social Protection Discussion Paper 0709*, World Bank.
- Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *American Economic Review*, 99(1): 112-145.
- Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 60: 531-542.
- McEwan, Patrick J. 2003. "Peer Effects on Student Achievement: Evidence from Chile." *Economics of Education Review*, 22:131-41.
- McCrary, Justin 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*, 142(2): 698-714.
- OECD (Organization for Economic Cooperation and Development). 2006. "*Equally Prepared for Life? How 15-year-old Boys and Girls Perform in School*". OECD. Paris.
- Patrinos, Harry A., and George Psacharopoulos. 1996. "Socioeconomic and Ethnic Determinants of Age-grade Distortion in Bolivian and Guatemalan Primary Schools." *International Journal of Educational Development*, 16 (1): 3-14.
- Robinson, Joseph P. 2008. "Evidence of a Differential Effect of Ability Grouping on the Reading Achievement Growth of Language-Minority Hispanics." *Educational Evaluation and Policy Analysis*, 30(2): 141-180.
- Sacerdote, Bruce. 2003. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics*, 116: 681-704.
- Urquiola, Miguel. 2006. "Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia." *Review of Economics and Statistics*, 88(1): 171-77.
- Urquiola, Miguel, and Eric Verhoogen. 2009. "Class-Size Caps, Sorting, and the Regression-Discontinuity Design." *American Economic Review*, 99(1): 179-215.
- Van der Klaauw, Wilbert. 2002. "Estimating the Effect of Financial Aid Offers on College Enrolment: A Regression-Discontinuity Approach." *International Economic Review*, 43(4):1249-1287.
- West, Martin R., and Ludger Wößmann. 2006. "Which School Systems Sort Weaker Students into Smaller Classes? International Evidence." *European Journal of Political Economy*, 22(4):944-968.
- Whitmore, Diana. 2005. "Resource and Peer Impacts on Girls' Academic Achievement: Evidence from a Randomized Experiment." *American Economic Review*, 95(2): 199-203.
- Zimmer, Ron. 2003. "A New Twist in the Educational Tracking Debate." *Economics of Education Review*, 22: 307-315.
- Zimmerman, David. J. 2003. "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." *Review of Economics and Statistics*, 85(1): 9-23.

TABLE 1
MEANS AND PROPORTIONS OF STUDENT AND TEACHER CHARACTERISTICS

Student characteristics		Class 1	s.e.	Class 2	s.e.	Difference	s.e.	
Sex	Female	0.524	(0.005)	0.458	(0.006)	0.066	(0.008)	
Age	(in years)	10.930	(0.009)	11.670	(0.014)	-0.740	(0.016)	
Race	White	0.306	(0.005)	0.264	(0.005)	0.042	(0.007)	
	Mixed	0.526	(0.005)	0.517	(0.006)	0.009	(0.008)	
	Black	0.097	(0.003)	0.143	(0.004)	0.046	(0.005)	
	East-Asian	0.027	(0.002)	0.034	(0.002)	-0.007	(0.003)	
	Indigenous	0.044	(0.002)	0.042	(0.002)	0.002	(0.003)	
Repeater	Never repeated	0.797	(0.004)	0.489	(0.006)	0.308	(0.007)	
	Once repeated	0.142	(0.004)	0.292	(0.005)	-0.150	(0.007)	
	Twice repeated	0.043	(0.002)	0.148	(0.004)	-0.105	(0.005)	
	Repeated 3 or more times	0.018	(0.001)	0.070	(0.003)	-0.052	(0.003)	
SES	Families with Bolsa Família	0.480	(0.005)	0.592	(0.006)	-0.112	(0.008)	
	HH with domestic aid	0.137	(0.004)	0.113	(0.004)	0.024	(0.005)	
Means	Num. of books	23.496	(0.322)	19.428	(0.330)	4.068	(0.463)	
	Num. of cars	0.608	(0.009)	0.503	(0.009)	0.105	(0.013)	
	Num. of computers	0.262	(0.005)	0.195	(0.005)	0.067	(0.007)	
	Num. of fridges	0.999	(0.005)	0.958	(0.006)	0.041	(0.008)	
	Num. of freezers	0.302	(0.006)	0.282	(0.007)	0.020	(0.009)	
	Num. of radios	1.342	(0.008)	1.286	(0.009)	0.056	(0.012)	
	Num. of TV	1.497	(0.008)	1.396	(0.009)	0.101	(0.012)	
	Num. of DVD players	0.849	(0.007)	0.786	(0.008)	0.063	(0.011)	
	Num. of bathrooms	1.246	(0.006)	1.175	(0.006)	0.071	(0.009)	
	Num. of washing machines	0.758	(0.007)	0.752	(0.007)	0.006	(0.010)	
	Num. of tumble dryers	0.168	(0.005)	0.163	(0.005)	0.005	(0.007)	
	Teacher characteristics		Class 1	s.e.	Class 2	s.e.	Difference	s.e.
	Sex	Female	0.983	(0.011)	0.965	(0.015)	0.018	(0.013)
Age	(in years)	40.495	(0.468)	40.094	(0.486)	0.401	(0.674)	
Race	White	0.456	(0.030)	0.477	(0.030)	-0.021	(0.042)	
	Mixed	0.420	(0.029)	0.399	(0.029)	0.021	(0.042)	
	Black	0.093	(0.017)	0.081	(0.016)	0.012	(0.024)	
	East-Asian	0.028	(0.010)	0.039	(0.012)	-0.011	(0.015))	
	Indigenous	0.004	(0.004)	0.004	(0.004)	0.000	(0.005)	
Highest edu. level	Secondary education	0.100	(0.018)	0.118	(0.019)	-0.018	(0.026)	
	Higher edu. – ped. degree	0.210	(0.024)	0.208	(0.024)	0.002	(0.034)	
	Higher edu. - regular	0.410	(0.029)	0.389	(0.029)	0.021	(0.041)	
	Higher edu. - licentiatuara	0.203	(0.024)	0.174	(0.022)	0.029	(0.033)	
	Higher edu. – other	0.076	(0.016)	0.111	(0.019)	-0.035	(0.024)	
	Salary (in R\$)	771.74	(22.803)	743.60	(23.754)	28.140	(32.920)	
	Years exp. in education	14.023	(0.360)	13.862	(0.375)	0.161	(0.520)	
	Years exp. at this school	8.227	(0.397)	7.257	(0.376)	0.970	(0.547)	
	Years exp. with this grade	4.697	(0.152)	4.764	(0.152)	-0.067	(0.213)	
	Part. in continued training	0.375	(0.028)	0.363	(0.029)	0.012	(0.040)	

Notes: The data from the upper panel is taken from the student background questionnaires, the data from the lower panel from the teacher questionnaires of PROEB 2007.

TABLE 2
MAIN ESTIMATION RESULTS

	(1)	(2)	(3)
Panel A: First stage			
Dependent variable: class rank			
	0.467*** (0.056)	0.453*** (0.057)	0.451*** (0.056)
R ²	0.326	0.370	0.403
Panel B: Reduced form			
Dependent variable: math test scores			
	-26.445*** (7.458)	-19.196** (7.646)	-19.513** (7.743)
R ²	0.405	0.482	0.485
Panel C: IV regression discontinuity results			
Dependent variable: Math test scores			
	-56.574*** (15.299)	-42.385*** (15.455)	-43.297*** (15.673)
R ²	0.410	0.485	0.489
Observations:	1,688	1,688	1,688
Repetition controls	no	yes	yes
Individual controls	no	yes	yes
Teacher controls	no	no	yes

Notes: The top panel reports the first stage regressions using OLS estimating equation (2). The middle panel reports the coefficient on math test score on the dummy equal 1 for the age rank larger than 0 (reduced form). Test scores are centred using school fixed effects in all specifications. The bottom panel reports IV estimates of the effect of being in the older class on math test scores, where being in the older class has been instrumented by a dummy for having an age rank larger 0. All specifications include a second-order polynomial in age. Specifications in column (2) include the whole set of predetermined individual and family characteristics, including sex, race and SES family characteristics, specifications in column (3) additionally include all predetermined teacher characteristics, including teacher sex, race, age, salary, variables on educational background and experience. Heteroskedasticity consistent standard errors are clustered by schools and reported in parenthesis. **, *** denote significance at the 5% and 1% level, respectively.

TABLE 3
RESPONSE OF TEACHING PRACTICES TO CLASS COMPOSITION

	(1)		(2)	
Teacher enforces student attention	-0.029	(0.021)	-0.022	(0.027)
Teacher corrects homework	-0.031	(0.020)	-0.048	(0.044)
Teacher availability to clarify doubts	-0.115***	(0.025)	-0.110***	(0.030)
Teacher explains until all students understand	-0.082***	(0.027)	-0.069**	(0.034)
Teacher gives opportunity to express oneself	-0.086***	(0.031)	-0.079**	(0.036)
Teacher helps more some students	0.251***	(0.043)	0.210***	(0.056)
Teacher interested in learning progress	-0.084***	(0.022)	-0.083***	(0.027)
Teacher needs to wait to start teaching	0.160***	(0.050)	0.168***	(0.059)
Teacher absenteeism	0.122***	(0.036)	0.108**	(0.044)
Fellow students classroom early	0.210***	(0.045)	0.212***	(0.059)
Fellow students are noisy and disruptive	0.115***	(0.044)	0.128***	(0.057)
Fellow students learn taught material	-0.132***	(0.029)	-0.139***	(0.039)
Fellow students pay attention in class	-0.027	(0.032)	-0.021	(0.039)
Number of observations	726		726	

Notes: Entries are OLS estimates on a dummy of classrank, where classrank=0 for the younger class and classrank=1 for the older class. Marginal effects reported. For each variable a separate regression has been estimated. The data comes from the student questionnaire of PROEB 2007. All regressions control for school fixed effects. Specification (1) does not include further controls; specification (2) includes controls for the socio-economic composition of the class, including peer racial and sex composition. Heteroskedasticity consistent standard errors are reported in parentheses. **, *** at the denote significance 5% and 1% level, respectively.

TABLE 4
RD ESTIMATES OF MATH TEST SCORES

	Ranks from threshold in months				
	1 month	2 months	3 months	4 months	5 months
	Estimated discontinuity at threshold				
Quadratic	-56.574*** (15.299)	-54.578*** (12.561)	-59.044*** (11.103)	-57.193*** (10.791)	-59.182*** (10.653)
Cubic	-55.477*** (15.551)	-54.467*** (12.622)	-59.560*** (11.106)	-57.188*** (10.842)	-58.416*** (10.722)
Quadratic with full controls	-43.297*** (15.673)	-43.762*** (12.446)	-45.216*** (11.259)	-43.600*** (10.980)	-43.066*** (10.675)
Cubic with full controls	-41.689** (16.299)	-43.753*** (12.45)	-45.625*** (11.274)	-43.769*** (11.031)	-42.726*** (10.749)
Number of observations	1688	3142	4547	5884	7223

Notes: Dependent variable is math test score and entries are estimates of the discontinuity including different range of observations in terms of the age rank indicated by the column heading. Entries for row (1) are the estimated coefficients of the RD from models that include a quadratic polynomial in age for the different range of observations. Row (2) includes a cubic polynomial. Rows (3) and (4) additionally include the full set of controls as in column (4) of table 2. Heteroskedasticity consistent standard errors are reported in parentheses. **, *** denote significance at the 5% and 1% level, respectively.

TABLE 5
HETEROGENEOUS TREATMENT ACROSS SCHOOLS

Difference in class means	(1)	(2)	(3)	(4)	(5)	(6)
Age variation	4.640*** (1.311)				4.988*** (1.589)	3.485** (1.446)
Age in months		-0.234 (1.819)			-1.695 (1.883)	-0.704 (1.625)
Years repeated			-3.739 (27.148)		-7.219 (24.341)	-2.149 (23.899)
Number of books in HH				-40.638*** (14.310)	-41.274** (19.331)	-30.246 (18.638)
Fraction of male students					20.911 (27.194)	26.985 (26.070)
Fraction of white students					-27.636 (38.310)	-21.781 (35.988)
Fraction of mixed students					-40.933 (26.400)	-35.811 (26.031)
Fraction of black students					29.564 (50.220)	23.887 (50.446)
Fraction of Asian students					-50.803 (125.308)	-81.925 (106.722)
HH with domestic workers					26.390 (50.451)	-1.419 (49.267)
HH with Bolsa Família					-32.108 (34.958)	-29.259 (35.020)
Number of bathrooms					-7.316 (34.912)	-2.999 (32.698)
Number of cars					-2.892 (27.685)	0.302 (25.877)
Number of computers					4.235 (50.686)	4.398 (43.518)
Number of fridges					-26.181 (28.694)	-16.989 (28.483)
Number of freezers					-60.881* (35.988)	-58.830* (33.266)
Number of radios					33.065 (25.721)	31.154 (24.689)
Number of washing machines					26.250 (23.605)	41.407* (22.327)
Number of tumble dryers					-75.724 (46.357)	-75.993* (44.057)
Number of DVD players					47.037 (32.465)	42.635 (32.030)
Number of TV sets					0.266 (26.057)	-3.199 (24.698)
Number of video players					48.602 (36.907)	34.795 (33.711)
School characteristics controls	No	No	No	No	No	Yes
R-squared	0.105	0.000	0.000	0.059	0.234	0.302

Notes: Dependent variables are the discontinuities in math test scores at the cut-off point on the school level. Specifications (1)-(6) report coefficients from a minimum distance estimator, weighting each observation by the inverse of the estimates' standard errors. Independent variables are differences in means of the peer values of socio-economic characteristics, class age and its distribution. Specification (6) contains controls for school characteristics as in table A1. Heteroskedasticity robust standard errors are reported in parenthesis. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

TABLE 6
RD ESTIMATES OF PREDETERMINED INDIVIDUAL AND FAMILY VARIABLES

		(1)		(2)	
		Individuals		Peers	
Fraction of	Age (in months)	0.442	(0.735)	8.157***	(0.796)
	Month repeated	0.728	(0.879)	7.487***	(0.457)
	Female	0.190	(0.127)	-0.088***	(0.019)
	White	0.008	(0.092)	-0.035	(0.023)
	Black	0.115**	(0.055)	0.089***	(0.018)
	Mixed	-0.037	(0.102)	-0.072**	(0.032)
	East-Asian	-0.026	(0.022)	0.011	(0.009)
	Indigenous	-0.076	(0.047)	-0.001	(0.009)
	Domestic helper	-0.020	(0.058)	-0.053***	(0.017)
Number of	Bolsa Família	0.165*	(0.099)	0.144***	(0.027)
	Bathrooms	-0.101	(0.098)	-0.129***	(0.033)
	Books	-4.314	(4.956)	-8.016***	(1.928)
	Cars	-0.167	(0.138)	-0.141***	(0.039)
	Computers	-0.031	(0.068)	-0.108***	(0.022)
	Fridges	0.096	(0.077)	-0.074**	(0.031)
	Freezers	-0.013	(0.087)	-0.052**	(0.025)
	Radios	0.195	(0.158)	-0.083	(0.052)
	Washing machines	0.080	(0.105)	-0.037	(0.033)
	Dryers	-0.057	(0.082)	0.014	(0.021)
	DVDs	0.125	(0.121)	-0.120***	(0.035)
	TV sets	-0.008	(0.141)	-0.194***	(0.042)
	Video players	0.080	(0.107)	-0.066**	(0.028)
	Number of observations		1,688		1,688

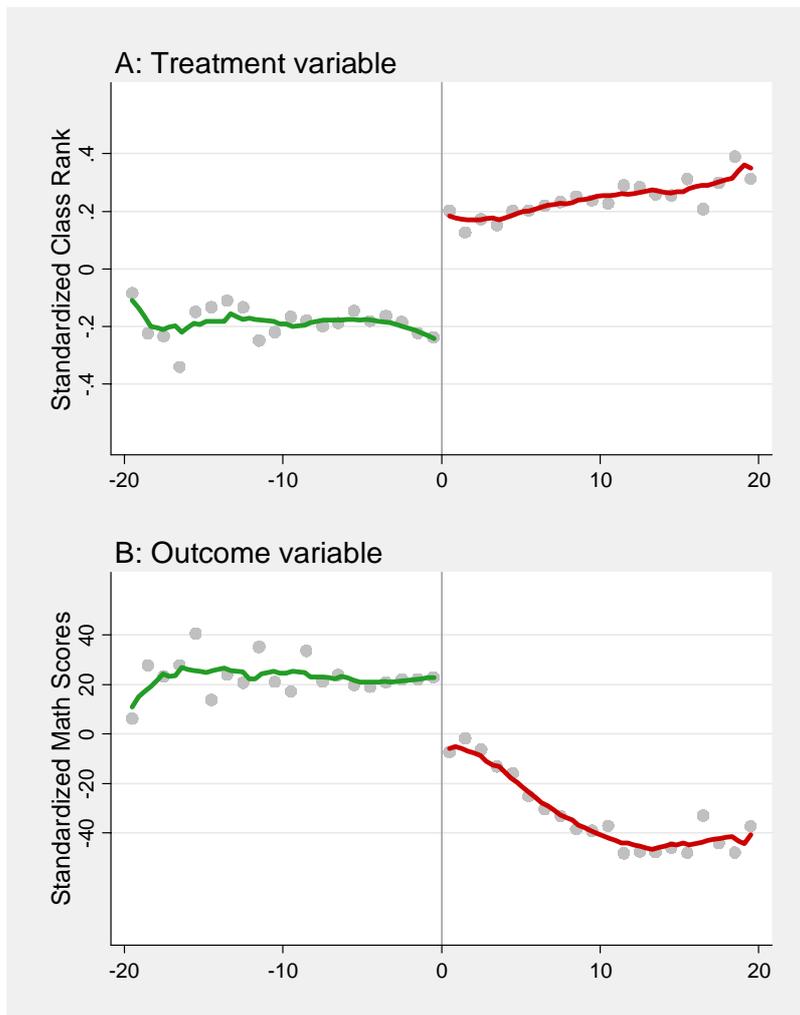
Notes: Entries are separate IV estimates of the class effect on student and family characteristics, where being in the second class has been instrumented by a dummy for having an age rank larger 0. For each variable a separate regression has been estimated. Column (1) reports the effect around the discontinuity point for the individual values of the characteristics; column (2) reports the estimates for the values of the peer group characteristics for the same individuals around the cut-off point. All regressions include a linear age control and a second-degree polynomial in age. Heteroskedasticity consistent standard errors, clustered on the school level are reported in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

TABLE 7
TEACHER AND CLASS CHARACTERISTICS

Dependent variables			
Teacher characteristics	Female	-0.018	(0.017)
	Age (in years)	-0.601	(0.886)
	White	-0.012	(0.060)
	Mixed	0.000	(0.060)
	Black	-0.004	(0.036)
	East-Asian	0.012	(0.023)
	Indigenous	0.004	(0.006)
	Higher education	-0.018	(0.038)
	Postgraduate degree	-0.016	(0.059)
	Years passed since graduating	0.257	(0.715)
	Teacher salary (in Brazilian Reais)	-47.321	(38.78)
	Participation in cont. training	-0.015	(0.048)
	Experience in education (in years)	-0.480	(0.686)
	Experience at current school (in years)	-1.048	(0.714)
Experience with current grade (in years)	0.030	(0.314)	
Class characteristics	Std. deviation of age (in months)	3.423***	(0.261)
	Class size	-2.880***	(0.359)
	Nonparticipation rate	0.071***	(0.020)
Class teacher statements	Frequency of class council meetings	0.102	(0.086)
	Quality of books	0.069	(0.050)
	Insufficient financial resources	0.005	(0.057)
	Insufficient pedagogic resources	0.049	(0.070)
	Disciplinary problems with students	0.139*	(0.078)
	% of planned curriculum taught	3.775***	(0.909)
	% of students to finish primary school	-0.284***	(0.090)
% of students to finish secondary school	-0.298**	(0.123)	
Number of observations		726	

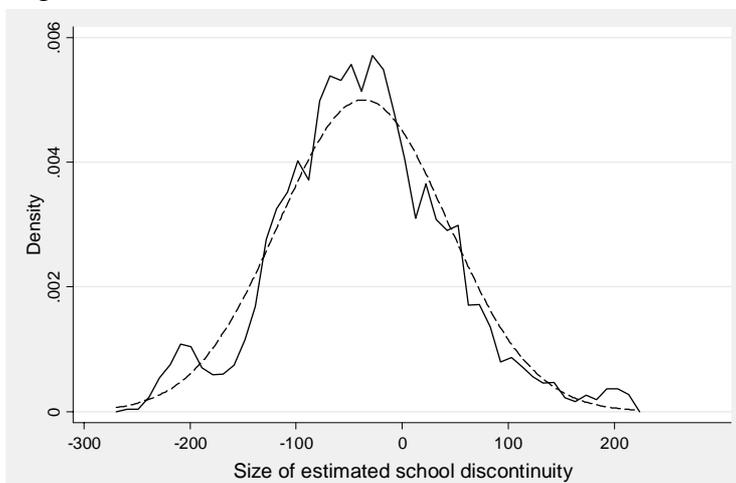
Notes: Entries are estimates of a linear probability model on a dummy of classrank, where classrank=0 for the younger class and classrank=1 for the older class. For each variable a separate regression has been estimated. The data comes from the teacher questionnaire of PROEB 2007 and the school census (for class characteristics). Class teacher statements come from the teacher questionnaire and relate to the specific class taught. Class size is computed using the official number of students enrolled in a class based on information from the school census. Non-participation rate is based on the difference between class size and number of students participating in the PROEB test. All regressions control for school fixed effects. Heteroskedasticity consistent standard errors are reported in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

Figure 1: Treatment and outcome variable discontinuity



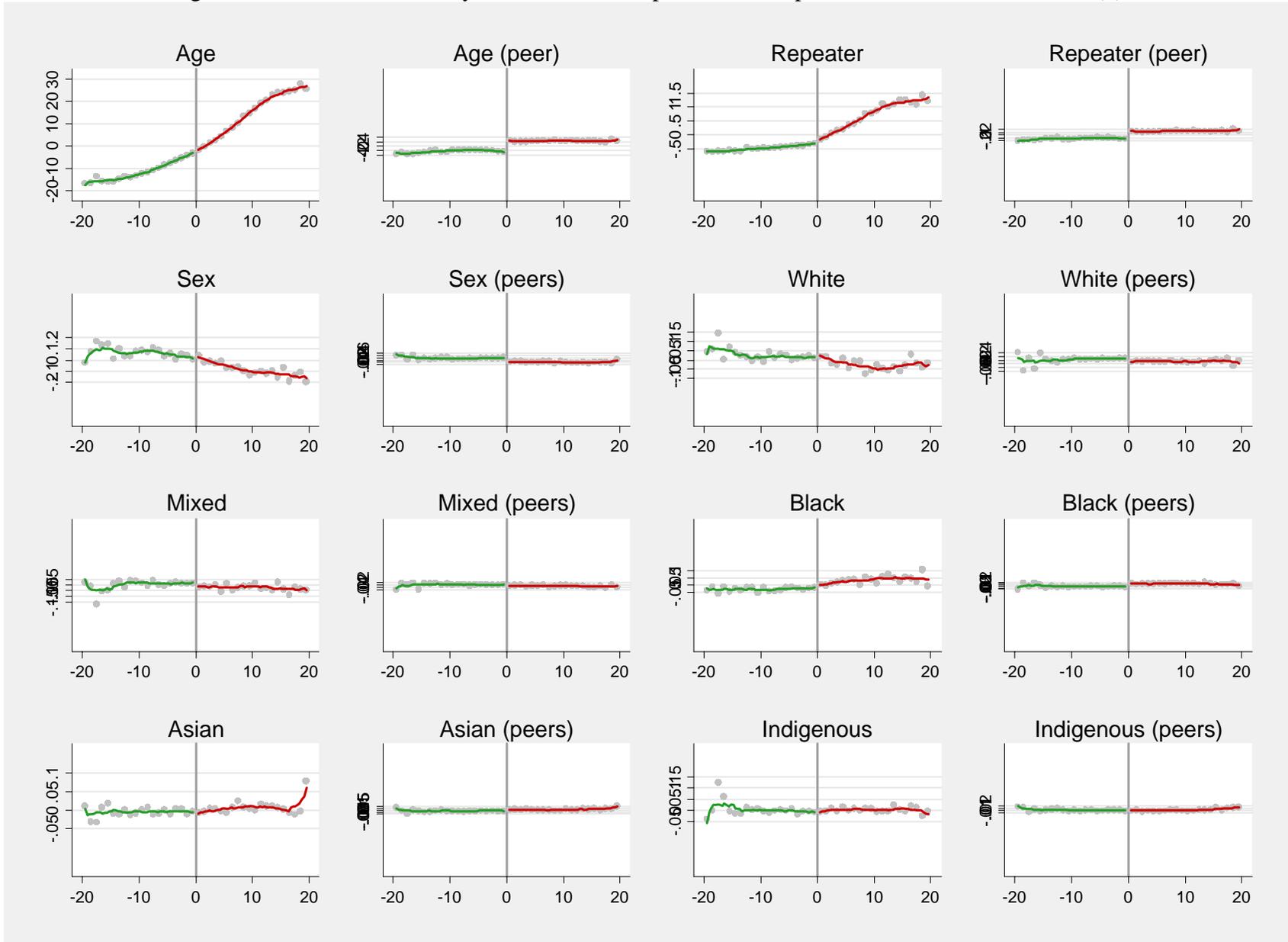
Notes: The graphs plot local averages of the standardized class rank of students and of the standardized math test score according to the age ranking in the cohort as distance of students from the cut-off point and local linear regression fits on both sides of the cut-off point using a rectangular kernel with a bandwidth of 3 months.

Figure 2: Distribution of estimated effects across schools



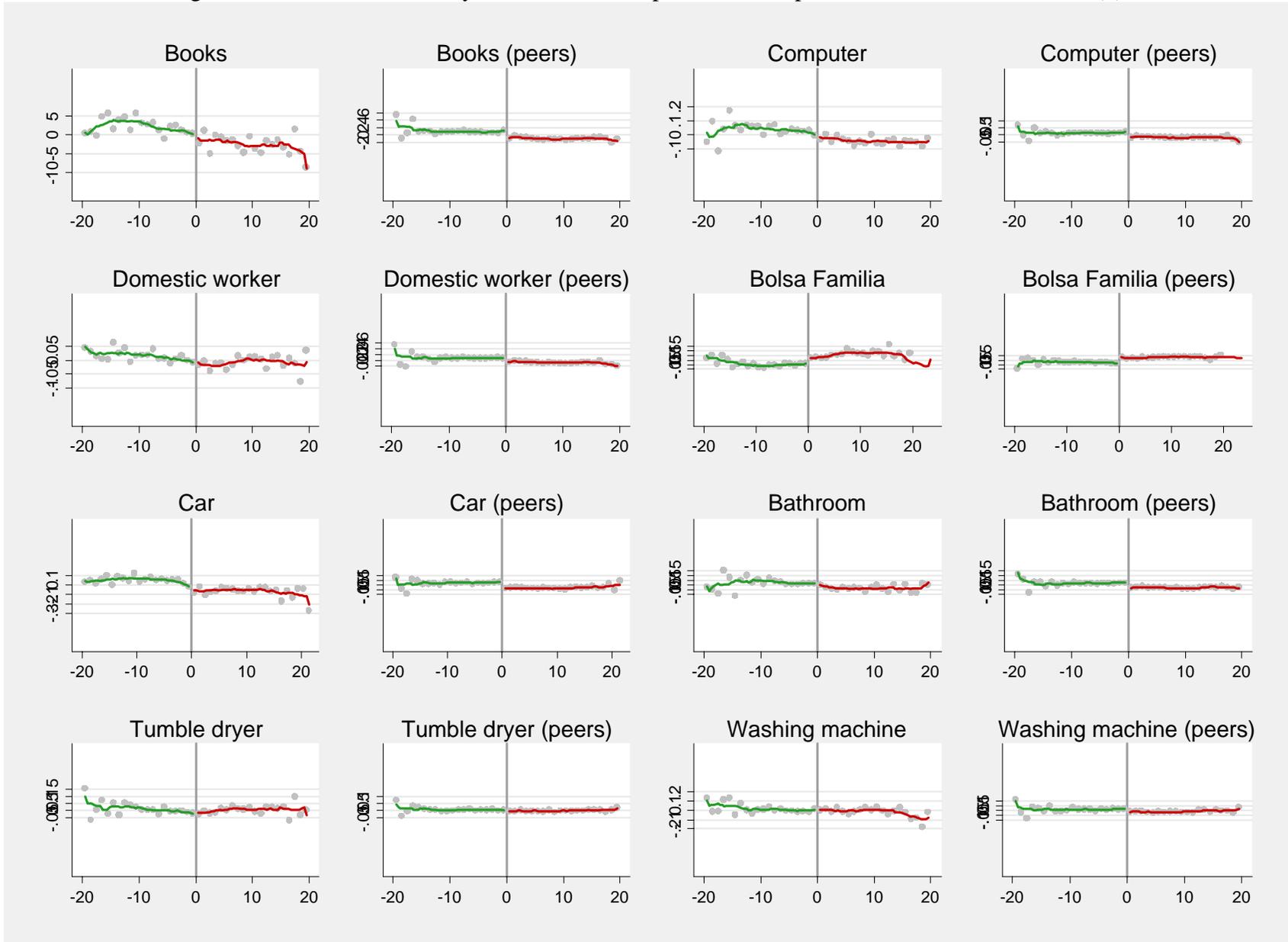
Notes: The graph plots kernel density estimates of school specific estimated discontinuities using a rectangle kernel with bandwidth 20. The dashed line is a normal distribution for comparison.

Figure 3: Test for discontinuity of individual and peer values of pre-determined characteristics (1)



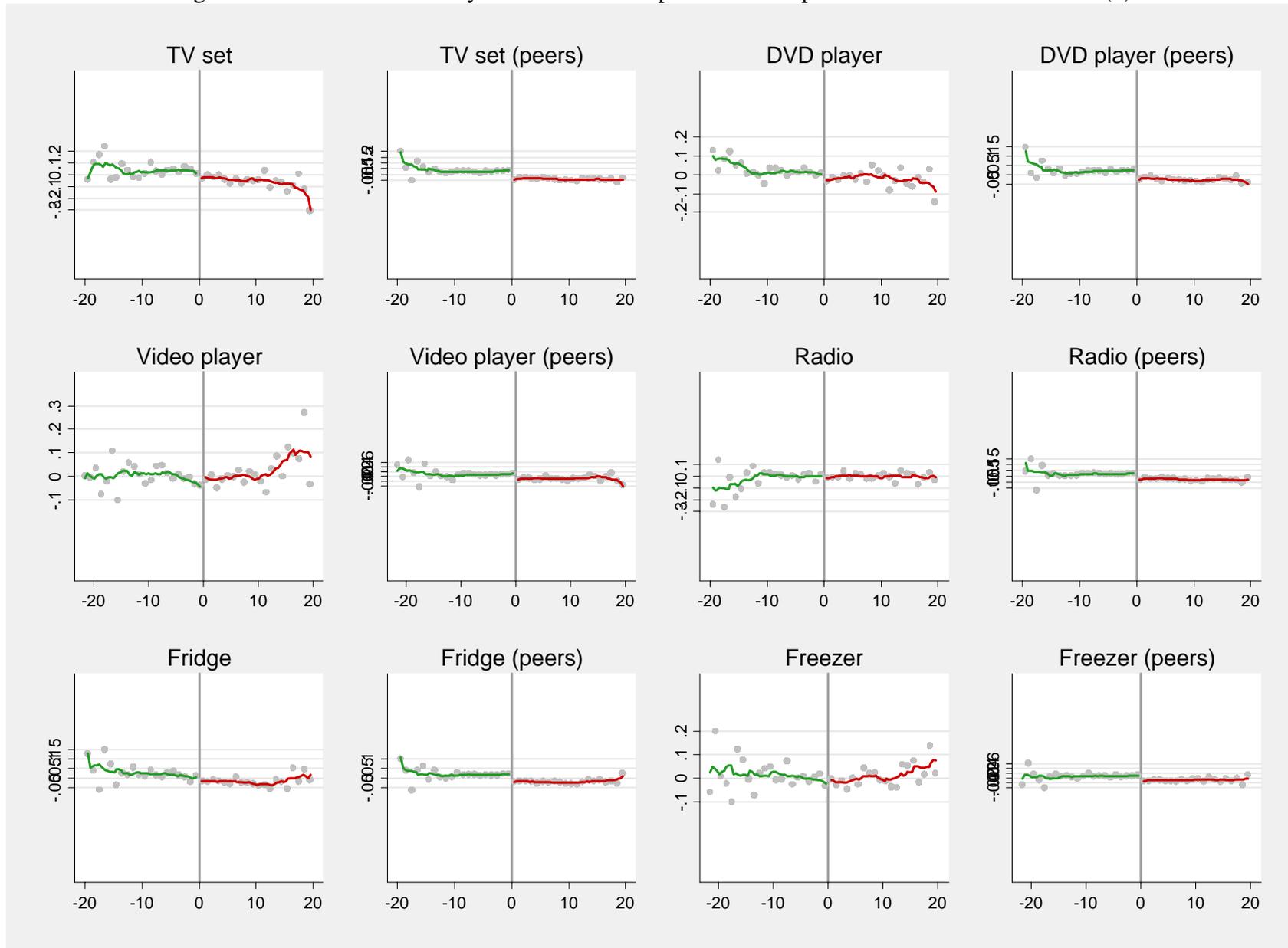
Notes: The graphs plot local averages of individual values (columns 1 & 3) and of the value for the peers of the individual students (columns 2 & 4) according to the age ranking in the cohort as distance of students from the cut-off point and local linear regression fits on both sides of the cut-off point using a rectangular kernel with a bandwidth of 3 months.

Figure 4: Test for discontinuity of individual and peer values of pre-determined characteristics (2)



Notes: The graphs plot local averages of individual values (columns 1 & 3) and of the value for the peers of the individual students (columns 2 & 4) according to the age ranking in the cohort as distance of students from the cut-off point and local linear regression fits on both sides of the cut-off point using a rectangular kernel with a bandwidth of 3 months.

Figure 5: Test for discontinuity of individual and peer values of pre-determined characteristics (3)



Notes: The graphs plot local averages of individual values (columns 1 & 3) and of the value for the peers of the individual students (columns 2 & 4) according to the age ranking in the cohort as distance of students from the cut-off point and local linear regression fits on both sides of the cut-off point using a rectangular kernel with a bandwidth of 3 months.

A1 DATA ANNEX

The variable of individual age of students has been created based on three questions related to age of the test takers in the student questionnaire of PROEB. Students need to provide their age in years, their birth month and whether or not they already passed their birthday in the current calendar year. This information together with the known test date of PROEB allows establishing the age of the children in years and months. Average age of students is 135.28 months, which is approximately 11.27 years. This is about $\frac{3}{4}$ of a year above the appropriate age at the end of 5th grade. Average age in the younger classes is 131 months and 140 months in the older class. The standard deviation of age in the cohort at 5th grade is 12.09 months. The distribution of age in the two classes differs quite considerably with a standard deviation of age in the younger classes of 10.02 month and 14.16 in the older classes. The histograms of figures A1 and figure A2 show the different distribution of age in the two classes. Both distributions are positively skewed, with the mass of the distribution concentrated to the left. This is due to age being naturally limited at the bottom with a minimum enrolment age of 5½, and the upper limit in age (maximum observed age is 15 years, which is almost 4 years above average age and 5 ½ years above the possible youngest age).

The substantial age-grade distortion in the student cohort can mostly be attributed to grade repetition of students. Every year repeated by a student contributes to the age variation based on the distribution of birth dates and the enrolment cut-off point at first grade. With 20% of students having repeated one year, 9% having repeated twice and 4% having repeated three or more times, repetition accounts almost exclusively for the age-grade distortion observed in the data (grade repetition accounts for approximately half a year in mean student age). The remainder is likely due to some late enrolment and school dropout with re-enrolment or school changes of students with reassignment at a lower grade. Unfortunately, I do not have available information on enrolment age for the cohort of interest. From the school census 2007 that contains information on age for individual students for first grade, I can infer that late enrolment was responsible for about 1.8 months, which is likely similar to late enrolment in the cohort of consideration that has enrolled 4 years earlier.

A2 ANNEX TO ORGANIZATION OF STUDENTS INTO CLASSROOMS

This annex gives some more details about the allocation mechanism of students into classes. The initial allocation of students into classes is completed before the start of the school year. With age of the students being readily observable from official documents, using student age to compose classes is an accessible way of allocating students into classes. As class composition of students is stable for at least the first five years of primary school, only migration between schools, drop-out and grade repetition affects the composition of the classes. Classes in which students have been sorted to make them heterogeneous in age, have an average age of 133.8 for the older classes and 130.0 months for the

younger classes and a t-test for the equality of the mean between the two classes reveals that there is no statistically significant difference in mean age. Schools, in which students are allocated to classes according to their relative age, have a mean age of 131.2 and 140.0 months, respectively.

To sort students to form homogeneous age classes, the school administration ranks students according to their age. With more than one class and a maximum class size rule of N students, the first class is formed by assigning students starting with the youngest student, being followed by the next youngest and so on, until the class size cap of N has been reached and the student $N+1$ is assigned to the next class etc. This rule is similar to the class-size function outlined by Angrist and Lavy in their paper on Maimonides rule (1999).⁶¹ Treating cohort size e_s as exogenous and with a maximum class-size of 25 students in Minas Gerais⁶² and under the assumption that the cohorts are divided into classes of equal size, class-size f_{sc} in school s and class c is a function of initial cohort size e_s , and f_{sc} is given as

$$f_{sc} = e_s / (\text{int}((e_s - 1)/25) + 1)$$

where the function $\text{int}(e_s - 1)$ gives the largest integer smaller or equal to $e_s - 1$ (Angrist and Lavy 1999). If cohort size is below the exact multiples of N , the maximum class-size rule does not bind and the cohort can be divided in an arbitrary way, as long as the maximum class size is not met in any of the classes. Although there is a maximum class size rule in Minas Gerais, there is some circumvention of this rule. This is obvious considering the very high cost from creating an additional class for a cohort just above multiples of the class size cap. With a given example entry cohort of 54 students strictly following the class size cap, average class size would turn out to be 18 in three classes compared to 27 for the two classes. It may often not be possible in particular for small schools with two classes per grade to accommodate an additional class in the given school space and employ additional teaching staff, so that schools with a cohort size slightly exceeding the class size caps are permitted to do so.

The possibility of grade repetition complicates the allocation of students and may lead to changes in the composition of the peer groups. Although grade repetition has been drastically reduced by the introduction of automatic grade promotion in Minas Gerais, table 1 reveals that there still exist a substantial number of students that have at least repeated one school year. This is most likely the case in third grade, for which there is an exception in the automatic grade promotion regime. Some grade repetition is also still possible in case of failure to complete a minimum number of school days during the school year. Grade repeaters will need to be accommodated in the following cohort and repeaters need to be allocated to a class. As repeaters are on average a year older than their new class mates, and in case of

⁶¹ See Urquiola and Verhoogen (2009) with a discussion on the validity of the approach by Angrist and Lavy (1999) due to strategic behaviour of schools close to the multiples of the class size cap in the case of private schools.

⁶² Law No 16.056 from 24th April 2006 establishes a maximum class-size for the first 5 years of fundamental schooling of 25 students; exceptions to that rule are only permitted in cases exceptional circumstances or transitory situations.

multiple repetition even more, they often rank on top of the age distribution in the cohort and naturally would be allocated to the older class, leading to the higher fraction of repeaters in these classes. As repeaters unlikely rank directly at the cut-off point due to being one year older, they do not pose a threat to the identification strategy. Formally testing for a discontinuity of the number of repeaters at the cut-off point confirms that there is no significant discontinuity in the number of years repeated at the threshold.

Although identifying assumptions for the regression-discontinuity are not violated by non-random choice of the allocation rule of students into classes, as identification relies on a local discontinuity within schools, it may be helpful to understand what drives the headmasters and school administrators to choose grouping students into classes at first grade in a particular manner. For that purpose, I have estimated a linear probability model, where the dependent variable is a binary variable with a value=0 if students are grouped heterogeneously in age and a value=1 if students are grouped homogeneously in age into classes. I use rich information on physical school characteristics and headmaster, teacher and mean student characteristics to learn about whether these are relevant determinants for the decision on the allocation rule. Specifically I estimate the following linear model: $Y = \beta_0 + \beta_1 S + \beta_2 D + \beta_3 T + \beta_4 P + u$, where Y is a binary outcome variable of choosing to sort students according to age making classrooms heterogeneous in age classrooms (Y=0) or homogenous in age (Y=1), S denotes school characteristics, D headmaster characteristics, T teacher characteristics, P mean characteristics of pupils in the cohort and u an idiosyncratic error term. Table A2 reports the estimated coefficients from the model; only few variables show a statistically significant effect on the grouping choice: absolute cohort size, the number of books in the parental household, the dummy on the existence of a secretariat and the participation in a computer literacy programme, the number of Pentium computers in the school and the professional experience of headmasters in years. With a larger cohort size, administrators tend to chose heterogeneous age sorting and with a student body with better socio-economic background (proxied by mean number of books in the HH) headmasters tend towards homogeneous age sorting. Also, headmasters with more years of experience tend to prefer homogeneous age sorting, but the size of the effect is negligible. Other coefficients are only marginally statistically significant at the 10% level. The selection of specific sorting schemes by the school administration nevertheless does not affect the identifying assumptions of the empirical strategy.

TABLE A1
MEANS AND PROPORTIONS SCHOOL AND HEADMASTER CHARACTERISTICS

School physical characteristics			
Means	Permanent class rooms	10.25	(0.19)
	Number of total staff	46.11	(1.15)
	Number of teaching staff	26.98	(0.98)
	Computers for students	10.11	(0.40)
	School books for 5th grade	290.13	(20.59)
	Class size	26.66	(0.01)
Proportions	Urban school	0.91	(0.02)
	State school	0.55	(0.03)
	Municipal school	0.45	(0.03)
	Filtered water	0.99	(0.01)
	Building shared with other school	0.10	(0.02)
	Headmaster office	0.90	(0.02)
	Faculty room	0.84	(0.02)
	School library	0.83	(0.02)
	Video collection	0.36	(0.01)
	TV room	0.98	(0.01)
	Video player	0.90	(0.02)
	DVD player	0.85	(0.02)
	Copy machine	0.37	(0.02)
	Kitchen	0.93	(0.01)
	Internet connectivity	0.59	(0.03)
	School canteen	0.54	(0.04)
	Computer laboratory	0.35	(0.02)
	Science laboratory	0.11	(0.02)
	Facilities for disabled children	0.82	(0.02)
	Public water supply	0.95	(0.01)
	Public energy supply	1.00	(0.00)
	Public sewage	0.83	(0.02)
	Waste collection	0.91	(0.01)
	Minimum income programme	0.98	(0.01)
TV escola (school TV programme)	0.44	(0.04)	
Public school transport	0.80	(0.03)	
School lunch	0.95	(0.01)	
Headmaster characteristics			
Sex	Female	0.86	(0.02)
Race	White	0.43	(0.03)
	Mixed	0.42	(0.03)
	Other	0.07	(0.01)
	Mean age (in years)	43.10	(0.05)
Highest edu. level	Secondary education	0.05	(0.12)
	Higher education – ped. degree	0.32	(0.03)
	Higher education – math	0.43	(0.03)

Higher education – literature	0.05	(0.01)
Higher education – other	0.15	(0.02)
Salary (in R\$)	1635.49	(38.85)
Years of experience in education	18.09	(0.21)
Years of experience at this school	6.21	(0.24)
Years of experience as headmaster	6.95	(0.26)
Participation in continued training	0.11	(0.02)

Notes: Data for the physical school characteristics comes from the annual Brazilian school census, headmaster characteristics come from the 2007 wave of PROEB.

TABLE A 2
LINEAR PROBABILITY MODEL

	coefficient	s.e.
<u>SCHOOL PHYSICAL CHARACTERISTICS</u>		
Urban school	0.00	(0.10)
State school	0.05	(0.08)
Number of permanent class rooms	0.00	(0.02)
Number of provisory class rooms	-0.09	(0.07)
Number of class rooms	0.02	(0.02)
Total number of staff	0.01	(0.01)
Total number of teachers	-0.01	(0.01)
Size of cohort	-0.01***	(0.00)
Headmaster office	-0.10	(0.10)
Secretarial office	-0.28**	(0.13)
Faculty room	0.11	(0.10)
Video collection	-0.22	(0.20)
TV room	0.00	(0.08)
Kitchen	-0.03	(0.09)
School canteen	-0.11	(0.07)
Computer lab	-0.05	(0.09)
Science lab	-0.07	(0.13)
Public energy supply	-0.19	(0.24)
Public water supply	0.09	(0.18)
Public sewage	0.04	(0.08)
Minimum Income Programme	0.21	(0.15)
TV escola	-0.08	(0.07)
Project Saude	0.10	(0.16)
Computer Literacy Programme	0.21**	(0.09)
Other federal programmes	-0.09	(0.09)
Other state programmes	0.08	(0.10)
Other municipal programmes	-0.06	(0.06)
Public school transport	-0.07	(0.08)
Number of video player	-0.02	(0.04)
Number of TV sets	0.00	(0.04)
Number of overhead projectors	-0.01	(0.06)
Number of printers	-0.02	(0.03)
Number of sound systems	0.02	(0.02)
Number of Pentium computers	0.02**	(0.01)
Number of 386/486 computers	0.02	(0.01)
<u>HEADMASTER CHARACTERISTICS</u>		
Headmaster male	-0.12*	(0.06)
Headmaster age	0.0002*	(0.00)
Headmaster highest education		
High school	-0.09	(0.11)
Higher edu - pedagogic degree	-0.05	(0.11)

Higher edu - normal	-0.13	(0.11)
Higher edu & teaching qualification	-0.09	(0.12)
Higher edu – other	-0.06	(0.14)
Experience in years as headmaster	0.0003**	(0.00)
Experience in years in education	-0.0006*	(0.00)
Continued training	0.02	(0.06)

TEACHER CHARACTERISTICS

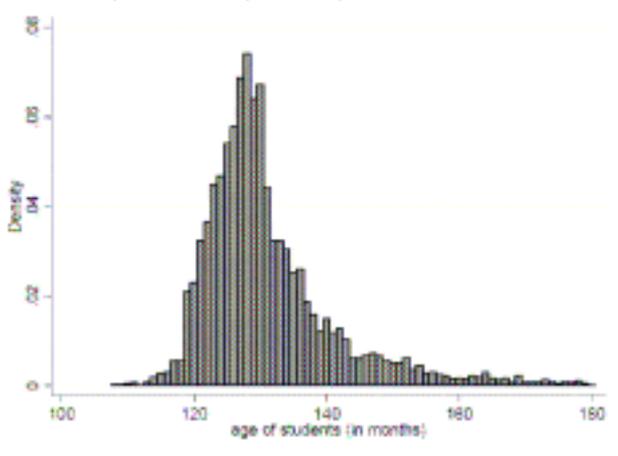
Teacher characteristics	Proportion male	0.05	(0.12)
	Higher education	0.07	(0.12)
Teacher highest edu.	High school	-0.13	(0.11)
	Higher edu. - pedagogic degree	0.13*	(0.07)
	Higher edu. - normal	-0.16	(0.12)
	Higher edu. & teaching qualification	-0.01	(0.06)
	Higher edu. – other	0.16	(0.13)
	Participation in teacher training	-0.01	(0.06)

STUDENTS CHARACTERISTICS

Student characteristics	Proportion Bolsa Família	0.04	(0.20)
	Mean books	0.01***	(0.00)
	Mean homework help	-0.01	(0.16)
	Proportion female	0.16	(0.25)
	Mean HH with domestic worker	0.43	(0.33)
	Proportion white	-0.13	(0.16)
	Mean automobiles	-0.11	(0.13)
	Mean computers	-0.14	(0.26)
	Mean times teacher absent	0.02	(0.10)
Constant		1.16***	(0.38)
Observations		363	
R-squared		0.29	

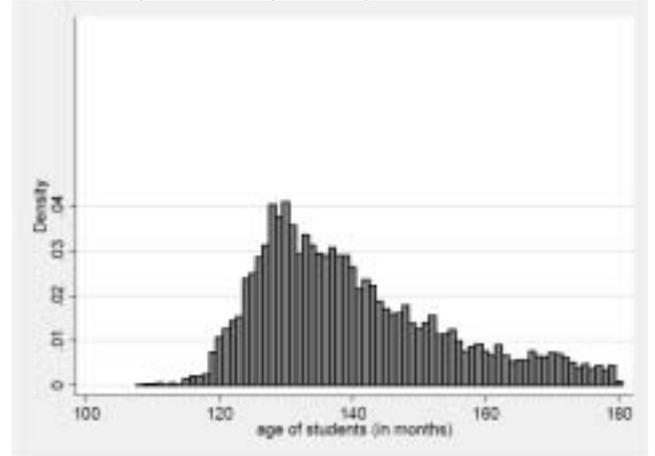
Notes: Standard errors in parenthesis, * significant at 10% level; ** significant at 5% level; *** significant at 1%

Figure A1: Age histogram for class 1



Notes: The graph plots the density of student age for class 1 (younger class), age is reported in months.

Figure A2: Age histogram for class 2



Notes: The graph plots the density of student age for class 2 (older class), age is reported in months.