CLASS ASSIGNMENT AND PEER GROUP EFFECTS: EVIDENCE FROM BRAZILIAN PRIMARY SCHOOLS

Martin Foureaux Koppensteiner*

01 October 2011

Abstract

Students in Brazil are typically assigned to classes based on their age ranking in their school grade. I exploit this rule to estimate the effects on maths achievement of being in a class with older peers for students in fifth grade of primary school. Because grade repetition is widespread in Brazil, the distribution of age is skewed to the right and hence age heterogeneity is typically higher in older classes. I provide evidence that heterogeneity in age is the driving factor behind the large negative estimated effect of being in an older class. Information on teaching practices and student behaviour sheds light on how class heterogeneity harms learning.

JEL: I20, I21
Keywords: Peer effects, regression discontinuity, educational production, group heterogeneity.

---

* I am very grateful to Marco Manacorda and Francesca Cornaglia for invaluable advice and guidance. I also thank Giacomo de Giorgi, Petra Todd, Lars Lefgren, Barbara Petrangolo, Randi Hjalmarsson, Habiba Djebbari, Maximo Torero, Michael Lechner and seminar participants at Queen Mary, Centre for Economic Performance LSE, Alicante, Leicester, Bangor, ZEW Mannheim, the North American Winter Meeting of the Econometric Society 2011, the EALE/SOLE 3rd International Conference 2010, the 13th IZA Summer School in Labor Economics 2010, the Congress of the European Economic Association 2010 and the Conference of the German Economic Association 2010 for useful comments. I am grateful to the Secretariat of Education in Minas Gerais, the Brazilian Ministry of Education and the National Institute for Educational Studies and Research (INEP) for the data. I thank Juliana Riani for assistance obtaining the data. The usual disclaimer applies.
Address: Department of Economics, University of Leicester, University Road, LE1 7RH, Leicester.
Email: mfkoppensteiner@le.ac.uk, telephone: +44 116 252 2467.

1. INTRODUCTION

The question of whether the composition of the peer group matters for the outcome of an individual member of the group has received considerable attention in numerous contexts where social interactions can be present. Peer effects have been studied in the context of schools, universities, workplaces, neighbourhoods and prisons, among others.[1] Due to the natural grouping of students into schools and classrooms, and the potential for education policies to affect the peer group composition, peer effects in education have received extensive attention by economists. Recent work goes beyond linear-in-means specifications and points to the potential relevance of the distribution of peer characteristics in explaining group effects (Hoxby and Weingarth 2006, Lyle 2009).

The identification of group effects is challenging due to conceptual problems as well as data limitations. In the education sphere, for example, an identification strategy for peer effects needs to address potential endogenous selection of students into schools and classes. With selection into groups, unobserved characteristics such as ability, parental support or students' effort are likely to be correlated among peers, and educational outcomes are therefore correlated within the peer group even in the absence of externalities. 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).

Randomized experiments are the first choice for overcoming the selection problem and there are a number of recent applications in this area (see Duflo, Dupas and Kremer (2011) on ability

[1] Recent studies include Mas and Moretti (2009) on productivity effects for supermarket cashiers; Bandiera, Barankay and Rasul (2010) on social networks and worker productivity in farm production; Bayer, Hjalmarsson and Pozen (2009) on the effect of juvenile offenders' serving time on other's subsequent criminal behaviour to name just a few. 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 (2010) 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 (2011) 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 in schools and classrooms (Zimmerman 2003 and Hanushek et. al 2003).

grouping in primary schools, and Whitmore (2005) on gender peer effects in higher education). Empirical strategies that exploit natural experiments, such as conditional random assignment of college roommates by Zimmerman (2003) and Sacerdote (2003), or the idiosyncratic variation in the gender or racial composition of a given cohort over time have also been used (Hoxby, 2000). There is little experimental or quasi-experimental evidence that overcomes the identification problems of peer group effects in primary or secondary education and even less evidence that specifically considers distributional features of peer groups that might affect educational achievement.

Group heterogeneity has not received much attention in the literature on peer effects. It has though been addressed in the literature on tracking (also referred to as streaming), where students are separated by academic ability into schools or classes.[2] Some recent research on the effects of tracking that addresses the endogeneity of tracking decisions, finds that tracking may benefit equally students from lower and higher achievement tracks. Figlio and Page (2002) show that tracking may actually help low-ability students without proposing a specific mechanism for this effect and Zimmer (2003) presents quasi-experimental evidence that a negative direct peer effect for low-achieving students is offset by the positive effects of achievement targeted instruction. Duflo, Dupas and Kremer (2011) use quasi-experimental assignment of pupils to classes to study the effect of tracking students on initial achievement among Kenyan primary school students. They find persistent positive effects across the achievement distribution of tracking students in a higher and a lower ability class. 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. 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. De Giorgi, Pellizzari and Woolston (2010) provide evidence on the effect of class heterogeneity on academic achievement and labour market outcomes in the setting of higher education. They find that the effect of the peer

[2] There is an extensive pedagogic literature on age, ability grouping, and academic tracking. 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.

distribution on student performance is non-linear and appears to be inverse U-shaped with respect to the dispersion of gender and ability in the group.

This paper provides quasi-experimental evidence from exogenous variation in peer group membership by using an assignment mechanism of students into classes which provides the basis for a regression discontinuity (RD) design. Brazilian primary school students are typically allocated to classes based on their relative age in the cohort. Using the age rank as a continuous assignment variable, this rule creates a discontinuity in the allocation to a class (peer group) for students close to the class size cap of the relatively younger class. I exploit this rule to compare outcomes of students at the margin of being assigned to an older versus a younger group in schools with two classes per cohort.

Using two-stage-least squares to estimate the discontinuity in a fuzzy RD setting, I find strong evidence for sizeable peer group effects. I estimate a negative effect from being in the relatively older class on maths test scores among students in fifth grade of around half of a standard deviation.

A major challenge in the present context is that the discontinuity cut-off - i.e. the size of the younger class - is potentially endogenous. If students are strategically allocated to classes based on their latent outcomes, variation in outcomes around the threshold is not 'as good as random' and differences in outcomes between those on the right and on the left of the cut-off do not provide a consistent estimate of the parameter of interest (Lee and Lemieux 2010). In the paper, though, I argue that assignment to the groups is largely predetermined (in 1st grade) and I find no evidence, based on a large array of observable covariates, of non-random sorting around the cut-off.[3]

Because I have data on more than 350 schools, I am able to estimate a separate parameter for each school and relate the magnitude of the estimated coefficient to differences in exogenous class characteristics across schools. This strategy allows me to identify which observable differences across classes, if any, drive the estimated gap in the attainment between barely eligible and barely

[3] Appendix A2 provides information on the initial assignment of students and the transition from one grade to the following grade.

ineligible pupils. Because, in Brazil, as in many other low- and middle-income countries, grade repetition is widespread, older classes tend typically to display larger variation in age. I find that the estimated group effect is driven mainly by the difference in the age dispersion across classes.

The remainder of this paper is organized as follows: Section 2 briefly describes the Brazilian educational system and the educational system in the state of Minas Gerais, which is the focus of this study. Section 3 describes the data. Section 4 presents the assignment mechanism of students to classes and introduces the identification strategy. Section 5 presents the main results and Section 6 presents tests for non-random sorting and for correlated effects. Section 7 gives an interpretation of the peer group estimates and Section 8 concludes.

2. THE EDUCATIONAL SYSTEM IN BRAZIL AND MINAS GERAIS STATE

Primary schooling in Brazil is compulsory and consists of nine years of schooling. Children who turn six years of age by March 31st of a given year are required to commence primary schooling in that year. 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. There exists a sizeable private sector engagement in the provision of primary schooling but, as private institutions charge substantial fees, access is limited to children from middle- and high-income families.[4] Public schools, in contrast, are free of charge at all ages.

In the public schools of Minas Gerais, which are the focus of this analysis, "normal" class size is set at 25 students per class.[5] When enrolment per grade is above 25 pupils, the school administration needs to make a choice on how to assign students to classes before the start of the school year. As, unlike innate ability or behavioural characteristics, age of students at the point of enrolment in first grade can be easily observed by school administrators, age sorting provides a

[4] Around 10% of school children in Minas Gerais attend private schools. Source: Brazilian school census 2007.
[5] Law 16.056 of 24th April 2006 limits class size to 25 students in the initial years of primary education (1st-5th grade) in all public schools in Minas Gerais. Exceptions are theoretically only allowed under special circumstances and during the transitional 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

convenient and widely used way of grouping students utilising observable characteristics at the time of entry into primary school.[6]

Students who progress regularly typically remain in their original class throughout primary school, so that, other than because of migration between schools and drop-out, assignment to classes is largely predetermined in first grade and not based on students' observable characteristics other than age.[7] Obviously, grade repetition may potentially lead to changes in the original class assignment. Although grade repetition has been reduced by the introduction of automatic grade promotion in Minas Gerais, Table 1 reveals that there still exist a substantial number of students who have repeated at least one school grade. Grade repeaters in first grade are, consistent with an assignment rule based on the age ranking of students in the cohort, usually allocated to the older class when repeating the grade in the following year. In succeeding grades, repeaters regularly are allocated to the older class as well. The propensity for repetition in subsequent grades is nevertheless also higher in the older classes, so that the in- and outflow of students into the classes largely cancel out each other and class size is hence unaffected by repetition.

3. DATA AND DESCRIPTIVE STATISTICS

For the purpose of this analysis, I use standardized test scores in mathematics of primary school students in public schools in Minas Gerais, a state in the Southeast and the second most populous state of Brazil. Educational standards in Minas Gerais are among the highest for the Brazilian states.[8] The primary source of data in this study is of PROEB (Programme of Evaluation of Basic Education), which provides maths test scores at the pupil level for all students in 5th grade in the state. I use the data for 2007, as this is the only year that contains detailed information on students' age. The test is carried out at all public schools in the state and test scores are standardized to a mean

[6] Grouping students according to their age may in fact at least partially coincide with grouping according to ability, as ability is likely to be correlated with age at time of primary school enrolment. See Cascio and Whitmore Schanzenbach (2007) and Angrist and Krueger (1991) for a discussion of student age and educational outcomes.
[7] The appendix A2 provides more information on the initial assignment of students.
[8] In the nation-wide school evaluation system of SAEB, 2005 mean maths performance of pupils from Minas Gerais is clearly above the Brazilian average, ranking first among Brazilian states (http://www.inep.gov.br/salas/download/prova_brasil/Resultados/Saeb_resultados95_05_UF.pdf).

of 500 with a standard deviation of 100. Participation is compulsory at school and individual levels, confirmed by a high student participation rate (93%). Surveyed pupils also answer a detailed socioeconomic questionnaire, which includes information on sex, month and year of birth, racial background and information on the socioeconomic background of the family.

In the following, I restrict the sample to schools with only two classes. This ensures that enough variation is available to identify sizeable group effects for students around the cut-off point, in particular with respect to variation in distributional features of the class composition.[9]

The data comprises 16,031 students from 363 public primary schools. Table 1 presents summary statistics for these data. The average age of students on the test day is 11.27 years, which is about nine months above the normal age for this grade. This age-grade mismatch is due to a combination of late enrolment and grade repetition. Students at these schools are overwhelmingly from deprived socioeconomic family backgrounds and 47% of the families of the students at these schools are recipients of *Bolsa Família*, the Brazilian conditional cash transfer programme for poor and very poor families, compared with around 25% in the total population.[10]

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 strategy at the school. The teacher questionnaire includes questions on individual characteristics, as well as statements on the students in class.

For part of the analysis, I complete the analysis with data from the 2007 School Census, which is conducted by the National Institute for the Study and Research on Education (INEP) on behalf of the Federal Ministry of Education (MEC) and comprises detailed information on school

[9] The focus on schools with two classes ensures that school administrators cannot establish *special classes* that do not follow the general assignment mechanism. With more than two classes the school administration may resort to forming separate classes in which students with specific characteristics are grouped, such as grade repeaters, and are separated from the other students in the cohort. As these *special classes* tend to be rather small, measures of age variation are also more susceptible to outliers (Lyle 2009).

[10] Families are eligible for *Bolsa Família* if per capita family income is not above R$ 120 per month (‘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.

characteristics for all primary schools in Brazil. The data appendix provides additional detailed information on the data sources and the variables used. Summary statistics from the census for the schools used in this analysis are presented in Table A1 in the appendix.

4. EMPIRICAL STRATEGY

The identification strategy used in this paper exploits the discontinuity in the assignment rule of students in schools with two classes. The treatment assignment mechanism is based 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 of receiving treatment is a discontinuous function of that variable at the class size cap $\overline{\mathrm{N}}$, the size of the youngest class.

Consider a simple reduced-form model of maths achievement

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

where Y_{is} denotes the outcome variable maths 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, I ignore at this stage any covariates one might want to include in the specification to reduce sampling variability in the estimator. Educational achievement measured in terms of test scores is assumed to depend on a smooth function $f(\cdot)$ of the student's age rank, 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 $\overline{\mathrm{N}}$ given by the maximum class size rule:

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

where ν_i is an error component.

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 and Lemieux 2010), so that assignment to either side of the discontinuity threshold is as good as random. In other terms, identification of the treatment effect relies on the assumption that just below and above the known cut-off point individuals are similar in observable and unobservable characteristics, but they are members of classes with different peer groups. In this way, the proposed RD strategy allows me to circumvent confounding effects induced by non-random sorting of individuals across groups that plagues the literature on spillover effects.

5. RESULTS

Before presenting the regression analysis, it is useful to show the raw data. The upper graph of Figure 1 plots standardized local averages of the class rank (1 or 2, to denote respectively group 1 or 2) in one month bins, where the age rank has been centred on the cut-off point of zero. 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 is around 0.5. The estimated increase in the rank is less than one, as not all schools choose to allocate students into homogenous classes. It appears that smaller schools deviate from this rule, but other than this, I find little systematic association between the probability of using the age-ranking rule and observable school and pupils characteristics.[11]

[11] In the Appendix I estimate a linear probability model, where the dependent variable is a binary variable with a value of 1 if student assignment is based on the age ranking and zero otherwise. I use the rich information in the two datasets on school, headmaster, teacher and students' characteristics to learn about the determinants of 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 takes a value of 1 for an allocation rule that sorts students into homogenous age classes and a value of 0 otherwise. 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 coefficients from the estimated model. Only few variables show a statistically significant effect at conventional levels of significance: cohort size, the existence of a headmaster's office, and the headmaster being of an Asian or Indigenous background and the mean number of fridges in student's families. With a larger cohort size, administrators are inclined to choose homogenous age sorting. The socioeconomic

In panel B of Figure 1, I plot local averages of maths test scores and the local linear regression lines on both sides of the cut-off point. The data show a very clear fall in maths test scores: the oldest pupil in the younger class shows an average attainment in maths that is 0.2 of a standard deviation higher than that of the younger pupil in the older class. Hence, Figure 1 suggests that being assigned to the older class significantly harms learning outcomes.

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. All specifications include school-fixed effects that account for observed and unobserved differences across schools which are common across classes. 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 rank. 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, similar to the observed discontinuity in panel A of Figure 1.

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

---

composition of students in the cohort and mean teacher characteristics do not seem to play a role in the choice of the assignment rule of students to classes. Other coefficients, such as the existence of a copy machine, the headmaster being of Indigenous background, the proportion of Asian students and the mean number of fridges in the student households are only statistically significant at the 10% level. In sum, there is little evidence of a systematic choice of the allocation rule based on headmaster, teacher or student characteristics.

.

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 maths test scores and significant at the 1% level.

Under the identifying assumptions 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 below $\overline{N}$ to just above $\overline{N}$.

To acquire some understanding of the distribution of effects across schools, I estimate school-specific discontinuities in maths test scores. As differences of mean peer variables between classes differ across schools, treatment also differs in respect of 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 of about -50. I will return to heterogeneous effects across schools in Section 7.

Table 3 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 a first robustness check. 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 significant way.

6. TESTS FOR NON-RANDOM SORTING

As already outlined, there are obvious threats to the identification assumption. Public knowledge of the allocation mechanism and the alleged penalty associated with treatment may invalidate the continuity assumption required for consistency of the RD estimator if 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 assignment of a student precisely around the cut-off point, the 'as good as random' assignment may fail.[12]

In the parents' case, a threat to the identification strategy arises from parents exerting pressure on school administrators to assign their child to the younger class at the time of initial enrolment or at a later stage. For the case of students close to the cut-off point, if the ability of parents to exert pressure and move their child to the younger class is systematically related to other unobserved determinants of maths achievement (e.g. the home learning environment or the support the student receives) this may invalidate the assumptions of the RD design.

Similarly, the school administration might manipulate class size so to move the youngest student in the older class to the younger class, or vice versa, based on some characteristics that are not necessarily observable to the econometrician and that are correlated with outcomes. In this case, the cut-off point would simply be shifted by one rank upwards/downwards. In reality this is unlikely to happen, as the allocation of students is decided before classes start at first grade, so that the school administration has no information on ability, race or socioeconomic background of the student other than administrative information such as age or sex that is to be found in the documents necessary for enrolment, such as a birth certificate.

In all cases, if manipulation occurred, whether due to schools or parents' pressure, pre-determined characteristics of students and their families would presumably no longer be balanced on either side of the discontinuity (van der Klaauw 2002).

In the following I use a rich array of information from the student questionnaire to formally test for the balancing properties of pre-determined student characteristics across the cut-off point.

[12] 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 uniformly distributed due to its nature as a relative rank, this test will not be informative in this analysis.

Figure A1 in the Appendix provides a 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. In Figure A1 (part 1), the graphs in columns 1 and 3 plot the individual level probability of being a girl and the probability of self-identifying with different ethnic groups. The probability of being a girl reduces smoothly with the age rank. The probability of being white, Asian or indigenous does not reveal any discontinuity at the threshold, while the probability of being a mixed or black student shows a minor positive increase at the cut-off point. The average number of months repeated also does not reveal a discontinuity, but different slopes of the local linear regression fits are apparent, these being induced by the different distribution of repeaters in the two classes. Columns 1 and 3 of Figure A1 (continued) present the same graphs for a wide range of predetermined socioeconomic characteristics. These variables appear well balanced on both sides of the cut-off point and there is no indication of a discontinuity in the means of these characteristics at the cut-off point. Among two additional proxies for the socioeconomic status of the family, the number of domestic workers employed and the fraction of families receiving *Bolsa Família*, only the latter shows a small difference around the threshold.

In a formal analysis, I estimate all predetermined characteristics of students using the same specification as for the main estimates in Table 2. Table 4 reports the RD estimates for these variables. Only the estimate for the probability of being a black student is significant at the 5% level.[13] None of the other household socioeconomic characteristics reveal a statistically significant difference at the threshold and most coefficients are small, confirming that the balancing properties of these predetermined characteristics are satisfied. 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.

[13] Choosing different specifications for the RD by including either only a linear polynomial term or a cubic term makes the estimate for this variable insignificant, 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 the significance of the estimates of any of the variables.

The inclusion of these additional individual and family controls in column 2 of Table 2 changes only modestly the estimates of the reduced-form regressions. The IV estimates at the bottom of the table are around 20% smaller than without these controls. The moderate reduction could likely be explained by model misspecification due to the inclusion of the set of predetermined variables. (Imbens and Lemieux 2008).

7. INTERPRETATION OF THE EFFECTS

A crucial question pertains to the channels through which the negative group effect operates. The substantial negative effect could either be driven by direct peer effects, e.g. through being with on average lower-performing classmates in the older class, or by indirect effects of the peer group composition that work through behavioural changes by students, teachers or schools to the class composition.

7.1 Exogenous peer characteristics and direct peer effects

In the literature, it is often assumed that peer characteristics such as sex, race and socioeconomic 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 who ask stimulating questions in class.

Column 2 of Table 4 reports the estimates of the difference in mean values of a number of peer variables for students around the cut-off point. 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. Unlike with the individual characteristics, I observe large and significant changes in peers' characteristics at the threshold. Peers 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.[14] The remainder is due to late enrolment at first grade and temporary drop-out from school followed by re-enrolment later.

Repeaters and students who enrol late at first grade often belong to families from a more deprived socioeconomic background (Patrinos and Psacharopoulos 1996 and Gomes-Neto and Hanushek 1994), which causes the socioeconomic indicators of peers to be systematically different between the two classes. RD estimates for many of these pre-determined characteristics show a statistically significant discontinuity in peer characteristics among students around the cut-off point.

Besides mean age, age dispersion in the class also differs considerably between the two classes. With the larger number of repeaters, age dispersion in the older classes is considerably greater than in the younger classes. The standard deviation of age is 40% greater (3.5 months) in the older classes (Table 5, row 1). Graphs 4 and 5 show the distribution of age of students for the two classes and give a graphical representation of the difference in the distribution of age between the classes.

Overall, students to the right of the cut-off point, while not being different from students just to the left on a whole range of individual and parental characteristics, have peer groups that not only consist of fewer girls, a higher fraction of blacks, a lower fraction of mixed students, and a higher share of children from more deprived socioeconomic background but also, due to widespread grade repetition, more heterogeneous classmates.

7.2 Indirect effects: responses of schools

Another concern for the estimation of class peer effects is, that correlated effects in the form of common shocks to the peer group (whether exogenous or endogenous) may bias the peer effect estimates. Although it is not possible to completely rule out the existence of any differences in the learning environments between the younger and older classes, I can nonetheless assess whether there

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

exist observable differences in a broad set of teacher and class characteristics, potentially in response to differences in the class composition.

Systematically different learning environments may arise from assigning teachers of different quality to either of the two classes. This may happen in a compensatory fashion, such that better teachers are allocated to weaker classes, which would lead to an underestimation of the 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. Headmasters are asked in the background questionnaire how they generally allocate teachers to classes. The vast majority (68%) of headmasters report allocating teachers in a non-systematic fashion to classes, either by means of a draw or by no specific criteria. Less than 2% of headmasters allocate more experienced teachers to stronger classes, and around 16% allocate the more experienced teachers to weaker classes. The remainder (13%) allows teachers to select the classes among themselves.[15]

To test whether there still are any systematic differences in teacher characteristics between the younger and older classes, I estimate teacher characteristics for the RD sample of students using the same specification as for the main estimates and the results are reported in Table 5. None of the teacher's characteristics, including sex, age, race, experience, education, training and earnings, reveal any significant difference between the two classes and the estimated coefficients are generally very small. This confirms that there is no evidence for strategic allocation of teachers. Including teacher characteristics as controls in the RD estimates (Table 2, column 3) also does not change the estimate for the peer effect in any relevant way.

Additional information from the teacher questionnaire about the allocation of teaching resources within the school to classes also provides some additional evidence that the estimates are not biased by common effects. Teachers report on the frequency of parent-teacher conferences, the quality of textbooks, and whether the provision of financial and pedagogic resources or of teaching

[15] Unlike in 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 so that there is less of an economic incentive to teach better classes. Details can be found in law No. 15.293 *Establishing the Careers of Professionals in Basic Education* in the state of Minas Gerais.

support staff for class teaching is insufficient. None of the variables on teacher characteristics or teaching resources in the classroom reported in Table 5 are significantly different between the two groups.

As outlined above, there is some concern about the difference in class size between the older and younger classes. The estimate in Table 5 reveals that the number of students in the older class is on average lower (by the order of four students) compared to the younger class. As class size may have an effect on student achievement, this may potentially lead to a bias in the estimation of the peer group effect. There is some agreement in the literature that smaller classes may be beneficial (see Angrist and Lavy 1999 and Urquiola 2006). The effects reported in the literature are nevertheless relatively small and mostly refer to a substantial reduction in the number of students per class. In the present case, the older class is on average smaller, so that - if anything - this may lead to a downward bias of the true peer group effect on student outcomes. As the difference in class size is rather limited, it is unlikely that it leads to any considerable bias of the estimates.

7.3 Indirect effects: responses of teachers and students

Despite the fact that teachers are observationally equivalent across classes, their teaching practices may differ as a consequence of teaching classes with a different composition of students. 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. The item responses that express levels of agreement with different statements on peer and teacher behaviour, ranging from 0 to 1, have been aggregated by averaging across all the standardized outcomes at the class level.[16] Table 6 reports the RD estimates using the aggregated variables and the specifications as for the estimates in Table 5.

[16] Because of missing values in item responses in this part of the student questionnaire, the RD estimates on individual values are less precise. As the response of teachers is not limited to the pivotal students at the threshold, but should equally affect the other students, using the aggregated data also seems sensible, despite the potential for compositional effects. Because the coefficients from the RD on individual data are very similar to the estimates in Table 6, compositional effects nonetheless do not seem to play a relevant role in this case.

The estimates reveal that students report significantly different teaching practices across classes. This is particularly remarkable as there are no differences in observable teacher characteristics.

Students in the older class report less often that their teacher is available to clarify doubts about the course content. The coefficient is -0.04 and statistically significant at the 1% level.[17] Similarly, students in the older class feel that the opportunity to express their opinion in class is substantially lower (-0.029, which is about 0.3 of a standard deviation of the mean). Further evidence of an effect on teaching practices through the impact on the distribution of instruction time is given by the difference in the answers on whether the class teacher helps some students more than others. The estimate for this variable shows a 0.084 difference between classes. Teachers in the older class are compelled to distribute their attention and instructional time more unequally, possibly devoting relatively more time to specific groups of students or addressing the same material targeted at different skills levels. With a more heterogeneous group, teachers may be less able to teach to the median students, as they need to specifically address the needs of students at the tails of the distribution. The distributional features of the class composition also result in teachers being less able to devote enough time until every student has comprehended the material (-0.027).

The higher dispersion in age and ability presumably demands that teachers address different skill levels separately. In support of this hypothesis, the proportion of the planned curriculum actually taught during the school year as reported by the teacher is about 6% lower for the older classes (Table 6).

In addition to the above findings on the differences in teaching practices, information from the student questionnaire also reveals significant differences in the behaviour of students. Students in older classes report more often that their classmates are noisy and disruptive (0.038).[18] With a more heterogeneous student composition teachers may need to spend more time on particular groups of

[17] Given the categorical nature of the answers to these questions, interpretation is not straightforward. To give an idea about the size of the effect, the point estimate is 0.49 of a standard deviation of the mean of the variable.
[18] The difference in class behaviour reported by students is confirmed by information from the teacher questionnaire. Teachers in the older classes are more likely to report disciplinary problems with their students (0.25) (Table 5).

students and more idle time for the remainder of students may also result in more disruptive behaviour.[19]

The probability of students leaving class early is also substantially higher in the older classes (0.070), which may also contribute to 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 due to noise (0.053).

The less favourable teaching environment may also have an effect on teacher motivation. Students of the older class report more often (0.041) that a teacher has been absent from school. The effect on absence of teachers may be interpreted as a response to the more deprived and demanding teaching environment. In turn, although difficult to quantify in terms of hours of instruction lost, teacher absence may also impact on the achievement of students, creating negative feedback effects between class composition, teacher and student behaviour.

Similarly to the findings of Lavy, Paserman and Schlosser (2008) the above results suggest that teaching practices respond to the peer composition and may be an important channel in explaining the negative peer effect for students close to the class threshold.

Table 5 also shows that the percentage of students who do not participate in the PROEB test, due to illness or other reasons, differs between the two classes. Although the non-response rate differs between younger and older classes for the peer group and is about 9% higher in the older classes, the non-response rate has a smooth transition across the discontinuity point. The size of the RD estimate for the non-participation rate at the threshold is very small and not statistically significant, so that the estimates cannot be confounded by differential non-response rate of students on either side of the cut-off point.[20]

[19] Interestingly, students from the entire age rank in the older class, not only marginal students close to the threshold, report a higher level of noise and disruption, which suggests that behavioural changes are not only due to the higher share of repeaters.
[20] The data appendix provides information on how the non-response rate on the class level and around the threshold has been established.

7.4 Opening the black-box of the peer-group effect: heterogeneous treatment across schools

The previous sections have discussed different potential channels through which the peer composition in this setting may lead to the drop in academic performance of students close to the cut-off point. It remains a challenging task to distinguish the precise role of the different characteristics of peers that lead to such a large disadvantage among students in older classes, whether directly or indirectly through behavioural adjustments.

The unique setup at hand with discontinuities in more than 350 schools allows – under some assumptions – the examination of the role of different observable characteristics of the peer group in explaining the gap in academic achievement. More precisely, the fact that the difference in the characteristics of peers between children in younger and older classes differs across schools can be used to gain an understanding of the role of the different channels. For students around the cut-off point, class characteristics, such as the socioeconomic composition of their peer group, are arguably quasi-random and the difference of these characteristics between classes varies across schools and can be related to the size of the test score difference across classes at the threshold.

I use a two-stage minimum-distance estimator that can be easily implemented using standard statistical packages.[21] In the first stage I estimate the size of the discontinuity in test scores at the cut-off and the differences in peer characteristics between the two classes by 2SLS separately for each school. In the second stage, the estimated discontinuities in test scores are used as dependent variable and are regressed on the estimated differences in class characteristics z_{cs}

$$b_s = \alpha_0 + \alpha_1 (z_{1s} - z_{0s}) + u_s \qquad (3)$$

where b_s are the estimated discontinuities in test scores for marginal students from the first stage.

Because the estimates of b_s are based on regressions using individual data, the minimum distance estimator is derived by minimizing the weighted difference between the auxiliary parameters from the first stage estimation, where the weights are equal to the reciprocal of the square of the standard errors of the first stage running minimum-distance weighted least squares. Given the

[21] Wolfowitz (1957) introduced the minimum-distance estimator. See Kodde et al. (1990) for details.

quasi-random allocation of students close to the threshold, the set of class characteristics are exogenous for the marginal students at each school, so that the variation in these differences across schools can be related causally to the size of the discontinuity in test scores. Under the assumption of homogenous treatment effects, by instrumenting all the variables of class characteristics by the probability of being on either side of the cut-off point, this procedure should deliver estimates of the effect of the class characteristics that are purged of the bias induced by non-random sorting of pupils across classes by a minimum-distance RD design.

Obviously, to the extent that there are other unobservable class level characteristics that affect outcomes and are correlated with the included regressors, the minimum distance estimates will still confound the effect of such variables with the effect of the included regressors. For example, if being older is also associated with lower innate ability, for example because older students have previously repeated a grade, but I am unable to measure ability, the measure of the average age of peers will also pick up the effect of having less able peers. It is consequently not possible to disentangle the effect of ability heterogeneity from the effect of age heterogeneity in this context.

Table 7 reports the coefficients of the above two-stage procedure.[22] Most of the independent variables of class characteristics are very imprecisely estimated and the direction of the effect is to some extent puzzling for some variables. The estimate of the difference in absolute age between the two classes on the test score gap is small and not statistically significant. Also, the coefficient for the difference in mean grades repeated by students in each class is small and not significant. Although a considerable part of the differences in the mean and the 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 negative effect of being in the older class in this setting and even shows a small negative effect on the absolute magnitude of the estimated discontinuity in test scores, though not being statistically significant.

[22] The dependent variable of the test score gap carries a positive sign, so that a larger positive value refers to a larger negative discontinuity in math test scores between class 1 and 2.

Other coefficients reveal a relatively unsystematic pattern: some of the differences in class characteristics are positively related to the size of the discontinuity, such as the fraction of male students, the fraction of black students, or the mean number of computers available at the homes of students, while other variables show a negative relationship, such as the fraction of white and mixed students or mean books in the students' households. None of these estimates is nevertheless statistically significant at conventional levels of significance. The coefficients for the mean number of washing machines and freezers are marginally significant at the 10% level of significance. Neither the sex, racial nor socioeconomic composition of students is relevant for explaining the differences in estimated peer effects across schools. Despite the pronounced differences in various socioeconomic peer characteristics, these do not seem to play a significant role in this context. The single significant variable for explaining heterogeneity in the size of the discontinuities across schools is the difference in the age dispersion between classes. A one month difference in the standard deviation of age explains 0.035 of a standard deviation in maths test scores. This is evidence for the importance of the age dispersion of the peers for academic achievement, in line with the findings of Hoxby and Weingarth's (2006) focus model of peer effects. The evidence from the minimum-distance procedure, which exploits the fact that differences in class characteristics across schools are quasi-random for students at the class threshold suggests, that the difference in the variation of age plays the key role in explaining the estimated strong negative effect for marginal students of being in the older class.

8. CONCLUSIONS

In this paper I introduce a novel way of identifying class peer effects using an RD design that exploits the rule which assigns students of a cohort to classes according to their ranking along the age distribution. The RD design allows us to compare students who are very similar in age but occur to be assigned to classes with either younger or older students. By exploiting this rule I provide evidence for strong negative effects on maths achievement for marginal students of being in a class with older peers. I find that marginal students who are assigned to the older classes have maths test scores that are around half of a standard deviation lower than those of students assigned to the younger classes.

Concurrently with being in different peer environments, marginal students are also either the oldest or the youngest in their respective classes and, apart from the effect from being assigned to classes with different peer characteristics and their distribution, there could be a separate pure relative age effect at work. It is nevertheless debatable whether conceptually there is a difference between a potential pure relative age effect and an age peer group effect and, given the identification strategy, these effects would by definition be practically indistinguishable. Moreover, there is no evidence for the existence of a separate pure relative age effect elsewhere in the literature.[23]

A potential threat to the identification strategy is that marginal students may be strategically allocated to classes. Assignment to classes is nevertheless largely predetermined in 1st grade and the composition of classes remains largely stable until 5th grade. When examining the balancing properties of a large array of observable characteristics, I find no evidence of non-random sorting around the cut-off. Furthermore, by examining the balancing properties of teacher characteristics and the provision of teaching resources across classes there is no evidence for behavioural adjustments by the school administration to the class composition, except for relatively small, significantly estimated adjustments in class size that may lead, if anything, to underestimating the true effect.

The unique setup of the RD design in this paper allows me to open the black-box of the estimated peer group effect. I use the fact that the estimated differences in peer characteristics between children in the younger and older classes vary across schools to get an understanding of the role of the potential channels at work. Using a two-step procedure I find that the negative effect is driven by the difference in the dispersion of age between the older and younger classes. The results

[23] Using experimental data from Project STAR, Cascio and Whitmore Schanzenbach (2007) do not find evidence for an effect from relative age on mean test scores. Elder and Lubotsky (2009) also show that a commonly postulated positive relationship between achievement and school entry age is primarily driven by the skills older children acquired prior to kindergarten rather than absolute or relative age effects. As the identification strategy employed in this paper is based on the discontinuity around the median age in the cohort, the estimated effects are not confounded by relative age effects at the extremes of the age distribution, i.e. being the youngest or oldest in the cohort, so that targeting the curriculum to a specific age group will not bias the estimated effects.

point out the importance of distributional features of the group composition for explaining peer effects in education.

Examining a set of complementary outcomes using information from the teacher and student questionnaires, I show that the group composition and presumably the age dispersion (and with it probably the dispersion in ability) is associated with differences in teaching practices. There is also evidence for behavioural differences across the two classes, with students in the older classes being more disruptive than their counterparts in the younger classes.

The results contribute to the debate on streaming or tracking of students into classes and schools and have direct policy implications. If the distribution of underlying observable student characteristics in classes has substantial effects on achievement, while changes in the composition could possibly be achieved at zero cost at the school level, there is a strong case for sorting students into classes aiming at a more homogenous class composition in terms of age or ability. Findings in the related literature point to a potential trade-off between direct and indirect peer effects from grouping students by age or ability. Zimmer (2003) finds that tracking in the US has a positive effect even on low-achieving students through more tailored instruction and can outweigh the negative direct effect on low-achievers from the absence in high quality peers. When grouping students according to ability, low-achieving students may no longer benefit from the presence of high-achieving peers, but instead may take advantage of the lower variation in ability, potentially leading to a more efficient teaching environment for all students. It is particularly important to consider this trade-off in educational systems with substantial age and ability heterogeneity, as is the case in many low- and middle-income countries.

The findings in this paper also contribute to the understanding of policies that aim at reducing the age variation in cohorts of students. Policies designed to reduce late enrolment in primary schools may have positive effects on all students by reducing the age heterogeneity in each cohort. Correspondingly, grade retention policies may have a significant impact on the dispersion of age in cohorts of students and may therefore affect achievement of all students. Retained students increase the dispersion of age in the cohort and may impose a negative externality on all students in the class regardless of the existence of direct and indirect effects of having (low-achieving) repeaters in the peer group.

The results may also contribute to the understanding of the organization of public schooling and the discussion on the consolidation of school districts (Andrews, Duncombe and Yinger 2002). Schools with larger student cohorts, given a maximum class size rule, may be able to sort students into more homogenous classes leading to efficiency gains in the provision of schooling. Likewise, there may be positive effects on student performance through reduced class heterogeneity from reducing maximum class size and hence increasing the number of classes per cohort for a given cohort size, because it may enable schools to more successfully sort students into homogenous classes (Lazear 2001).

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.

Andrews, Matthew, William Duncombe, and John Yinger. 2002. “Revisiting Economies of Size in American Education: Are We Any Closer to a Consensus?” *Economics of Education Review*, 21(3): 245-262.

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.

Carrell, 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.

De Giorgi, Giacomo, Michele Pellizzari, and William Gui Woolston. 2010. “Class Size and Class Heterogeneity.” *Journal of the European Economic Association*, forthcoming.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review,* 101(5): 1739–74.

Elder, Todd, and Darren Lubotsky. 2009. "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers," *Journal of Human Resources,* 44(3): 641-683.

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.

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.

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.

Kodde, David A., Franz, C. Palm, and Gerard A. Pfann. 1990. "Asymptotic Least-Squares Estimation Efficiency Considerations and Applications." *Journal of Applied Econometrics,* 5: 229-243.

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. 2011. "Mechanisms and Impacts of Gender Peer Effects at School." *American Economic Journal: Applied Economics.* 3(2): 1-33.

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. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(2): 281-355.

Lindert, Kathy, Anja Linder, Jason Hobbs, and Bénédict 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.

Lyle, David S. 2009. "The Effects of Peer Group Heterogeneity on the Production of Human Capital at West Point." *American Economic Journal: Applied Economics*, 1(4): 69-84.

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.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*, 142(2): 698-714.

Ministry of Education. 2004. "*Ensino Fundamental de Nove Anos – Orientações Gerais*." Secretariat of Basic Education, Federal Brazilian Ministry of Education. Brasília.

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.

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.

Whitmore, Diana. 2005. “Resource and Peer Impacts on Girls’ Academic Achievement: Evidence from a Randomized Experiment.” *American Economic Review*, 95(2): 199–203.

Wolfowitz, Jacob. 1957. “The Minimum Distance Method.” *The Annals of Mathematical Statistics*, 28(1): 75-88.

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

Panel A: Student characteristics		Younger class		Older class	
	Class size	24.738	(0.287)	21.868	(0.302)
Age	(in years)	10.930	(0.009)	11.670	(0.014)
Sex	Female	0.524	(0.005)	0.458	(0.006)
Race	White	0.306	(0.005)	0.264	(0.005)
	Mixed	0.526	(0.005)	0.517	(0.006)
	Black	0.097	(0.003)	0.143	(0.004)
	East-Asian	0.027	(0.002)	0.034	(0.002)
	Indigenous	0.044	(0.002)	0.042	(0.002)
Repeater	Never repeated	0.797	(0.004)	0.489	(0.006)
	Repeated once	0.142	(0.004)	0.292	(0.005)
	Repeated twice	0.043	(0.002)	0.148	(0.004)
	Repeated 3 or more times	0.018	(0.001)	0.070	(0.003)
SES	Family with Bolsa Família	0.480	(0.005)	0.592	(0.006)
	Household employs domestic worker	0.137	(0.004)	0.113	(0.004)
	Number of books	23.496	(0.322)	19.428	(0.330)
	Number of cars	0.608	(0.009)	0.503	(0.009)
	Number of computers	0.262	(0.005)	0.195	(0.005)
	Number of fridges	0.999	(0.005)	0.958	(0.006)
	Number of freezers	0.302	(0.006)	0.282	(0.007)
	Number of radios	1.342	(0.008)	1.286	(0.009)
	Number of TVs	1.497	(0.008)	1.396	(0.009)
	Number of DVD players	0.849	(0.007)	0.786	(0.008)
	Number of bathrooms	1.246	(0.006)	1.175	(0.006)
	Number of washing machines	0.758	(0.007)	0.752	(0.007)
	Number of tumble dryers	0.168	(0.005)	0.163	(0.005)
Panel B: Teacher characteristics					
Sex	Female	0.983	(0.011)	0.965	(0.015)
Age	(in years)	40.495	(0.468)	40.094	(0.486)
Race	White	0.456	(0.030)	0.477	(0.030)
	Mixed	0.420	(0.029)	0.399	(0.029)
	Black	0.093	(0.017)	0.081	(0.016)
	East-Asian	0.028	(0.010)	0.039	(0.012)
	Indigenous	0.004	(0.004)	0.004	(0.004)
Highest	Secondary education	0.100	(0.018)	0.118	(0.019)
educational degree	Higher education – pedagogic degree	0.210	(0.024)	0.208	(0.024)
	Higher education - regular	0.410	(0.029)	0.389	(0.029)
	Higher education and teaching qualification	0.203	(0.024)	0.174	(0.022)
	Higher education – other	0.076	(0.016)	0.111	(0.019)
	Earnings (in R$)	771.74	(22.803)	743.60	(23.754)
	Years of experience in education	14.023	(0.360)	13.862	(0.375)
	Participation in continuing education	0.375	(0.028)	0.363	(0.029)

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

TABLE 2
MAIN ESTIMATION RESULTS

	(1)	(2)	(3)
	Panel A: first stage Dependent variable: class rank		
	0.467***	0.453***	0.451***
	(0.056)	(0.057)	(0.056)
R^2	0.326	0.370	0.403
	Panel B: reduced form Dependent variable: maths test scores		
	-26.445***	-19.196**	-19.513**
	(7.458)	(7.646)	(7.743)
R^2	0.405	0.482	0.485
	Panel C: IV regression discontinuity results Dependent variable: maths test scores		
	-56.574***	-42.385***	-43.297***
	(15.299)	(15.455)	(15.673)
R^2	0.410	0.485	0.489
Number of student observations	1,688	1,688	1,688
School fixed effects	yes	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 maths test score on the dummy equal 1 for the age rank larger then 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 maths test scores, where being in the older class has been instrumented by a dummy for having an age rank larger than 0. All specifications include a second-order polynomial in the age rank. Specifications in column (2) include the whole set of predetermined individual and family characteristics, including sex, race, repeated years 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. ** and *** denote significance at the 5% and 1% level, respectively.

TABLE 3
RD ESTIMATES OF MATHS TEST SCORES

	Ranks from threshold in months				
	1 month	2 months	3 months	4 months	5 months
	Estimated discontinuity at threshold				
Quadratic	-56.574***	-54.578***	-59.044***	-57.193***	-59.182***
	(15.299)	(12.561)	(11.103)	(10.791)	(10.653)
Cubic	-55.477***	-54.467***	-59.560***	-57.188***	-58.416***
	(15.551)	(12.622)	(11.106)	(10.842)	(10.722)
Quadratic with full controls	-43.297***	-43.762***	-45.216***	-43.600***	-43.066***
	(15.673)	(12.446)	(11.259)	(10.980)	(10.675)
Cubic with full controls	-41.689**	-43.753***	-45.625***	-43.769***	-42.726***
	(16.299)	(12.45)	(11.274)	(11.031)	(10.749)
Number of student observations	1,688	3,142	4,547	5,884	7,223

Notes: The dependent variable is the maths test score and entries are estimates of the discontinuity including the 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 the age rank for the different range of observations. Row (2) includes a cubic polynomial in the age rank. 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. ** and *** denote significance at the 5% and 1% level, respectively.

TABLE 4
RD ESTIMATES OF PREDETERMINED INDIVIDUAL AND FAMILY VARIABLES

		(1)		(2)	
		Individuals		Peers	
	Age (in months)	0.442	(0.735)	8.157***	(0.796)
	Grades repeated (in months)	0.728	(0.879)	7.487***	(0.457)
Fraction of:	Female	0.190	(0.127)	-0.088***	(0.019)
	White	0.008	(0.092)	-0.035	(0.023)
	Mixed	-0.037	(0.102)	-0.072**	(0.032)
	Black	0.115**	(0.055)	0.089***	(0.018)
	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)
	Bolsa Família	0.165*	(0.099)	0.144***	(0.027)
	Parental homework support	0.027	(0 .054)	-0.066***	(0.016)
Number of:	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 student 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 than 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 specifications include a second-order polynomial in the age rank Heteroskedasticity consistent standard errors, clustered on the school level are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.

TABLE 5
CLASS AND TEACHER CHARACTERISTICS

Dependent variable			
Class characteristics	Std. deviation of age (in months)	4.012***	(0.381)
	Class size	-4.162***	(0.583)
	Non-participation rate (at threshold)	0.006	(0.004)
	Non-participation rate (of peers)	0.093***	(0.022)
Teacher characteristics	Female	-0.087*	(0.049)
	Age (in years)	-1.607	(1.615)
	White	-0.005	(0.101)
	Mixed	-0.048	(0.103)
	Black	0.025	(0.060)
	East-Asian	0.020	(0.033)
	Indigenous	0.009	(0.009)
	Higher education degree	0.030	(0.077)
	Postgraduate degree	-0.034	(0.103)
	Years passed since graduation	-0.108	(0.226)
	Earnings (in Brazilian Reais)	-69.176	(56.943)
	Participation in continuing education	-0.015	(0.091)
	Experience in education (in years)	-0.395	(0.259)
	Teacher has other source of income	-0.089	(0.093)
Teaching resources	Frequency of parent-teacher conferences	0.068	(0.135)
	Quality of textbooks	0.178	(0.098)
	Insufficient financial resources	-0.024	(0.080)
	Insufficient pedagogic resources	-0.063	(0.108)
	Insufficient teaching support staff	0.036	(0.102)
Number of student observations		1,688	

Notes: Entries are separate IV estimates of the class effect on class and teacher characteristics, where being in the second class has been instrumented by a dummy for having an age rank larger than 0. For each variable a separate regression has been estimated. The data come 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 calculated using the official number of students enrolled in a class based on information from the School Census. The *non-participation rate (at threshold)* is based on the difference in the distribution of students of age ranks between the school census and PROEB test takers. The *non-participation rate of peers* is based on the difference between class size and number of students participating in the PROEB test. The variable *quality of textbooks* ranges between 0 and 1, with the value 1 given for the best quality and 0 for the lowest. All regressions control for school fixed effects. Heteroskedasticity consistent standard errors are reported in parentheses. * and *** denote significance at the 10% and 1% level, respectively.

TABLE 6

RESPONSE OF TEACHING PRACTICES TO CLASS COMPOSITION

Disciplinary problems with students	0.253**	(0.113)
Rate of planned curriculum taught	-0.059***	(0.019)
Rate of students expected to finish primary school	-0.097***	(0.023)
Rate of students expected to finish secondary school	-0.133***	(0.031)
Teacher availability to clarify doubts	-0.039***	(0.008)
Teacher explains until all students understand	-0.027***	(0.009)
Teacher gives opportunity to express oneself	-0.029***	(0.010)
Teacher helps more some students	0.084***	(0.015)
Teacher interested in learning progress	-0.028***	(0.007)
Teacher needs to wait to start teaching	0.053***	(0.017)
Teacher absenteeism	0.041***	(0.012)
Fellow students leave classroom early	0.070***	(0.015)
Fellow students are noisy and disruptive	0.038***	(0.015)
Fellow students learn taught material	-0.044***	(0.010)
Fellow students pay attention in class	-0.009	(0.010)
Teacher enforces student attention	-0.010	(0.007)
Teacher corrects homework	-0.020	(0.013)
Number of student observations	1,688	

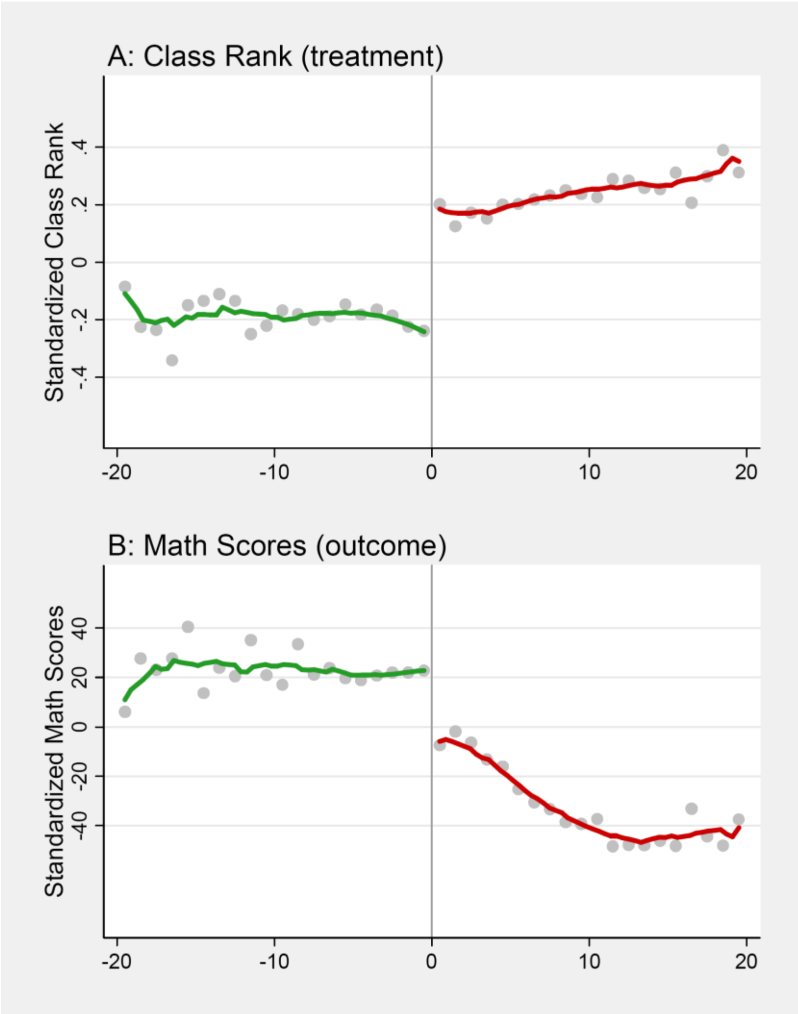
Notes: Entries are separate IV estimates of the class effect on the response of teachers and students to the class composition, where being in the older class has been instrumented by a dummy for having an age rank larger than 0. For each variable a separate regression has been estimated. The variables in the top panel are from the teacher questionnaire. The variable *disciplinary problems with students* is a dummy taking a value 1 if teachers report that there are problems with the discipline of students. The variables from the bottom two panels come from the student questionnaire of PROEB 2007. The variables have been recoded from categories ranging from "totally disagree" to "totally agree" on a scale from 0-1 and aggregated on the class level. All regressions control for school fixed effects. Heteroskedasticity consistent standard errors, clustered on the school level, are reported in parentheses. ** and *** denote significance at the 5% and 1% level, respectively.

TABLE 7
HETEROGENEOUS TREATMENT ACROSS SCHOOLS

Difference in class means		
Age dispersion	3.485**	(1.446)
Mean age (in months)	-0.704	(1.625)
Mean grades repeated (in months)	-2.149	(23.899)
Fraction of male students	26.985	(26.070)
Fraction of white students	-21.781	(35.988)
Fraction of mixed students	-35.811	(26.031)
Fraction of black students	23.887	(50.446)
Fraction of Asian students	-81.925	(106.722)
Fraction of households with domestic workers	-1.419	(49.267)
Fraction of households receiving Bolsa Família	-29.259	(35.020)
Mean books	-30.246	(18.638)
Mean bathrooms	-2.999	(32.698)
Mean cars	0.302	(25.877)
Mean computers	4.398	(43.518)
Mean fridges	-16.989	(28.483)
Mean freezers	-58.830*	(33.266)
Mean radios	31.154	(24.689)
Mean washing machines	41.407*	(22.327)
Mean DVD players	42.635	(32.030)
Mean TV sets	-3.199	(24.698)
Mean video players	34.795	(33.711)
Teacher controls	yes	
Number of observations (discontinuities):	363	
R^2	0.302	

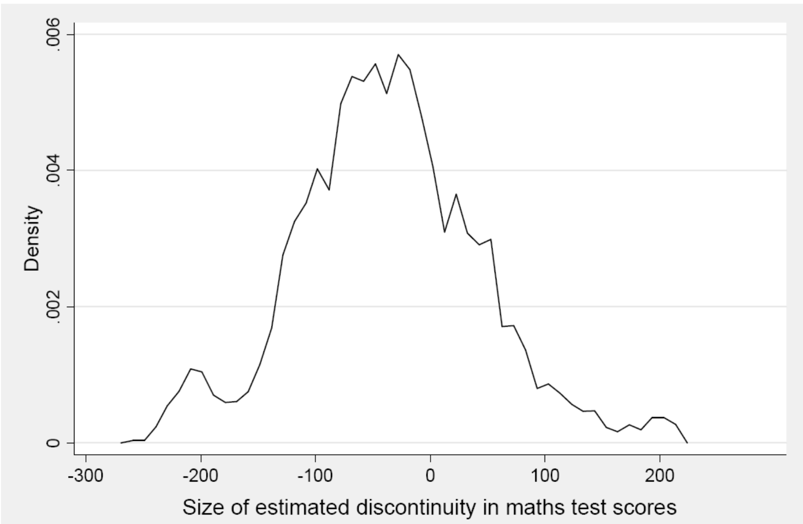
Notes: The dependent variables are measures of the absolute size of the discontinuities in math test scores at the cut-off point on the school level estimated by 2SLS. The entries report coefficients from the second stage of the minimum distance estimation, where the weights are equal to the inverse of the standard errors of the estimates of the first stage. Independent variables are the estimated differences in means of the peer values of socioeconomic characteristics, class age and its distribution. Heteroskedasticity robust standard errors are reported in parenthesis. * and ** denote significance at the 10% and 5 level, respectively.

FIGURE 1: LOCAL AVERAGES AND LOCAL LINEAR REGRESSION OF TREATMENT AND OUTCOME VARIABLE



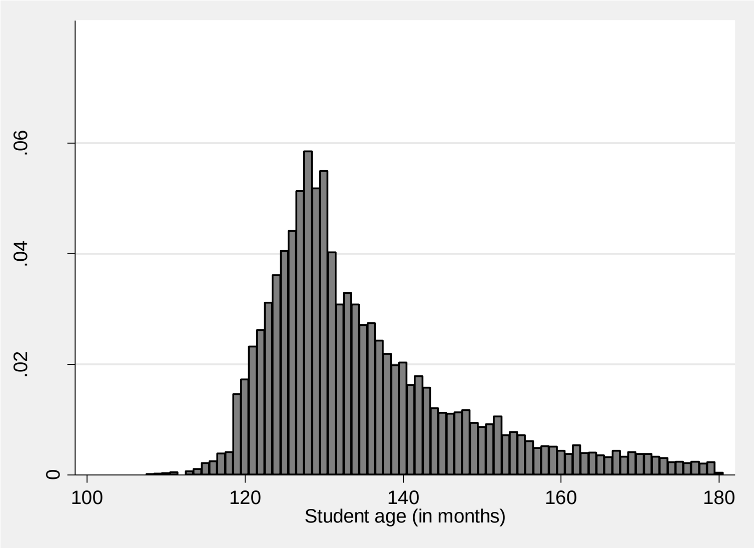
Notes: The graphs plot local averages of the standardized class rank of students and of the standardized maths test score according to the age ranking in the cohort as distance of students from the cut-off point and local linear regression fits on both sides of the cut-off point using a rectangular kernel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

FIGURE 2: DISTRIBUTION OF RD ESTIMATES ACROSS SCHOOLS



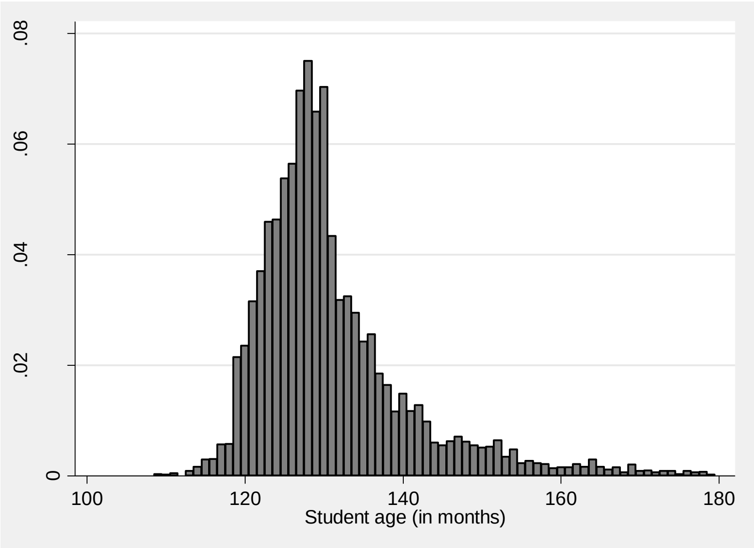
Notes: The graph plots kernel density estimates of school specific estimated discontinuities using a rectangle kernel with a bandwidth of 20.

FIGURE 3: AGE DISTRIBUTION IN THE COHORT



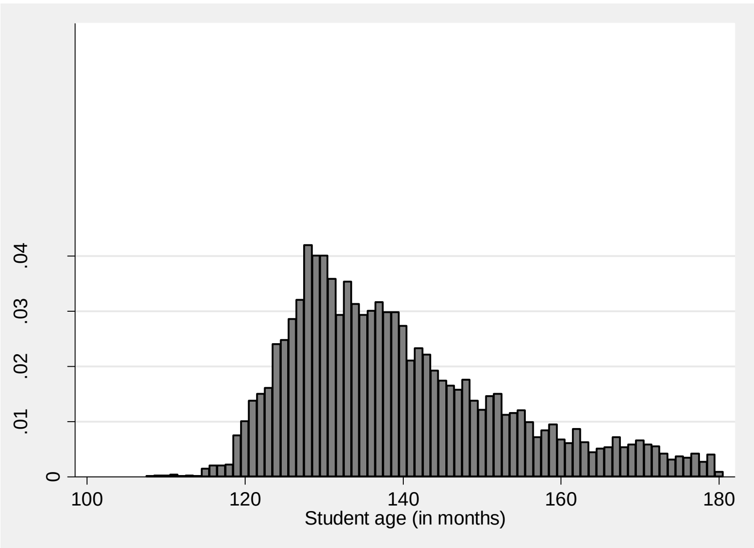
Notes: The graph plots the density of student age for all students in the cohort, age is reported in months.

FIGURE 4: AGE DISTRIBUTION IN YOUNGER CLASSES



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

FIGURE 5: AGE DISTRIBUTION IN OLDER CLASSES



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

A1. DATA APPENDIX

This appendix describes the variables of students, teachers, class, headmaster and schools used in this paper.

Outcome variable: Maths test score

The PROEB test score for mathematics has been constructed from a battery of 40 multiple choice questions covering four areas: space and shapes, size and measurement, algebraic operations, and treatment of information. For each question, students are offered 4 possible answers, of which one is correct. The test scores have been standardized to a mean of 500 and a standard deviation of 100. The test is administered in November, close to the end of the school year.

Student socioeconomic characteristics

All information on the socioeconomic background of students comes from a socioeconomic questionnaire which is a supplement to the maths test. *Racial affiliation* is self-reported by students, as well as all other information on the background characteristics of the students and their families.

The dummy variable *Bolsa Família* reports whether the family is a recipient of cash-transfers from the federal programme and takes a value of 1 if the family is a recipient.

The dummy variable *domestic worker* records whether the family employs one or more domestic workers (part-/full-time).

The variables on the *number of books, cars, computers, fridges, freezers, radios, TVs, DVD players, bathrooms, washing machines and tumble dryers* are numeric and can take the values “0”, “1”, “2” or “3 and more”. The value of “3 and more” has been coded with a value of 3.

The variable *individual age* of students has been created based on three questions related to age. Students need to provide their age in years, their month of birth and indicate whether or not they have already passed their birthday in the current calendar year. This information, together with the test date of PROEB, allows the age of the children in years and months to be established. The average age of students is 135.28 months, which is approximately 11.27 years. This is about 9 months above the appropriate age at the end of 5th grade. Average age in the younger classes is 131 months and in the older class 140 months. 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 months and of14.16 months in the older classes. Figure 3 plots the density of age in the entire cohort and shows how the age distribution is skewed to the right. The histograms of Figures 4 and 5 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 because age is censored at the left tail with a minimum enrolment age of 5½, and the upper age limit. The maximum observed age is 15 years, which is almost 4 years above the average age and 5½ years above the possible youngest age. The substantial age-grade distortion in the student cohort can mostly be attributed to grade repetition by 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 1st grade. With 20% of students having repeated one year, 9% having repeated twice and 4% having repeated three or more times, repetition accounts almost wholly 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 to be due to some late enrolment and school dropout with re-enrolment or change of school by students who are then reassigned to 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 1st grade, I can calculate that late enrolment is responsible for about 1.8 months, which is likely to be similar to the effect of late enrolment in the cohort of consideration that had enrolled 4 years earlier.

Teacher characteristics, statements of teachers on class teaching environment and class characteristics

The information on teacher characteristics comes from two sources and these are matched by school and class identifiers and the subject the teachers teach. All information on socioeconomic characteristics (in panel B of Table 1) comes from the annual Brazilian School Census that collects information on school, teacher and headmaster characteristics from all Brazilian schools. The variables *years passed since graduation* and the different variables on teacher professional experience have been transformed using midpoints of the ranges reported in the questionnaire. Salary of teachers is reported in Brazilian Reais (R$1 was worth approximately US$0.58, as of 10th September 2010) and is calculated from the mid-points of the salary ranges given in the questionnaire.

The information on the teaching environment and student behaviour comes from the background questionnaire of PROEB that is completed by all teachers. *Frequency of class council meetings* is reported as *never, once, twice and three times and more*, the last of which has been recoded with a value of 3, and reports on the adequacy of financial and pedagogic resources for class teaching are dummy variables taking a value of 1 if teachers think that resources are insufficient and 0 otherwise. Teacher statements about the progress of teaching and students have been included from the Prova Brasil 2007 teacher background questionnaire. The percentages of the *planned curriculum taught*, and students finishing primary and secondary school, have been calculated using the midpoint of the percentage ranges reported by teachers.

The variable on standard deviation of age in the classes is calculated using individual student age. The variable class size is based on information from the official School Census that reports the number of students in each class.

The *Non-participation rate* at the class level is based on the difference between official student numbers as recorded by the School Census and the number of students participating in the PROEB test at the class level. The *Non-participation rate at the threshold* is established using information from the school census on the complete age distribution of all students at school. The difference between official numbers and numbers of students taking the PROEB test for the same age rank (as in monthly intervals) at the school level informs about the missing students and the non-response rate of students for each age rank at schools.

Student evaluation of teaching practices and classroom environment

The information about teaching practices and the classroom environment in Table 6 come from the student background questionnaire of PROEB. The variables report means on the class level. The bottom four variables refer directly to classmates of students, whereas the top variables refer to teaching practices and teacher behaviour. Students report their level of agreement to statements about teaching practices and the behaviour of their classmates, from strongly disagree to strongly agree. The categories have been recoded to range from 0 (strong disagreement) to 1 (strong agreement). Table A3 reports the mean and standard deviation for these variables on the class level.

School characteristics and headmaster characteristics

The information on physical school characteristics comes from the annual School Census.

The dummy variable *urban school* takes a value of 1 for being in an urban setting, and 0 for a rural setting.

The dummy variable *state school* takes a value of 1 for the school being under the direct administration of the state secretariat of education in Minas Gerais and 0 for a municipal school which is under administration of the municipal secretariat of education.

The variables of *headmaster office, faculty room, school library, video facilities, TV room, copy machine, printer, overhead projector, school kitchen, internet access, computer and science lab, filtered water, public water supply, public sewerage and sport facilities* are all dummy variables taking a value of 1 when the facilities exist at the school and 0 otherwise.

Normalization on school and class level

As mentioned in the text, each of the regressions includes school fixed effects. For this purpose all variables used for the RD analysis have been normalized to have a mean of 0 at school level. Furthermore, the ranking of students has been centred on a cut-off point of 0, reporting the age rank as distance from the cut-off point.

A2. APPENDIX ON INITIAL CLASS ASSIGNMENT AND CLASS TRANSITION

Primary education in Brazil is divided into two stages. The first stage (*initial years*) comprises five years and the second stage (*final years*) the remaining four years of primary education. During the initial years a single class teacher (*professor regente*) teaches the entire curriculum covering all subjects (mathematics, Portuguese, science, history, geography), whereas classes are taught by specialized teachers separately for different subjects in the last four years of primary school.[24] The aim of the initial years, besides the achievement of curriculum targets, is to establish social and emotional ties and to build the capacity of students in interacting with other children of similar age and with adults.[25] To facilitate this aim, all subjects are taught by a single class teacher and students remain in their originally assigned class formed at first grade throughout the first five years of primary school. It is therefore informative to learn about the initial assignment of students into classes and the transition of students from grade to another. As PROEB only focuses on the cohorts tested (5th and 9th grade of primary school) there is nevertheless no individual data for the initial class assignment at first grade for the cohort of interest. With a change in the data collection method of the Brazilian school census in 2007, information on individual students rather than aggregated class and school data is collected from the year 2007. The school census contains information on individual characteristics on age, sex and the racial attribution of students and permits to test whether or not the balancing properties of these predetermined characteristics are satisfied for the entry cohort of 2007. Table A4 reports the RD estimates for these characteristics for the 2007 entry cohort of first graders.

[24] Details are outlined in Resolution SEE No 1086 of the State Secretariat of Education Minas Gerais.
[25] Brazilian Ministry of Education (2004).

The coefficients of the RD estimates for sex or the racial attributes of students at the threshold are relatively small and not statistically significant, confirming that the predetermined characteristics are balanced across the threshold for first grade students of the entry of cohort of primary school.

Using the 2007 and 2008 census I can follow the cohorts of students from school year 2007 to school year 2008.[26] The transition information includes the records of students being promoted from 1st to 2nd, from 2nd to 3rd, from 3rd to 4th and from 4th to 5th grade and I have pooled all the cohorts together. Over 80% of students remain in the class with the same peers. Conditional on regular transition, 96% of students remain in the same peer environment. Regressing the probability of being in class j (j=1/2) at time t-1 on the probability of being in group j at time t, conditional on age rank at age t-1 does not reveal a significant difference for remaining with the same class for students that rank close to the cut-off point, so that students close to the threshold have the same (high) probability to stay with the same class the next year compared to students further away from the threshold.

[26] For this exercise the student information is available only for a restricted sample of schools. Not all schools have recorded consistently the student information across years in the school census, so that the information on the transition of classes is only available for 55 schools.

TABLE A1

MEANS AND PROPORTIONS SCHOOL AND HEADMASTER CHARACTERISTICS

Physical school characteristics			
Means	Permanent class rooms	10.250	(0.190)
	Number of total staff	46.110	(1.150)
	Class size	23.199	(0.217)
Proportions	Urban school	0.910	(0.020)
	State school	0.553	(0.030)
	Municipal school	0.447	(0.030)
	Headmaster office	0.897	(0.016)
	Faculty room	0.844	(0.019)
	School library	0.825	(0.020)
	Video facilities	0.356	(0.010)
	TV room	0.979	(0.007)
	Video player	0.902	(0.015)
	DVD player	0.847	(0.019)
	Copy machine	0.370	(0.025)
	Printer	0.903	(0.017)
	Overhead projector	0.788	(0.023)
	School kitchen	0.926	(0.013)
	Internet connectivity	0.589	(0.028)
	Computer laboratory	0.355	(0.025)
	Science laboratory	0.106	(0.016)
	Facilities for disabled children	0.820	(0.020)
	Filtered water	0.989	(0.005)
	Public water supply	0.950	(0.011)
	Public energy supply	0.997	(0.003)
	Public sewerage	0.828	(0.019)
	Waste collection	0.913	(0.015)
	Sport facilities	0.606	(0.027)
Headmaster characteristics			
Sex	Female	0.860	(0.020)
Race	White	0.452	(0.028)
	Mixed	0.427	(0.028)
	Black	0.068	(0.014)
	Asian	0.046	(0.014)
	Indigenous	0.007	(0.005)
	Age (in years)	43.100	(0.054)
Highest educational level	Secondary education	0.050	(0.123)
	Higher education – pedagogic degree	0.318	(0.026)
	Higher education – maths	0.428	(0.028)
	Higher education – literature	0.053	(0.013)
	Higher education – other	0.151	(0.020)
	Earnings (in R$)	1635.49	(38.85)
	Years of experience in education	18.090	(0.209)
	Years of experience at this school	6.210	(0.241)
	Years of experience as headmaster	6.949	(0.258)
	Participation in continuing education	0.114	(0.020)

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
CHOICE OF CLASS ASSIGNMENT RULE

		coefficient		s.e.
SCHOOL PHYSICAL CHARACTERISTICS				
	Urban school	0.025		(0.121
	State school	0.006		(0.06
	Number of permanent class rooms	0.01		(0.01
	Total number of staff	-0.001		(0.002
	Size of cohort	-0.008	***	(0.003
	School library	0.029		(0.081
	Headmaster office	-0.224	***	(0.067
	Faculty room	0.05		(0.083
	Video facilities	0.038		(0.091
	TV room	-0.133		(0.189
	Copy machine	0.078		(0.051
	Printer	-0.079		(0.086
	Overhead projector	-0.009		(0.07
	School kitchen	0.069		(0.081
	Internet access	0.085	*	(0.05
	Computer lab	-0.111	*	(0.056
	Science lab	-0.041		(0.09
	Filtered water	-0.019		(0.101
	Public water supply	0.145		(0.252
	Public sewerage	0.022		(0.08
	Sport facilities	-0.007		(0.048
HEADMASTER CHARACTERISTICS				
	Male	-0.019		(0.060)
	Age	-0.001		(0.002)
	Mixed	0.066		(0.054)
	Black	0.052		(0.080)
	Asian	0.167	**	(0.078)
	Indigenous	0.036	**	(0.144)
Highest education obtained	High school	0.117		(0.188)
	Higher education - pedagogic degree	0.014		(0.129)
	Higher education – normal	-0.077		(0.126)
	Higher education & teaching qualification	0.033		(0.152)
	Higher education – other	0.044		(0.132)
	Experience in years as headmaster	-0.001		(0.002)
	Experience in years in education	0.000		(0.000)
	Continuing education	-0.055		(0.0760)
	Earnings	0.000		(0.000)
MEAN TEACHER CHARACTERISTICS				
	Proportion male	-0.113		(0.180)

Table A2 cont.

High school	-0.149		(0.116)
Higher education – pedagogic degree	-0.056		(0.104)
Higher education – regular	-0.081		(0.107)
Higher education and teaching qualification	-0.128		(0.153)
Higher education – other	-0.126		(0.126)
Earnings	0.000		(0.000)
Mean experience in education	0.004		(0.005)
Proportion Bolsa Família	0.147		(0.227)
Mean books	0.008		(0.006)
Proportion female	0.000		(0.275)
Mean HH with domestic worker	0.260		(0.471)
Proportion white	-3.325		(1.983)
Proportion mixed	-2.365		(1.951)
Proportion black	-2.856		(1.830)
Proportion Asian	-4.507	*	(2.294)
Proportion Indigenous	-2.984		(1.853)
Mean automobiles	-0.258		(0.176)
Mean computers	0.091		(0.298)
Mean fridges	-0.539	**	(0.259)
Mean freezers	0.381		(0.273)
Mean radios	0.050		(0.153)
Mean washing machines	0.096		(0.154)
Mean tumble dryer	0.144		(0.403)
Mean DVD players	0.109		(0.200)
Mean TV sets	-0.013		(0.155)
Mean bathrooms	0.317		(0.229)
Mean videos	-0.264		(0.250)
Constant	4.157	*	(1.878)
Observations	363		
R-squared	0.236		

Notes: The coefficients come from a linear probability model on the selected assignment rule of students into classes, where the outcome is a dummy taking a value of 1 if students are assigned to classes using their relative age to form *homogenous* classes, and 0 otherwise. Heteroskedasticity robust standard errors are reported in parenthesis. *, ** and *** denote significance at the 10%. 5% and 1% level, respectively.

TABLE A3
MEANS OF STUDENT STATEMENTS ON TEACHING PRACTICES AND PEER BEHAVIOUR

	Mean	Std. Dev.
Teacher enforces student attention	0.918	(0.067)
Teacher corrects homework	0.788	(0.119)
Teacher availability to clarify doubts	0.904	(0.079)
Teacher explains until all students understand	0.891	(0.083)
Teacher gives opportunity to express oneself	0.850	(0.098)
Teacher helps more some students	0.254	(0.147)
Teacher interested in learning progress	0.917	(0.072)
Teacher needs to wait to start teaching	0.581	(0.166)
Teacher absenteeism	0.237	(0.130)
Fellow students leave classroom early	0.269	(0.150)
Fellow students are noisy and disruptive	0.527	(0.140)
Fellow students learn taught material	0.866	(0.087)
Fellow students pay attention in class	0.652	(0.108)

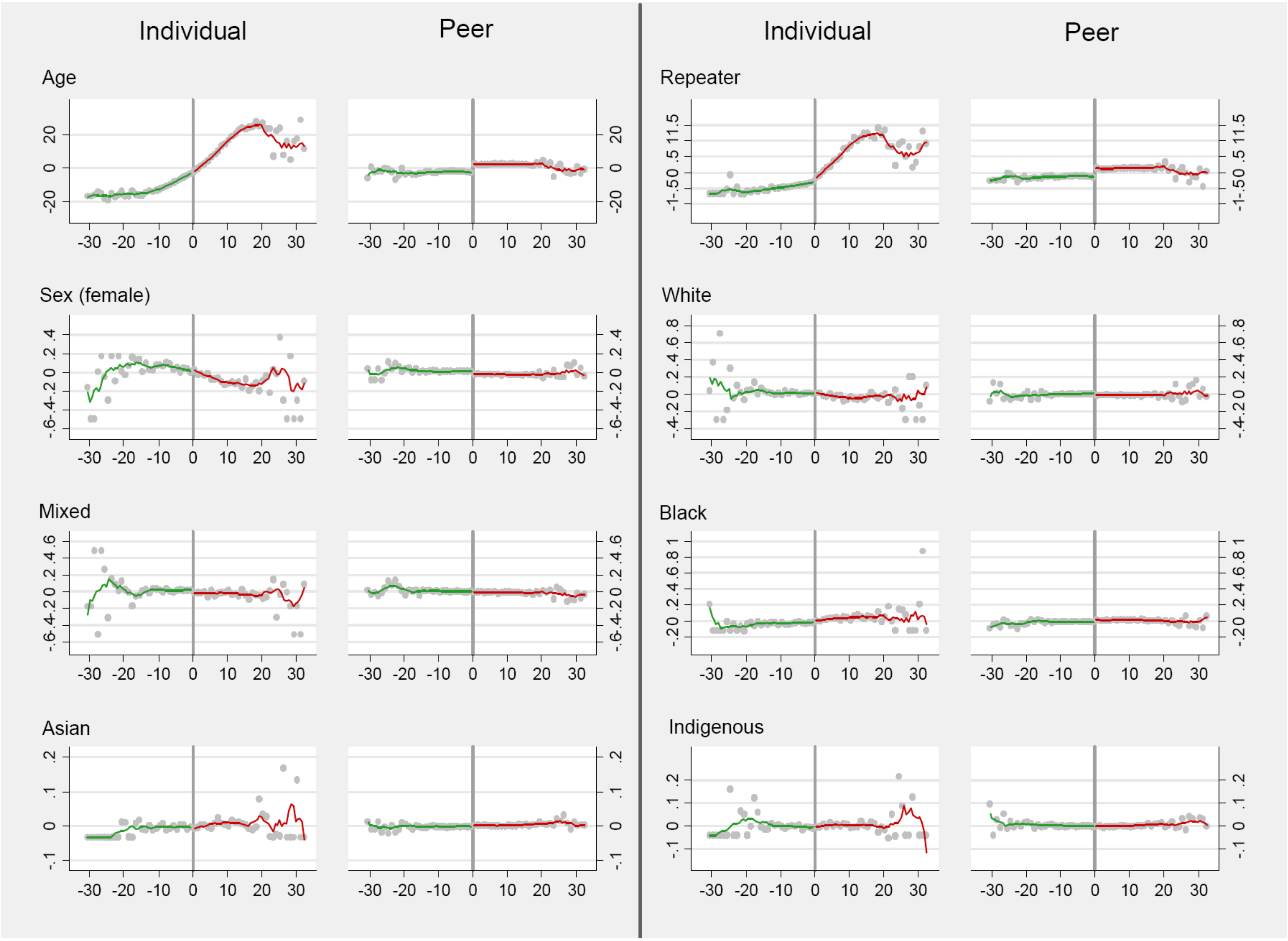
Notes: Entries are means of the standardized categorical answers to the student questionnaire aggregated on the school level. The data comes from the student questionnaire of PROEB 2007. Standard deviations reported in parenthesis.

TABLE A4
RD ESTIMATES OF PREDETERMINED INDIVIDUAL CHARACTERISTICS OF THE 2007 ENTRY COHORT

Sex	-0.069	(0.070)
White	0.006	(0.087)
Mixed	-0.047	(0.053)
Black	0.050	(0.060)
Asian	0.024	(0.034)
Indigenous	0.004	(0.006)

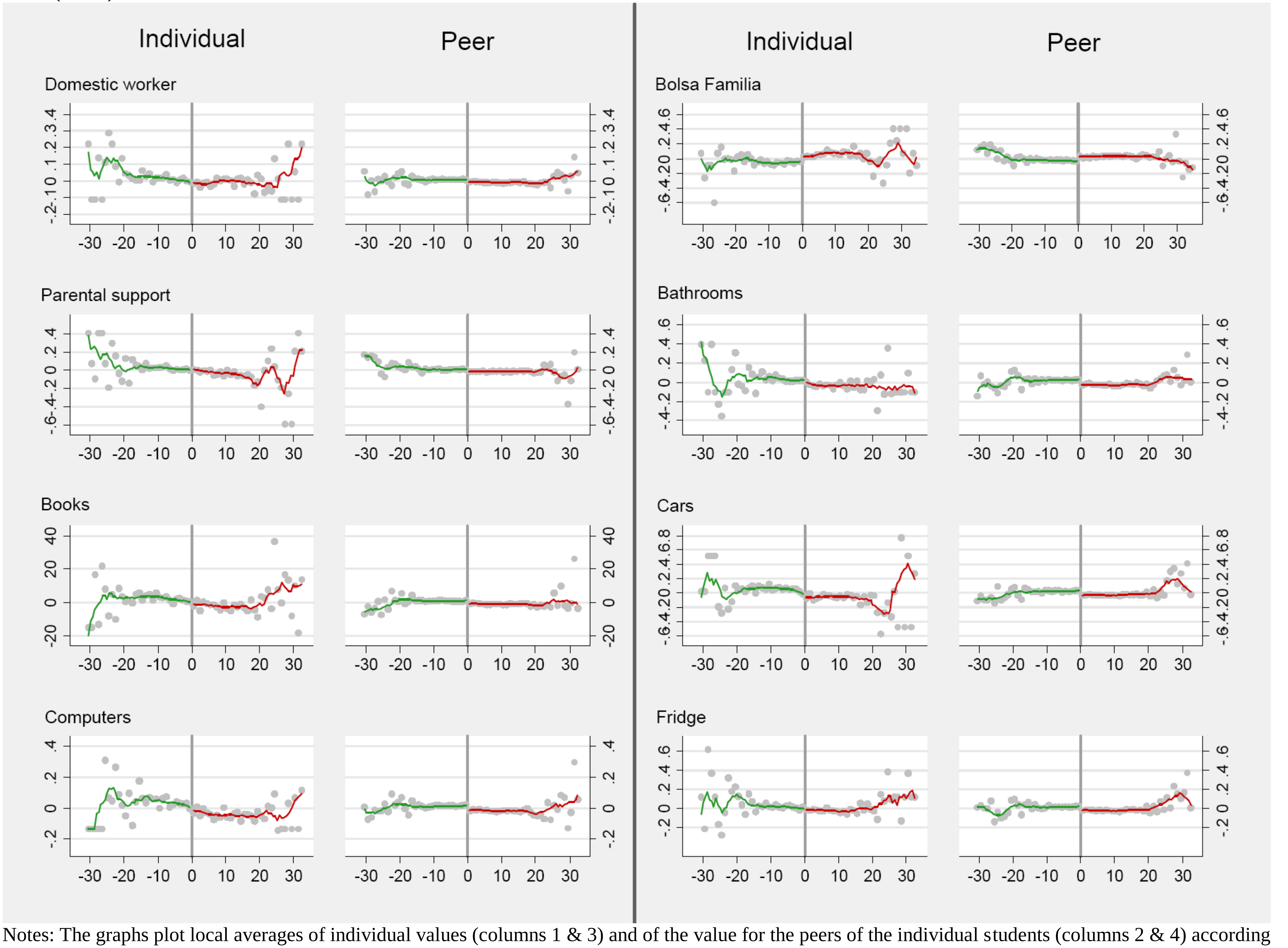
Notes: Entries are separate IV estimates of the class effect on student characteristics of first grade students of the school entry cohort of 2007, where being in the second class has been instrumented by a dummy for having an age rank larger than 0. The data comes from official records of the 2007 school census. For each variable a separate regression has been estimated. All specifications include a second-order polynomial in the age rank of students. Heteroskedasticity consistent standard errors, clustered on the school level are reported in parentheses.

FIGURE A1: TEST FOR DISCONTINUITY OF INDIVIDUAL AND PEER VALUES OF PRE-DETERMINED CHARACTERISTICS (1)



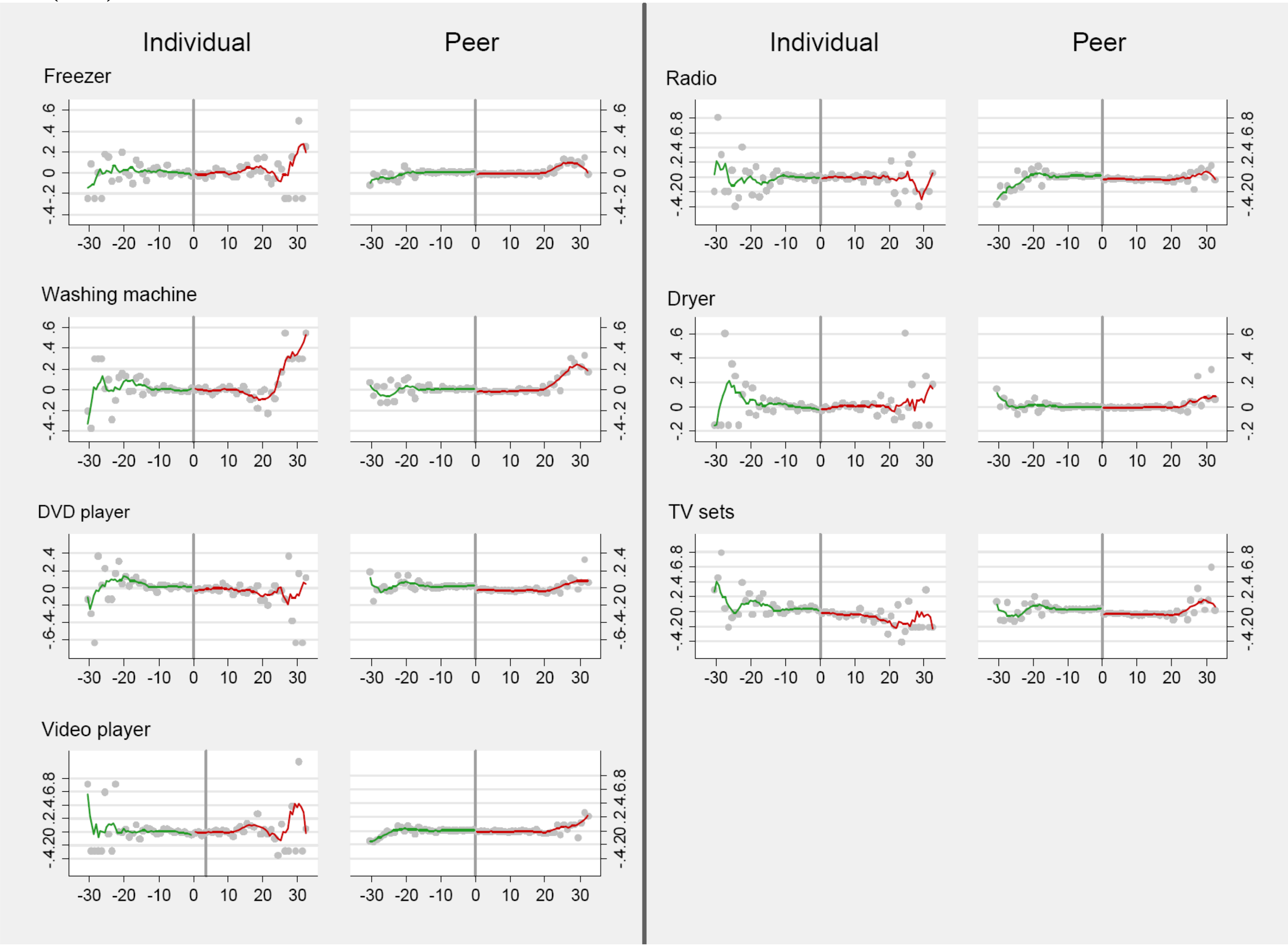
Notes: The graphs plot local averages of individual values (columns 1 & 3) and of the value for the peers of the individual students (columns 2 & 4) according to the age ranking in the cohort as distance of students from the cut-off point and local linear regression fits on both sides of the cut-off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

FIGURE A1 (cont.): 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 A1 (cont.): 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.