I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11211

The Effect of Multigrade Classes on Cognitive and Non-Cognitive Skills: Causal Evidence Exploiting Minimum Class Size Rules in Italy

Daniele Checchi
Maria De Paola

DECEMBER 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11211

The Effect of Multigrade Classes on Cognitive and Non-Cognitive Skills: Causal Evidence Exploiting Minimum Class Size Rules in Italy

Daniele Checchi
University of Milan and IZA

Maria De Paola
University of Calabria and IZA

DECEMBER 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Effect of Multigrade Classes on Cognitive and Non-Cognitive Skills: Causal Evidence Exploiting Minimum Class Size Rules in Italy*

We analyse how schooling in multigrade classes affect the formation of student cognitive and non-cognitive skills. Our identification strategy is based on some institutional features of the Italian educational system establishing a minimum number of students per class. Classes that do not reach the minimum number of pupils are organized in multigrade classes. In addition, the Italian law also establishes a maximum number of students for multigrade classes, leaving a narrow range in the number of students per class which is very similar to small class size in single grade classes when the number of students enrolled in the grade is just above the minimum number of students per class. Using census data on 5th grade Italian students, we find that pupils in multigrade classrooms obtain worse test scores both in literacy and numeracy standardized tests compared to comparable pupils in single grade classroom. While the effect is small and not always statistically significant for the literacy score, we find a large and highly statistically significant effect on the numeracy score. No effect is found on grades assigned by teachers. We also find that pupils placed in multigrade classes tend to have a more external centred locus of control. Our results are robust to different specifications including controls for class size and a number of student and school characteristics.

JEL Classification: I21, I28, C36

Keywords: multigrade classes, mixed-age classes, cognitive skills, non-cognitive skills

Corresponding author:
Daniele Checchi
University of Milan
Via Festa del Perdono 7
Milan, 20122
Italy

E-mail: daniele.checchi@unimi.it

* Paper presented to the Invalsi conference on "I DATI INVALSI: UNO STRUMENTO PER LA RICERCA" (Rome, 20/9/2016); we thanks the participants (and especially Marco Paccagnella) for helpful comments. We also thanks Gianna Barbieri (MIUR Statistical office) for providing institutional data, and Patrizia Falzetti (INVALSI) for kindly merging this information with the public use file of students' test scores.

1. Introduction

Multigrade classes, where students from two (or more) adjacent grades are grouped within one classroom, are common in many developing and developed countries. According to UNESCO (2005 Agenda for Educational Planning) approximately one third of all classes across the world are multigrade classes. In 2007, about 28% of schools in the United States were adopting this type of educational practice when the number of pupils was too small. The incidence of multigrade classes is also high in many European countries, especially in less populated areas. For instance, in France about 37 per cent of primary school pupils are in such classes. In Finland and in the Netherlands multigrade classes prevail over single-grade ones (Mulkeen and Higgings, 2009).

The use of multigrade classes often responds to the need of providing school services at the student's place of residence at a reasonable cost. In fact, in many developed countries the presence of this type of school organization is typical of rural or mountainous areas that in recent years, especially in some countries, have experimented a drastic reduction in resident population: in these circumstances multigrade classes allow schools to remain located closer to the families they serve since there are not enough children to fill a conventional monograde class.

However, cost-saving considerations are to be evaluated against the effects that multigrade classes may produce on student outcomes. From a theoretical point of view, the effects produced by grade-mixing on student achievement can be either positive or negative. On the one hand, such a diverse environment in terms of age, skills and maturity can foster cognitive skills. On the other hand, the fact that teachers are forced to jump from one program to another and to interact with pupils with different needs and different skills might slow down the learning process.

It could also be that, since teachers might be induced to adopt a different teaching approach and students are interacting in a very peculiar environment, multigrade classes affect students' non-cognitive skills. For instance, it could be that multigrade classes involve a more individualized teaching style that might affect socio-emotional skills. In addition, since multigrade classes are more heterogeneous in terms of pupils' age it could become more difficult for students to understand whether the results they get compared to those obtained by their peers are due to ability and effort or are instead related to age. If success or un-success depends on being the younger or the oldest in a group, why bother expending any effort? Age heterogeneity can also affect the quality of relationships among peers and through this channel affect self-esteem, locus of control and so on.

Furthermore, grade mixing could produce differentiated effects depending on the characteristics of students. For example, it may be beneficial for younger students who come into contact with older students, and instead produce negative effects on the latter, who are forced to interact with classmates with a level of skills below their own. However, in many cases, since multigrade classes are typically formed combining adjacent grades, the peer group changes over the school cycle. For instance, in Italy, it is quite common to

have the first three grades (1st, 2nd and 3rd) grouped in a classroom and the last two (4th and 5th grade) in another classroom. Then, a student who starts primary school in a multigrade class during her/his first year interacts with older peers but in the following years she/he finds himself in a class with younger peers. This implies that the cumulative effect of multigrade classes is the sum of the effect of sharing the classroom with higher grade mates and the effect of sharing the classroom with lower grade ones. If these effects go in opposite directions the net effect might be either positive or negative (or even nil).

The empirical literature trying to figure out which of the effects produced by multigrade classes is more important and to understand if students of a mixed-grade classroom are actually penalized or advantaged is scarce. This is mainly due to non-random selection into multigrade classes. For instance, schools might be more likely to adopt multigrade classes if they expect to obtain better results or if the teaching staff is more sympathetic with this type of educational practice. Teachers might try to avoid multigrade classes because of the higher effort required by teaching different programs and selection might not be random. In addition, the assignment of students to multigrade classes might depend on their unobservable characteristics. The few studies that have tried to solve these problems with proper techniques show mixed results.[1] Sims (2008) uses an instrumental variable strategy based on class size caps imposed by the California Class Size Reduction Program and shows that multigrade classes negatively affect test scores in Grades 2 and 3. Negative effects both on final grade attainment and labour market participation are also found by Gerhardts et al. (2016) who exploit a natural experiment deriving from the abolition of parochial schools in Germany. Instead, Thomas (2012), adopting a school fixed effects method, finds that first graders are not harmed by being in a multigrade classes. Finally, Leuven and Ronning (2014), exploiting discontinuous grade mixing rules in Norwegian junior high schools, show that the presence of younger peers decreases achievement, while the reverse occurs in case of older peers. Even more scant is the evidence on the effects of multigrade classes on non-cognitive skills. The only paper looking at non-cognitive skills is Sattari (2016) who shows that placing students in multigrade classrooms causes more behavioural problems.

Our analysis contributes to this literature by providing additional evidence of the effects of multigrade classes both on cognitive and non-cognitive skills. At this aim we use a very rich dataset covering the entire populations of students attending the 5th grade in Italian public schools and providing information on students' performance in standardized test scores, grades assigned by teachers and on a measure of locus of control. Our identification strategy takes advantage of some institutional features of the Italian school legislation. According to existing rule (DM 331/98) primary school classes must consist of no less than 10 children[2] (a maximum of 25 pupils per class is also defined). Classes that do not respect the minimum number of pupils are organized in multigrade classes. However, according to the Italian law, the possibility

[1] A review of earlier studies is provided by Veenman (1995) who surveys 56 papers and concludes that pupils in multigrade classrooms show results that are similar to those reached by pupils in classrooms that track pupils by grade. However, as pointed out by Mason and Burns (1997), many of these studies do not address sorting of pupils and teachers into multigrade classes.

[2] A 2009/10 reform increased the minimum size to 15 in hilly and plain areas (leaving the threshold of no less than 10 children in mountain areas and small islands). This reform was rolled out one grade per year, starting with grade 1. Students considered in our data remain subject to the old rules.

to form a multigrade class depends not only the number of students enrolled in a given grade, but is also related to the number of students enrolled in adjacent grades. In fact, the law establishes that multigrade classes are subject to a maximum number of 12 students per class (with a minimum of 6). This additional constraint implies that class size in multigrade classes is very similar to class size in single grade classes with a number of students enrolled in the grade just above the cut-off point of 10. Exploiting these rules we build a binary multigrade predictor taking the value of 1 when the two following conditions are met: 1) the number of students enrolled in the grade is smaller than 10; 2) the total number of students enrolled in adjacent grades is smaller than 13.[3] Even if these rules are not strictly respected (since school principal had some margin for flexibility) they represent a source of exogenous variation that we exploit to identify the effect of multigrade classes.

In addition, to compare similar classes and minimize problems related to endogeneity in class size, we only focus on small schools that have no more than a classroom per grade. In these schools class size can be considered exogenous and mainly determined by variation in cohort size. Since there is only a classroom per grade there is less room for parents' and school administrators' choices and variation in class size is mainly related to natural randomness in population (Hoxby, 2000). In addition, to minimize the impact of class size we restrict our analysis to classes with no more than 13 students and no less than 5 students. This restriction, as we show in the appendix of the paper, does not affect our results that remain substantially unchanged also when considering the whole sample, but reassures us that the estimated effect is not driven by differences in class size.

We find a negative effect on students' performance both in literacy and numeracy standardized test scores. While the effect is not always statistically significant for the literacy score, we find a large and highly statistically significant effect on the numeracy score. Students placed in multigrade classes obtain a score in numeracy of about half a standard deviation lower than students in single grade classes. This effect holds true when we control for class size and for a number of students and school characteristics. Since our results pertain to students attending the 5th grade who are likely to have been in a multigrade class also in the previous years of their educational process (unfortunately we do not have information on this), it is probable that the negative effect we find is the cumulated effect of having attended a substantial part of the primary school cycle in a multigrade environment.

On the contrary, we find a negative but statistically insignificant effect when we look at grades assigned by teachers. This might depend on the fact that teachers' evaluations, in spite of standardized test scores, are the results of a more complex assessment process, which reflects the objective level of skills achieved by students, but also a number of other factors such as the perceived student effort, motivation, behaviour as well as parents' expectations (OECD, 2012; 2013). Then, it could be that teachers of multigrade classes, aware of the more complex environment faced by their students, tend to reward more generously their effort. We find that this is especially true for teachers working in the Southern part of the country, while

[3] The second condition suggests that in each adjacent class there are 6 or 7 students per grade. As a consequence a multigrade class can be formed by joining two adjacent classes while respecting the minimal and maximal threshold.

for schools located in the North we find that multigrade classes harm students' skills also when measured using teachers' assessment.

Thanks to a survey proposed to students taking the INVALSI test, we are also able to analyse the effect of multigrade classes on locus of control, a psychological trait that has received considerable attention both in the psychological and economic literature. Locus of control captures '*a generalised attitude, belief or expectancy regarding the nature of the causal relationship between one's own behaviour and its consequences*' (Rotter, 1966). Individuals who believe that life's outcomes are due to their own efforts have an internal locus of control, while those believing that outcomes are due to external factors (e.g. luck) have an external locus of control (Gatz and Karel, 1993). Locus of control has been shown to explain a wide range of social and economic outcomes, such as educational attainment, earnings, unemployment and job search behaviour, life satisfaction and health investments (see for instance, Caliendro et al., 2015; Cebi, 2007; Chiteji, 2010; Coleman and Deleire, 2003; Osborne Groves, 2005). In addition, individuals with an internal locus of control are more able to cope with unanticipated life events such as health shocks and unemployment (Schurer, 2011; Caliendo *et al*., 2015). In line with an emerging literature that describes non-cognitive skills as resulting from educational attainment, parental investments and policy interventions (Almlund et al., 2011), we consider the impact of educational inputs on individual locus of control. We find that students placed in multigrade classes are more likely to have an external locus of control. However, the effect is not robust when, instead of focusing on small classes, we consider the whole sample of students.

The rest of the paper is organized as follows. Section 2 provides a description of the institutional setting of Italian schools and explains the rules followed to decide whether teaching activity will take place in a multigrade or in a single grade class. Section 3 presents our estimation strategy and discusses possible threats to the validity of our research design providing a number of checks on our first stage. Estimates of the effects of multigrade classes on cognitive skills are reported in Section 4, while in Section 5 we analyse the effects on locus of control. Section 6 presents some robustness checks. Section 7 concludes.

2. Institutional background and Data

In 2015 about 51 thousand children (1.5% of the whole population of students enrolled in primary school) attended primary school in a multigrade grade classroom in Italy. The majority of these students lives in inner areas that are at a considerable distances from main towns and offer only limited access to essential services such as education, mobility, health, etc. In 2016, one quarter of the Italian population lived in these areas, which represent about 60 per cent of the national territory with 4261 municipalities with an average population of 1.500 inhabitants. In recent years, these municipalities have suffered a strong population decline and demographic ageing which has led to smaller and smaller schools with a low number of students per class (in 2016, in 25% of these municipalities the average number of pupils per class in primary school was less than 10). In these areas multigrade classes can be a cost-effective tool to retain locally provided

education and to avoid to pupils and their families the cost associated to attending schools located in a different municipality.

Teachers in Italian primary schools are required to have obtained a university degree in Education and they have to teach an identical nation-wide curriculum, defined for each grade. This implies that students are taught the same curriculum by teachers with similar qualification, irrespective of whether they attend a single grade class or a multigrade one. Rules about the number of teachers in each class have changed over time; until 1990 there was a single teacher for each class, in the period going from 1990 to 2008 there were two teachers per class, while in the school year 2008-2009 a new reform (*riforma Gelmini*) has reintroduced the possibility to have a single teacher for each class (*maestro unico*). The cohort of students we consider in this study (enrolled at the first grade in 2007-2008) experienced two different teachers specialized in the main subjects. Teaching organization is the same for multigrade and single grade classes.

The allocation of students to multigrade classes is decided following a mix of rules and discretion. In the Italian educational system, until 2008/09 school year, primary school classes were subject to a minimum size of 10 and a maximum of 25 (Decreto Ministeriale 331/98). In 2009/10 the minimum and the maximum were increased to 15 and 27 respectively (Decreto del Presidente della Repubblica 81/2009), with the exception of schools in mountain areas and small islands where the minimum number of pupils was retained at 10. If the number of students in a specific grade was lower than that threshold, school heads were required to form mixed-grade classes (*pluriclassi*), grouping students of adjacent grades. The Italian law also established that multigrade classes cannot be composed by more than 12 students (changed to 18 students in 2009). The 2009 reform was rolled out one grade per year starting with grade 1 and students considered in our data were subject to the old rule. On these basis, we build a dummy variable *Predicted Multigrade* taking the value of one when the number of students enrolled in 5^{th} grade is smaller than 10 and if the total number of students enrolled in grades 4 and 5 is smaller than 13. Since both the cap on the minimum number of students per class and the cap on the maximum number of students in multigrade classes could be applied with a discretionary margin of 10% above/below the numbers set by law, *Predicted Multigrade* does not perfectly predict student placement in a multigrade class.

To investigate the effect of multigrade classes on student outcomes we rely on data from the Italian National Assessment Program, INVALSI, a government agency that carries out a yearly testing of student attainment in literacy and numeracy. The evaluation covers the entire population of students attending 2^{nd} and 5^{th} grade (primary school), as well as 8^{th} and 10^{th} graders (lower and upper secondary schools respectively). The dataset provides information not only on standardized test score results but also on the marks assigned by math and language teachers.

In our work we focus on primary schools because the adoption of multigrade classes is more diffuse at this level of education (quite often small municipalities do not have secondary schools). We restrict our analysis to students in the 5^{th} grade since their potential outcomes include standardized test scores and a

number of non-cognitive skills.[4] Data are from the 2011-12 wave for which we also have information on the actual number of multigrade classes within a school, thanks to additional administrative data (*Rilevazione integrativa, Ministero dell'Istruzione, dell'Università e della Ricerca*, MIUR). .

In addition, to reduce the risk of self-selection of students and teachers in different classes and to better identify students effectively attending a multigrade class we restrict to schools that have no more than a classroom for the 5th grade. To compare students attending classes of similar size, we also exclude from our analysis classes with more than 13 students and with less than 5 students (respectively the maximum and the minimum number of students allowed in multigrade classes, applying the discretionary margin of 10%). Finally, we restrict our sample to students that undertook both the language and the math test.[5] Our final sample consists of 16,665 observations in 1,879 schools. This amounts to about 16% of pupils attending the 5th grade in Italian primary schools in 2011-12.

Using the information on the actual number of multigrade classes in the school, provided by administrative sources, we try to infer whether students in our dataset were effectively placed in a multigrade class. Since we do not have detailed information at class level we rely on two different definitions. The first one considers as students attending a multigrade class only students who are enrolled in a school in which there are two multigrade classes. Following this rule we build a dummy variable *Actual Multigrade1*. As we are focusing on small schools and on small classes it is very likely that this definition catches students who are effectively attending a multigrade class. According to this measure about 24% of students in our sample are placed in a multigrade class. On the contrary, if we look at the whole sample of students (Table A1 in the Appendix) we find that about 1% of students are placed in a multigrade class. This percentage seems consistent with the total number of students attending a multigrade class in Italian primary schools. However, we are misplacing those students attending a multigrade class in a school in which the total number of students enrolled is so low that teaching activity is organized in a single multigrade class (covering from grade 1 to grade 5). These students end up included in the control group.

The second definition, *Actual Multigrade2,* is wider and considers as students attending a multigrade class those who attend a school in which there is at least a multigrade class.[6] Under this second definition, the percentage of students placed in multigrade classes increases to 38% when considering this measure, while if we consider the whole sample we find that about 2% of student are placed in a multigrade class. This definition allows the inclusion of all students attending a multigrade class in the "treated group", but it is likely that also some students actually not attending a multigrade class are considered as "treated".

[4]Using data on test scores of 2nd grade students we find that being placed in a multigrade class negatively affects student performance. The effect is smaller in magnitude compared to what we find for 5th graders. For 2nd grade students no information is available on non-cognitive skills. Results available from the authors.

[5]As the language and math tests were held on different days if we do not restrict the sample to students undertaking both tests we end up with two slightly different samples for the language and math outcomes depending on the number of students that were absent during one of the two tests. Results do not change when using full available samples.

[6] We have also experimented with a very restrictive definition considering as students attending a multigrade class only students who are enrolled in a school in which there is at least a multigrade class and in which the total number of students enrolled in the school divided by the number of multigrade classes in the school is smaller than the maximum of students allowed for multigrade classes. We find qualitatively the same results reported in the paper but the effects are larger in magnitude.

As regards the outcome variables, the public use files provide two alternative measures of student performance at standardized tests: a) the fractions of correct answers in literacy and numeracy multiple choice tests (*Literacy Score* and *Numeracy Score*); b) scores computed by INVALSI applying the IRT Rasch model to students' answers in the tests, in order to account for different difficulties of single items (*Rasch Literacy Score* and *Rasch Numeracy Score*)[7]. Since the data come from a national test which is common to all schools, the performance of students attending the same grade are by construction comparable across schools in different geographical areas of the country. In order to avoid problems deriving from score manipulation, both these measures are expressed as "cheating-corrected" test scores.[8]

We also have information on the marks assigned by teachers in the two main subjects in the primary school program: Italian Language and Mathematics. Data collected by INVALSI allow the distinction between "written marks" and "oral marks". We have considered oral marks (*Teacher Marks Numeracy, Teacher Marks Literacy*), but results do not change qualitatively if we consider the written marks or the average value of written and oral marks.[9] Teachers' marks and INVALSI test scores are positively correlated (the correlation between *Rasch Literacy Score* and Italian Language oral mark is 0.50, while the correlation between *Rasch Numeracy Score* and Maths oral mark is 0.33) but there are some relevant differences. The INVALSI tests are identical across schools while marks given by teachers are based on tests autonomously set by each teacher. Then, while INVALSI scores are comparable across schools, this is not the case for teachers' marks. In addition, while INVALSI tests assess student performance on an absolute grading scale, teachers might adopt relative marking which might also be affected by class composition.

Apart from measures of cognitive skills, the INVALSI dataset allows also to build some measures of non-cognitive skills. Using the survey submitted to students (Student Questionnaire) the same day of one of the two tests, we consider eight questions allowing us to construct a locus of control measure, i.e. the extent to which a person believes her\his actions affect her\his outcomes. Five of these questions refer to successful situations[10] and three of them to unsuccessful ones[11]. The student is asked to attribute each situation to: 1)

[7] In such a case these scores are standardized to have a mean of 200 and a standard deviation of 40.

[8] As documented by Angrist et al (2017) and Bertoni et al. (2013) many schools follow a "cheating to the test" practices. Since cheating significantly affects the reliability of test scores, INVALSI has developed a statistically solution to purge the data from this problem. This method exploits the statistical properties of the distribution of answers given in classes where the test is taken under the supervision of external examiners (randomly assigned to selected classes and schools with the task of monitoring), and calculates a continuous class-level probability of manipulation (similar to that estimated in Angrist et al. 2017). This probability is based on the variability of intra-class percentage of correct answers, modes of wrong answers, etc.; the resulting estimates are used to "deflate" the raw scores in the test. For a detailed description of the method see INVALSI (2010).

[9] The correlation between written and oral mark is 0.9619 and 0.9712 for Italian Language and Mathematics respectively (p-value 0.000).

[10] This is the list of questions: "1) The teacher asked you to make a picture and you did it very well. How did you do?; 2) The teacher asks you to repeat a story you read together in class and you did it very well. How did you do?; 3) On the first day of school, the teacher asks you to tell what you did during the holidays, you tell it so well that all your schoolmates have fun. How did you do?; 4) At the recital at the end of the year you performed your part so well that everyone applauded. How did you do?; 5) The teacher asked you to do a math exercise on the blackboard and you did it very well. How did you do?"

[11] This is the list of questions. "1) The teacher asks you to write a theme, but you make many mistakes. Why? 2) The teacher asks you to repeat a poem you've learned, but you do not remember it very well and make a lot of mistakes. Why? 3) The teacher asks you to do a work for Christmas, but it comes out very bad and you had to do it again. Why?"

help or lack of help from others; 2) lucky or unlucky circumstances; 3) easiness or difficulty of the task; 4) own ability; 5) own effort. The choice of the first three options denotes an external locus of control (outcomes depend on luck or external factors), while the choice of the last two options is considered as an indicator of an internal locus of control (outcomes depend on own ability and effort). Then for each of these questions we build a dummy variable taking the value of 1 when the student picks one of the first three possible answers and zero otherwise. Using students' answers to these questions we build three different measures of locus of control. The first, *External Locus of Control,* is based on the whole set of questions and takes values from 0 (when the student answering to the 8 questions never chooses one of the first three options) to 8 (when the student chooses one of the first three options for each of the 8 questions). The second *External Locus of Control Positive* is based only on questions proposing successful situation and takes values from 0 to 5. The third *External Locus of Control Negative* is instead based exclusively on questions proposing unsuccessful situations and takes values from 0 to 3.

The dataset at hand also provides information on a number of children and parents characteristics (gender, citizenship, attendance of pre-primary school, parent working status and education). Information on the family background of the student are used by INVALSI to build an indicator of socioeconomic status (called ESCS-Economic and Social Cultural Status)[12], out of which principal component analysis is applied, obtaining a variable with zero mean and unitary standard deviation. We also have information on whether the student is younger or older than a regular students (we build a dummy variable for students who went to school one year before the suggested age, *Early Enrolled*, and dummy variable for students who entered the school one year after or repeated one or more years, *Late Enrolled*)

We also have information on the number of students enrolled in each grade at the beginning of the school year. For single grade classes the number of enrolled students in the grade corresponds to class size, instead for multigrade classes class size is constructed considering the number of enrolled students in the different grades composing the class.[13]

As regard school organization we know whether the class follows a full or part-time schedule and on the basis of this information we build a dummy variable *Full time* for those classes whose schedule is organized in entire days (8am-4pm usually) instead that only in the morning.

Finally, we have information on the region in which the school is located and on a number of different school catchment area characteristics (population size, extension and altitude).

[12]This indicator is built in accordance to the one proposed in the OECD-PISA framework and considers parents' occupation, educational attainment and possession of educational resources at home (for instance, the number of books). For a detailed description see Ricci (2010), http://new.sis-statistica.org/wp-content/uploads/2013/09/RS10-SP-The-Economic-Social-and-Cultural-Background-a-continuous-index-for-the-Italian-Students-of-the-fifth-grade.pdf

[13]To have information on students enrolled in 4th grade in the academic year 2011-2012, we consider the information provided by the 2012-13 INVALSI wave when these students are in 5th grade (for no test is undertaken in fourth grade). Since, as we explain in more detail below, retention is quite rare in Italian primary schools, this number is likely to be quite close to what it was the year before. Then, for multigrade classes we calculate class size summing the number of students in 5th grade to the number of students enrolled in 4th grade: if this number does not reach the minimum class size imposed for multigrade classes we also add the number of students enrolled in 3rd grade, obtained from the 2013-14 INVALSI wave.

In Panel (a) of Table 1 we report descriptive statistics for the whole sample used in our analysis.[14] *Predicted Multigrade* takes an average value 0.339, implying that according to rules defined by the Italian law about 34% of students in our sample should be placed in a multigrade class. The actual number of students in our sample effectively attending a multigrade class is however smaller and equal to 24% (*Actual Multigrade1*).

Descriptive statistics for the sample of students placed in single grade classes are reported in Panel (b) of Table 1, while in Panel (c) are reported descriptive statistics for students in a multigrade class (Actual Multigrade1).

If we compare students in the sample used for our main analysis (Table 1, Panel (a)) with the total population of students enrolled in grade 5th (see Table 1A in the Appendix of the paper) we find a number of differences with respect to individual background that depend on the fact that the students we consider in our analysis live in small villages, typically characterized by poorer economic conditions. In fact, students in our sample are from less wealthy families, less likely to have an immigrant background and to have attended pre-primary school. They obtain worse scores both in literacy and numeracy, while grades assigned by teachers are only slightly lower compared to those observed for the whole population. On the contrary, there are not relevant differences as regards the percentage of students regularly enrolled.

When we compare students in our sample placed in single grade classes (Table 1, Panel (b)) with those placed in multigrade ones (Table 1, Panel (c)) we find that students in multigrade classes are comparable in terms of a number of observable characteristics, such as gender, father and mother nationality. There are, however, some statistically significant differences in terms of ESCS index, percentage of regular students, class size, and attendance of pre-primary school. Most notably, multigrade class tend to be smaller and to have a higher percentage of students from poorer background (the difference is however small -0.040 statistically significant at 10 percent level). In addition, the average score obtained both in literacy and numeracy standardized tests is lower for students in multigrade classes compared to students in single grade classes.

[14] In the Appendix of the paper we report descriptive statistics for the whole sample of 5th grade students undertaking the INVALSI standardized test.

Table 1. Descriptive statistics

	Panel (a) Whole sample			Panel (b) Single Grade			Panel (c) Actual Multigrade1		
	Mean	St. Dev.	Obs.	Mean	St. Dev.	Obs.	Mean	St. Dev.	Obs.
Actual Multigrade1	0.241	0.428	16.665						
Actual Multigrade2	0.386	0.487	16.665						
Predicted Multigrade	0.333	0.471	16.665	0.206	0.405	12.647	0.757	0.429	4.018
Rasch Literacy Score	197.363	41.525	15.155	198.291	41.787	11.535	194.403	40.544	3.620
Literacy Score	70.361	21.507	16.665	70.799	21.339	12.647	68.980	21.972	4.018
Teacher Mark Literacy	7.577	1.148	15.509	7.481	1.117	11.686	7.505	1.081	3.771
Rasch Numeracy Score	173.589	41.649	14.891	175.414	41.136	11.304	167.837	42.722	3.587
Numeracy Score	49.407	22.081	16.665	49.613	22.297	12.647	48.758	21.378	4.018
Teacher Mark Numeracy	7.486	1.109	15.457	7.571	1.156	11.721	7.595	1.124	3.788
External Locus of Control	2.256	1.674	16.091	2.240	1.677	12.203	2.307	1.666	3.888
Ext Locus Control Positive	1.359	1.165	16.220	1.353	1.168	12.303	1.377	1.154	3.917
Ext.Locus Control Negative	0.902	0.963	16.353	0.892	0.958	12.402	0.933	0.977	3.951
Female	0.491	0.500	16.665	0.491	0.500	12.647	0.488	0.500	4.018
Regularly Enrolled	0.940	0.237	16.665	0.936	0.245	12.647	0.954	0.210	4.018
Early Enrolled	0.018	0.134	16.665	0.021	0.143	12.647	0.010	0.099	4.018
Late Enrolled	0.041	0.199	16.665	0.043	0.203	12.647	0.036	0.186	4.018
Pre Primary School	0.115	0.319	16.665	0.123	0.329	12.647	0.087	0.282	4.018
ESCS index	-0.016	0.977	16.665	-0.006	0.992	12.647	-0.046	0.926	4.018
Italian Father	0.854	0.353	16.665	0.852	0.355	12.647	0.862	0.345	4.018
Italian Mather	0.837	0.370	16.665	0.836	0.371	12.647	0.840	0.367	4.018
Full time	0.165	0.371	16.665	0.168	0.374	12.647	0.155	0.362	4.018
# student enrolled in grade 5	10.218	2.615	16.665	11.069	1.929	12.647	7.540	2.684	4.018
Class size	10.525	2.181	16.665	11.076	1.906	12.647	8.793	2.082	4.018
Southern regions	0.419	0.493	16.665	0.433	0.495	12.647	0.373	0.484	4.018

Notes: The data are drawn from the Invalsi websites (downloaded on 7/2/2016)

3. Estimation Strategy

Distinguishing the effect of grade mixing from the effect produced by other factors that are both relevant for student cognitive and non–cognitive skills and for being placed in a multigrade class is not an easy task. In order to recover the causal impact of multigrade classes on student performance we exploit the rule that sets to 10 the minimum number of students in a class and also requires that multigrade classes cannot be composed by more than 12 students. As these rules introduce a plausibly exogenous variation in treatment status, we use them to identify the impact of multigrade classes on student outcomes. In fact, the application of these rules implies that the probability of attending a mixed-grade class is a discontinuous function of the number of students enrolled in a class and also depends on the number of students enrolled in adjacent grades. Then, we apply an instrumental variable strategy that uses the class size rule as an instrument for being placed in a multigrade class. We estimate the following model:

[1] $$Y_{ij} = \beta_0 + \beta_1 \text{Actual Multigrade}_{ij} + \beta_2 \text{Class Size}_{ij} + \beta_3 X_{ij} + \lambda_k + \varepsilon_{ij}$$

[2] $$\text{Actual Multigrade}_{ij} = \phi_0 + \phi_1 \text{Predicted Multigrade}_{ij} + \phi_2 \text{Class Size}_{ij} + \phi_3 X_{ij} + \lambda_k + \upsilon_{ij}$$

where in equation [1] Y_{ij} is the outcome variable of interest (alternatively the performance of student i enrolled in class j in Literacy and Numeracy or her\his indicator of locus of control); $\text{Actual Multigrade}_{ij}$ is a dummy variable indicating whether the student is attending a multigrade class; Class Size_{ij} is the number of students attending class j; X_{ij} is a vector of individual and school characteristics (gender, ESCS index, mother and father immigrant status, pre-primary school attendance, Early Enrolled, Late Enrolled, Full Time); λ_k are regional fixed effects and ε_{ij} is a random error term. Equation [2] represents the first stage of the relationship between student actual placement in a multigrade class and $\text{Predicted Multigrade}$, that is a dummy variable taking the value of one when the number of students enrolled in 5th grade class j attended by student i is smaller than 10 and when class size in a multigrade class does not exceed the maximum number of 12 students allowed by the law.

The rule establishing a specific maximum number of students for multigrade classes allows us to avoid problems typically encountered by studies that exploit minimum class size rules. These works have to take into account that when a multigrade class is formed there is also a sharp change in class size. In our case, since according to the Italian law, multigrade classes cannot be composed by more than 12 students, class size is similar in multigrade and single grade classes that are just above the minimum threshold of 10 students. In our sample, average class size in multigrade classes is of 8.8, while it is equal to 11 students in single grade classes. However, as described in equation [1], in our estimates we control for class size and to avoid problems that might derive from class size endogeneity we have restricted our analysis to small schools that have no more than one classroom for 5th grade and to classes composed by no more than 13 students and no less than 5 students (the maximum and minimum class size imposed by the reform for multigrade classes, applying the discretionary margin of 10%).

Given the type of schools considered in our analysis is also unlikely that school managers behaved in such a way as to change the number of enrolled students in a given grade when it was near the cut-off point. Also because the composition of a multigrade class also depends on the number of students enrolled in adjacent grades. They might be interested in such manipulations to avoid the reduction of the number of teachers working in the school or (at the opposite extreme) to get rid of undesirable teachers. However, the minimum class size rule applies when the number of enrolled students becomes quite low and this is likely to occur in small and/or isolated municipalities where it is difficult to attract new students. In addition, in Italy grade retention in primary school is very rare and it is very unlikely that teachers and school managers use this variable to reach the minimum class size rule. In our sample 94% of students are regular in their school path, while about 4.1% and 1.8% of them are late or early entrants in their educational path. Delays in students' educational process are mostly due to the fact that non-Italian students at the entry of their stay in the country are often placed in grades lower than those corresponding to their age, in order to improve the

mastery of Italian language. In fact, when we only consider native students who typically start school at the expected age, the percentage of students being late in their educational career drops to 2%.

In order to try to understand whether there has been manipulation in the number of students enrolled in 4th and 5th grades, in Figure 1 and Figure 2 we present the density of grade enrolment in each grade for our sample and the whole sample (including also classes with more than 13 students), respectively. As shown in Figure 1, presenting the density of grade enrolment, there are no suspect discontinuities.

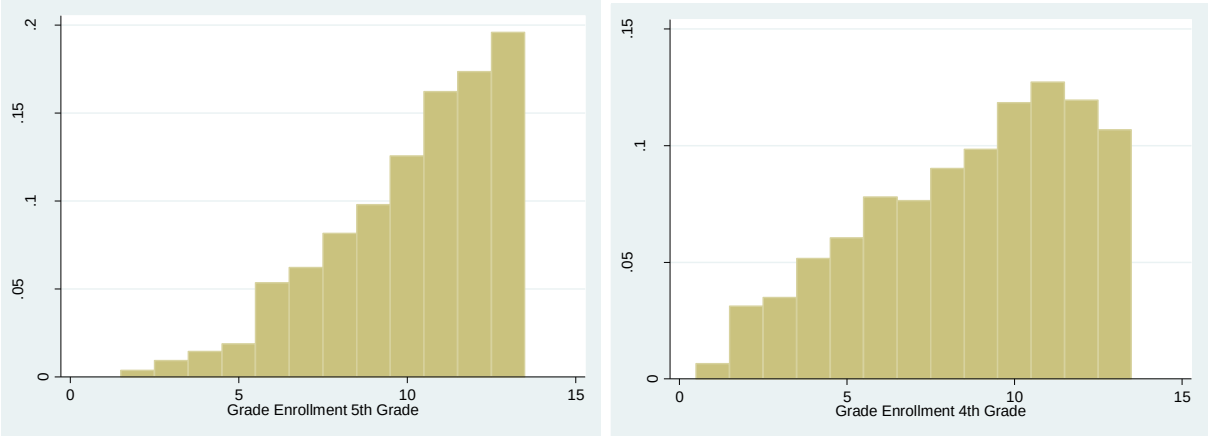

Figure 1. Density of Grade Enrolment, small classes sample

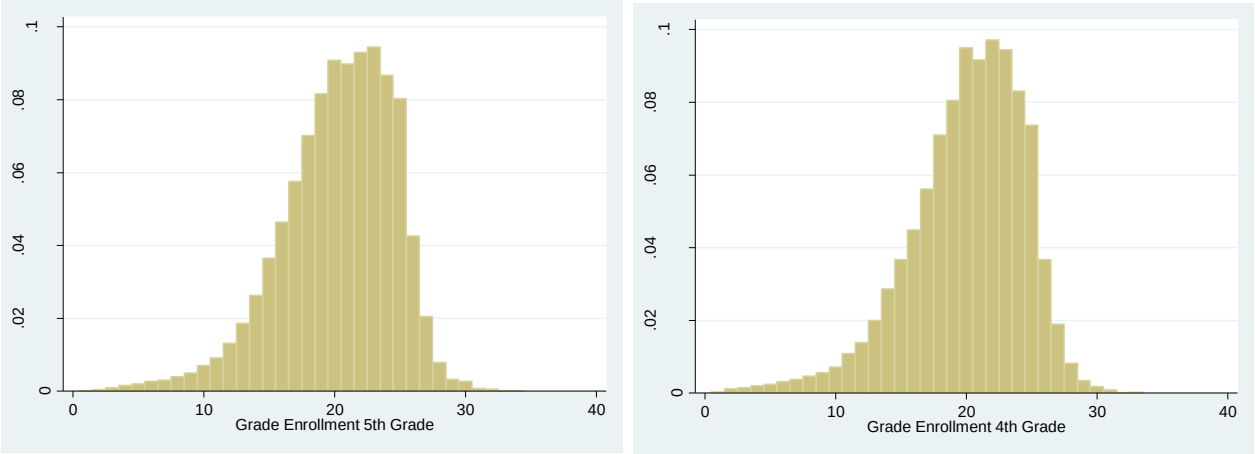

Figure 2. Density of Grade Enrolment, full sample

We have also checked whether our instrument is correlated to student and school characteristic. Even if testing exclusion restriction is not possible since it involves the structural error that is never observable, if there is no relationship between the instrument and observable student and school characteristics one may reasonably expect that also unobserved student and school characteristics be uncorrelated with the instrument. Similarly, to Sims (2009) and Sattari (2016), we regress students' and schools' characteristics on $PredictedMultigrade_{ij}$. Results of our regression are reported in Table 2, where we test whether the

$PredictedMultigrade_{ij}$ is predictive of *Female*, ESCS, *Father Born Abroad*, *Mather Born Abroad*, while controlling for class size. Overall, Table 2 confirms that these variables fail to show a statistically significant correlation with instrument status. However, since not all variables are balanced, we control for these variables in the regressions to avoid any bias due to the lack of balance.

Table 2. Differences in predetermined characteristics. 5th grade – Italy 2011-12

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Female	Regularly Enrolled	Early Enrolled	Late Enrolled	Pre-primary school	ESCS index	Italian Father	Italian Mather	Full time
Predicted Multigrade	-0.034** (0.014)	0.024*** (0.008)	-0.023*** (0.008)	-0.001 (0.008)	-0.008 (.0013)	0.030 (0.042)	0.003 (0.015)	-0.000 (0.015)	-0.032 (0.030)
Obs	16,665	16,665	16,665	16,665	16,665	16,665	16,665	16,665	16,665

The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

As explained in the previous section, we see an imperfect correspondence with the Assigned Treatment and Actual Multigrade. This is due both to the fact that effective treatment is measured with error and to the fact that schools may deviate from the rule. Based on the First Stage Equation, in Figures 3, we plot the probability of attending a mixed-grade class against class size when considering the whole sample of students attending the 5th grade in schools with no more than 5 classrooms (estimates for this sample are reported in the Appendix of the paper). The circles are the means of the probability of effectively attending a mixed grade class for a given Class Size, while the red dots are the predicted values from the first stage equation. As it is possible to see on the left hand side of the graph in Figure 3, the probability of effectively attending a mixed grade class for students in classes below the cut-off point is about 0.5 while it drops 0.01 for students in classes with a number of enrolled students above the threshold.

A similar picture emerges also when focusing on the sample we use in our main analysis. In this case, as shown in Figure 4, the probability of effectively attending a mixed grade class is respectively 0.67 and 0.015 respectively for students in classes below and above the cut-off point.

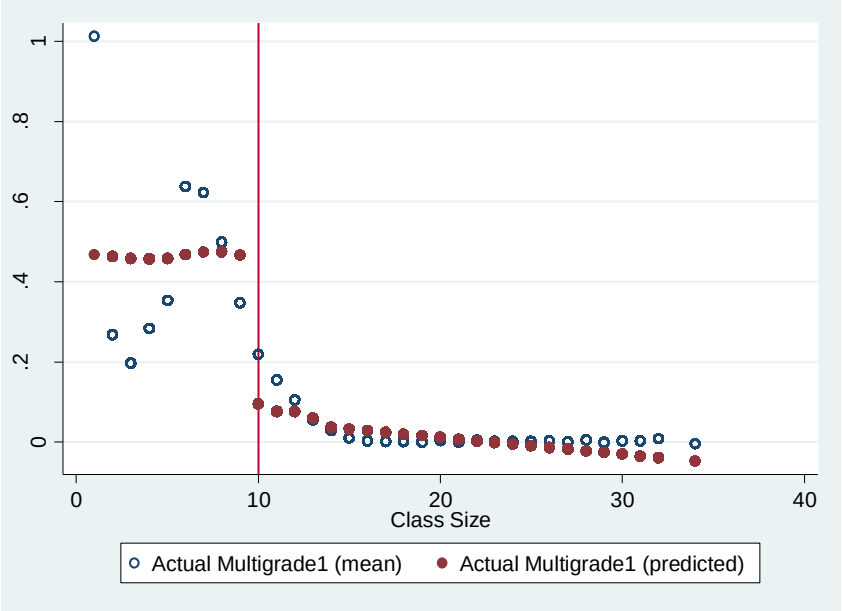

Figure 3. First Stage relationship: Grade Enrolment and Predicted Actual Multigrade. Full sample

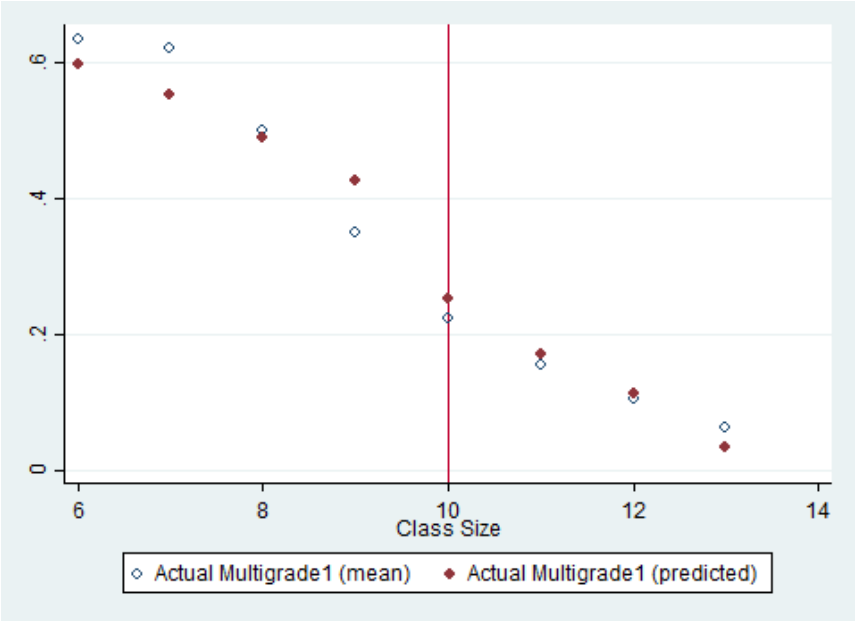

Figure 4. First Stage relationship: Grade Enrolment and Predicted Actual Multigrade. Small classes sample

4. The effect of multigrade classes on cognitive skills

In this section, we report our main results of the effect of being placed in a multigrade class on some measures of student cognitive abilities. The INVALSI dataset provides information on both standardized tests and teacher evaluations. Then, for each student we observe both the marks assigned by math and language teachers and the INVALSI standardized test score results obtained in the same areas during the

same school year. We firstly analyse the impact of multigrade classes on the INVALSI standardized test scores and then we focus our attention on marks assigned by teachers.

4.1. The effect of multigrade classes on standardized test scores

We begin our analysis using as outcome variables students' performance in standardized test scores. Initially we focus on student performance in the Italian language test and then we turn our attention to their performance in math.

In Table 3 we report instrumental variable estimation results when considering as dependent variable the Italian Language Score. In the Panel B of the Table we report First Stage estimation results. Standard errors are robust to heteroscedasticity and are allowed for clustering at the class level. In all specifications we control for regional fixed effects.

The First Stage shows that Predicted Multigrade strongly determines the effective treatment, since the First Stage F-statistics is always greater than 200.

In the first two specifications of Table 3 our dependent variable is the fraction of correct answers in the Literacy test. In specification (1) we control for school and individual characteristics. We find that being placed in a multigrade class produces a negative effect on student performance in the Literacy test. The effect becomes statistically insignificant when we control for class size (column 2). Very similar results are found in specifications (3) and (4) where we replicate the first two specifications of Table 3 but consider as outcome variable the Rasch score, which also take into account the different degree of difficulty of questions. Class size does not produce any statistically significant effect.

The effects of control variables are consistent with the findings presented in the existing literature. Females tend to perform better than males. Students with a better socio-economic background obtain better results compared to students who are from more disadvantaged families. In addition, students with Italian parents perform better than students whose parents were born abroad.

Table 3. TSLS Estimates of Multigrade Classes on Literacy Scores

	(1)	(2)	(3)	(4)
	Literacy Score	*Literacy Score*	*Rasch Literacy Score*	*Rasch Literacy Score*
Panel A: Two-Stage Least Squares Estimates				
Actual Multigrade	-7.480***	-7.042	-6.787***	-6.651
	(1.990)	(4.854)	(2.513)	(6.467)
Class Size		0.047		0.015
		(0.441)		(0.608)
Female	2.340***	2.340***	7.721***	7.721***
	(0.356)	(0.356)	(0.666)	(0.666)
Pre Primary School	-0.704	-0.687	-1.678	-1.673
	(0.749)	(0.760)	(1.207)	(1.222)
Early Enrolled	-1.383	-1.357	-3.505	-3.499
	(1.733)	(1.730)	(3.185)	(3.169)
Late Enrolled	-5.880***	-5.877***	-17.309***	-17.308***
	(0.902)	(0.902)	(1.898)	(1.900)
Full Time	-0.766	-0.764	-0.482	-0.481
	(1.207)	(1.208)	(1.571)	(1.574)
ESCS index	2.323***	2.327***	9.110***	9.111***
	(0.272)	(0.274)	(0.403)	(0.408)
Italian Father	1.690***	1.681***	5.675***	5.672***
	(0.630)	(0.635)	(1.294)	(1.299)
Italian Mather	0.352	0.349	3.790***	3.789***
	(0.608)	(0.609)	(1.184)	(1.183)
Observations	16665	16665	15155	15155
Panel B: First Stage				
Predicted Multigrade	0.412***	0.303***	0.416***	0.301***
	(0.006)	(0.011)	(0.006)	(0.012)
Class Size		-0.028***		-0.030***
		(0.003)		(0.002)
First Stage F Statistics	452.934	74.2263	411.692	65.740

Notes: The Table reports TSLS estimates. Standard errors in parentheses are corrected for heteroscedasticity. In all regressions we control for regional dummies (20 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

In Table 4 we report results obtained when considering as dependent variable student cognitive skill in math. More precisely in the two first specification of the Table we consider as outcome variable the number of correct answers in numeracy test, while in specifications (3) and (4) the *Rash Numeracy Score* is considered. We find that being placed in a multigrade class reduces student score in numeracy. The effect remains also when controlling for class size: students in multigrade classes obtain a *Numeracy Scores* of about half a standard deviation lower compared to students in single grade classes. A decrease of half a standard deviation would move people who were originally at the mean, which is also about the median of the Numeracy Score, down to the third decile. Class size does not produce any statistically significant impact.

The negative effect holds true when we consider as outcome variables the Rasch score. The magnitude of the effect is of about 0.4 of a standard deviation. In order to evaluate the magnitude of the estimated effect, it is useful to consider that the effect of being assigned to a multigrade class on the *Rasch Numeracy Score* corresponds to the effect produced by an increase in *ESCS index* of about 3 standard deviations.

Qualitatively the same results are found, both for the Literacy and Numeracy test scores, when we run our regressions considering the whole sample of students undertaking the INVALSI standardized test and attending schools with no more than five classrooms (see Table 2A in the Appendix of the paper). In addition, these results are robust when controlling for a number of characteristics of the municipality in which the school is located, such as population, extension in squared kilometres, altitude (results not reported but available upon request).

We checked whether the effect of multigrade classes is heterogeneous according to student gender and economic background, but we do not find statistically significant differences.

Table 4. TSLS Estimates of Multigrade Classes on Numeracy Scores

	(1)	(2)	(3)	(4)
	Numeracy Score	*Numeracy Score*	*Rasch Numeracy Score*	*Rasch Numeracy Score*
Panel A: Two-Stage Least Squares Estimates				
Actual Multigrade1	-8.857***	-12.311***	-16.988***	-15.294**
	(1.916)	(4.716)	(3.013)	(7.574)
Class Size		-0.359		0.181
		(0.409)		(0.687)
Female	-1.821***	-1.822***	-2.464***	-2.462***
	(0.346)	(0.348)	(0.580)	(0.579)
Pre Primary School	-0.367	-0.470	3.382**	3.428**
	(0.725)	(0.740)	(1.331)	(1.348)
Early Enrolled	-3.021*	-3.199*	8.876***	8.949***
	(1.798)	(1.790)	(2.968)	(2.966)
Late Enrolled	-5.834***	-5.843***	-7.131***	-7.129***
	(0.880)	(0.886)	(1.622)	(1.619)
Full Time	1.061	1.067	-0.090	-0.091
	(1.040)	(1.057)	(1.743)	(1.741)
ESCS index	3.275***	3.239***	5.112***	5.129***
	(0.247)	(0.254)	(0.420)	(0.428)
Italian Father	2.629***	2.693***	2.393**	2.357**
	(0.695)	(0.708)	(1.184)	(1.186)
Italian Mather	1.268**	1.301**	1.453	1.448
	(0.627)	(0.637)	(1.099)	(1.097)
Observations	16665	16665	14891	14891
Panel B: First Stage				
Predicted Multigrade	0.399***	0.297***	0.406***	0.294***
	(0.006)	(0.011)	(0.007)	(0.012)
Class Size		-0.026***		-0.029***
		(0.002)		(0.002)
First Stage F Statistics	409.171	72.369	367.605	62.104

Notes: The Table reports TSLS estimates. Standard errors (reported in parentheses) are corrected for heteroscedasticity. In all the regressions we control for regional dummies (20 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

4.2. The effect of multigrade classes on grades assigned by teachers

In this section we investigate the effect of multigrade classes on student performance as assessed by the teachers through the marks assigned on the subject. In Table 5 we replicate the same specification we have estimated in the previous analysis but considering as outcome variables marks assigned by teachers in

Literacy (columns 1 and 2) and Numeracy (columns 3 and 4), respectively. We do not find any statistically significant effect of being assigned to a multigrade class on teacher assessment of students' skills. When we control for class size we find a negative effect but not statistically significant at conventional levels.

Table 5. TSLS Estimates of Multigrade Classes on grades assigned by teachers

	(1)	(2)	(3)	(4)
	Teacher Mark Literacy	*Teacher Mark Literacy*	*Teacher Mark Numeracy*	*Teacher Mark Numeracy*
Panel A: Two-Stage Least Squares Estimates				
Actual Multigrade	0.063	-0.222	0.095	-0.228
	(0.067)	(0.160)	(0.069)	(0.166)
Class Size		-0.025**		-0.030**
		(0.012)		(0.012)
Female	0.240***	0.239***	0.064***	0.064***
	(0.018)	(0.018)	(0.018)	(0.018)
Pre Primary School	-0.034	-0.043	-0.019	-0.030
	(0.028)	(0.029)	(0.029)	(0.030)
Early Enrolled	0.086	0.073	0.102*	0.086
	(0.066)	(0.067)	(0.061)	(0.062)
Late Enrolled	-0.538***	-0.537***	-0.470***	-0.470***
	(0.048)	(0.048)	(0.052)	(0.052)
Full Time	-0.039	-0.037	-0.019	-0.017
	(0.038)	(0.038)	(0.037)	(0.037)
ESCS index	0.305***	0.302***	0.311***	0.308***
	(0.010)	(0.010)	(0.010)	(0.011)
Italian Father	0.129***	0.135***	0.093**	0.100***
	(0.035)	(0.035)	(0.036)	(0.036)
Italian Mather	0.232***	0.234***	0.216***	0.217***
	(0.032)	(0.032)	(0.033)	(0.032)
Observations	15457	15457	15509	15509
Panel B: First Stage				
Predicted Multigrade	0.401***	0.297***	0.401***	0.298***
	(0.006)	(0.012)	(0.006)	(0.012)
Class Size		-0.027***		-0.027***
		(0.012)		(0.002)
First Stage F Statistics	334.745	38.012	333.421	39.568

Notes: The Table reports TSLS estimates. Standard errors (reported in parentheses) are corrected for heteroscedasticity. In all the regressions we control for regional dummies (20 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

All in all, we find that students attending multigrade classes perform significantly worse in Numeracy standardized tests, while the effect on Literacy even if negative is smaller and in some cases not statistically significant. On the other hand, we do not find any effect of being placed in a multigrade class on teachers' assessment of student skills pointing to the fact that teachers my try with more generous evaluations to compensate for the more difficult environment faced by students who interact with peers of different grades.

5. The effect of multigrade classes on non-cognitive skills

Individual success is both determined by cognitive and non-cognitive abilities (or personality traits). Locus of control has received particular attention by both psychologists and economists who have shown its relevance for social and economic success.

In order to assess the reliability of our measures of external locus of control we have tried to see whether they behave similarly to what described in the literature. One well documented result is the negative correlation between cognitive ability and external locus of control (see Cebi, 2007; Baron and Cobb-Clark, 2010). This relationship is confirmed by our data as we find a negative correlation between student performance in literacy and numeracy and our measures of external locus of control. The correlation between the Numeracy Score and External Locus of Control is –0.127 (p-value 0.000), while the correlation between the Language Score and External Locus of Control is –0.107 (p-value 0.000).

In Table 6 we report instrumental variable estimation results when considering different measures of external locus of control as dependent variable. In the first two specifications of Table 6 our dependent variable is External Locus of Control. We find that students attending a Multigrade class are more inclined to have an external locus of control. The effect is statistically significant at the 10 percent level both when we do not control for class size (column 1) and when we control for it (column 2). In columns (3) and (4) we consider as dependent variable students tendency to attribute unsuccessful situations to external factors (External Neg). We find a positive and statistically significant correlation between Actual Multigrade and External Neg, implying that students attending a multigrade class are more inclined to attribute unsuccessful situations to luck or other factors behind their own control (column 3). This result holds true also controlling for class size (column 4). On the other hand we do not find any statistically significant correlation between Actual Multigrade and student tendency to attribute successful situations to external factors (columns 5 and 6).

The sign of the effects remains the same also when considering the whole sample of students but the effects become more imprecisely estimated and statistically not significant at conventional levels.

Table 6. TSLS Estimates of Multigrade Classes on External Locus of Control

	(1)	(2)	(3)	(4)	(5)	(6)
	External Locus of Control	*External Locus of Control*	*External Locus of Control Neg*	*Externa Locus of Control Neg*	*Externa Locus of Control Pos*	*External Locus of Control Pos*
Panel A: Two-Stage Least Squares Estimates						
Actual Multigrade1	0.142*	0.348*	0.104**	0.236*	0.030	0.140
	(0.086)	(0.209)	(0.049)	(0.122)	(0.057)	(0.137)
Class Size		0.021		0.014		0.011
		(0.019)		(0.011)		(0.012)
Female	-0.321***	-0.321***	-0.163***	-0.163***	-0.159***	-0.159***
	(0.027)	(0.027)	(0.015)	(0.015)	(0.019)	(0.019)
Pre Primary School	0.088**	0.094**	0.012	0.016	0.073**	0.076**
	(0.044)	(0.044)	(0.026)	(0.026)	(0.030)	(0.030)
Early Enrolled	0.165	0.176	0.058	0.065	0.110	0.116
	(0.112)	(0.112)	(0.068)	(0.068)	(0.073)	(0.073)
Late Enrolled	0.461***	0.463***	0.236***	0.236***	0.233***	0.234***
	(0.072)	(0.072)	(0.043)	(0.043)	(0.051)	(0.051)
Full Time	0.061	0.060	0.025	0.024	0.032	0.032
	(0.048)	(0.048)	(0.027)	(0.027)	(0.031)	(0.031)
ESCS index	-0.171***	-0.169***	-0.054***	-0.053***	-0.118***	-0.117***
	(0.015)	(0.015)	(0.008)	(0.009)	(0.010)	(0.010)
Italian Father	-0.093*	-0.096*	-0.057*	-0.059**	-0.025	-0.027
	(0.054)	(0.054)	(0.030)	(0.030)	(0.037)	(0.037)
Italian Mather	-0.106**	-0.108**	-0.057**	-0.058**	-0.056*	-0.057*
	(0.048)	(0.049)	(0.027)	(0.028)	(0.034)	(0.034)
Observations	16091	16091	16353	16353	16220	16220
Panel B: First Stage						
Predicted Multigrade	0.400***	0.297***	0.401***	0.297***	0.400***	0.297***
	(0.006)	(0.011)	(0.006)	(0.011)	(0.006)	(0.011)
Class Size		0.027***		0.027***		0.027***
		(0.001)		(0.001)		(0.001)
First Stage F Statistics	404.554	71.164	408.881	71.961	405.862	71.057

Notes: The Table reports TSLS estimates. Standard errors (reported in parentheses) are corrected for heteroscedasticity. In all the regressions we control for regional dummies (20 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

6. Robustness Checks

In this section we perform a number of robustness checks. Firstly, even if we dealt cheating problems by using the "cheating-corrected" test scores, to be reassured that our results are not driven by different cheating behaviour in single and multigrade classes we only considered schools located in the Northern part of the country where cheating problems have been proved to be less important (see for instance Paccagnella and Sestito, 2014). Secondly, to this point our measure of Actual Multigrade has consisted of students attending small classes in small schools where there were at least two multigrade classrooms. We examine the robustness of our results to an alternative measure that includes students enrolled in schools where there is only one multigrade class.

In Table 7 we report estimation results for the specifications with the full set of regressors while restricting the sample to schools located in Northern regions and considering as outcome variables

alternatively the Literary and Numeracy scores (columns 1 and 2), the marks assigned by teachers (columns 3 and 4) and our indicator of external locus of control (column 5).

As regard to student performance in Literacy and Numeracy test scores we find results that are consistent with those found with the sample including the full country, even if the magnitude of the effects is slightly larger (we find very similar results also when considering as outcome variables *Rasch Literacy Score and Rasch Numeracy Score).* As shown in columns 3 and 4, when focusing on schools in the Northern part of the country, we find that Actual Multigrade produces a negative and statistically significant effect on student performance also when considering marks assigned by teachers as a measure of their cognitive skills (the effect is negative but not statistically significant when we restrict the sample to schools located in the South).

Results are also consistent with the previous section when we consider External Locus of Control as outcome variable (column 5): once again we find that being placed in a multigrade class increases students inclination to attribute success and failure to external factors.

Table 7. TSLS Estimates of Multigrade Classes on test scores, teachers' marks and locus of control. Schools located in Northern regions

	(1)	(2)	(3)	(4)	(5)
	Literacy Score	*Math Score*	*Teacher Mark Literacy*	*Teacher Mark Numeracy*	*External Locus Control*
Actual Multigrade2	-9.752	-14.845**	-0.410*	-0.432*	0.542*
	(6.425)	(5.936)	(0.236)	(0.244)	(0.286)
Class Size	0.014	-0.572	-0.048**	-0.053**	0.039
	(0.539)	(0.512)	(0.021)	(0.021)	(0.026)
Female	2.437***	-2.572***	0.229***	0.041*	-0.311***
	(0.430)	(0.425)	(0.023)	(0.025)	(0.034)
Pre Primary School	-0.623	-0.550	-0.026	-0.027	0.109**
	(0.745)	(0.813)	(0.036)	(0.038)	(0.052)
Early Enrolled	-3.571	-5.636**	-0.083	-0.046	0.195
	(2.633)	(2.586)	(0.140)	(0.119)	(0.214)
Late Enrolled	-5.943***	-6.977***	-0.481***	-0.391***	0.401***
	(1.042)	(1.069)	(0.060)	(0.067)	(0.089)
Full Time	-0.114	1.113	-0.046	-0.049	-0.022
	(1.193)	(1.123)	(0.052)	(0.049)	(0.063)
ESCS index	2.272***	3.594***	0.267***	0.279***	-0.145***
	(0.256)	(0.279)	(0.013)	(0.014)	(0.019)
Italian Father	2.681***	2.649***	0.208***	0.189***	-0.119*
	(0.739)	(0.851)	(0.046)	(0.048)	(0.071)
Italian Mather	0.880	1.576**	0.287***	0.273***	-0.166***
	(0.693)	(0.784)	(0.042)	(0.043)	(0.062)
Observations	9687	9687	9050	9084	9371
Panel B: First Stage					
Predicted Multigrade	0.273***	0.273***	0.262***	0.261***	0.273***
	(0.015)	(0.015)	(0.016)	(0.016)	(0.015)
Class Size	-0.033***	-0.033***	-0.035***	-0.035***	-0.033***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
First Stage F Statistics	27.771	27.914	33.368	33.368	33.106

Notes: The Table reports TSLS estimates. Standard errors (reported in parentheses) are corrected for heteroscedasticity. In all the regressions we control for regional dummies (20 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

We now turn to the robustness of our results to an alternative definition of Actual Multigrade. In Table 8 we report results obtained when using as indicator of students being actually placed in a multigrade class the dummy variable Actual Multigrade2. We find qualitatively similar results even if the magnitude of the effects is somehow larger pointing to the fact that when using Actual Multigrade1 some students actually attending a multigrade class were attributed to the control group.

Table 8. TSLS Estimates of Multigrade Classes on test scores, teachers' marks and locus of control. Alternative definition of Actual Multigrade

	(1)	(2)	(3)	(4)	(5)
	Literacy Score	*Math Score*	*Teacher Mark Literacy*	*Teacher Mark Numeracy*	*External Locus of Control*
Actual Multigrade2	-8.119	-13.766**	-0.260	-0.265	0.387*
	(5.658)	(5.344)	(0.187)	(0.191)	(0.237)
Class Size	-0.099	-0.644	-0.035**	-0.039**	0.029
	(0.540)	(0.510)	(0.018)	(0.018)	(0.023)
Female	2.336***	-1.828***	0.239***	0.064***	-0.320***
	(0.358)	(0.350)	(0.018)	(0.018)	(0.027)
Pre Primary School	-1.039	-1.311	-0.059*	-0.046	0.113**
	(0.857)	(0.847)	(0.033)	(0.034)	(0.048)
Early Enrolled	-1.506	-3.746**	0.060	0.073	0.186*
	(1.755)	(1.811)	(0.068)	(0.063)	(0.113)
Late Enrolled	-6.019***	-6.178***	-0.541***	-0.473***	0.468***
	(0.908)	(0.905)	(0.048)	(0.052)	(0.072)
Full Time	-0.747	1.113	-0.024	-0.002	0.056
	(1.211)	(1.054)	(0.038)	(0.037)	(0.047)
ESCS index	2.324***	3.272***	0.303***	0.308***	-0.169***
	(0.277)	(0.256)	(0.010)	(0.011)	(0.015)
Italian Father	1.708***	2.808***	0.137***	0.101***	-0.099*
	(0.642)	(0.709)	(0.035)	(0.037)	(0.054)
Italian Mather	0.305	1.190*	0.232***	0.216***	-0.105**
	(0.610)	(0.643)	(0.032)	(0.033)	(0.049)
Observations	16665	16665	15457	15509	16091
Panel B: First Stage					
Predicted Multigrade	0.264***	0.264***	0.257***	0.262***	0.265***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Class Size	-0.042***	-0.042***	-0.043***	-0.042***	-0.043***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
First Stage F Statistics	46.117	46.117	40.292	41.975	46.307

Notes: The Table reports TSLS estimates. Standard errors (reported in parentheses) are corrected for heteroscedasticity. In all the regressions we control for regional dummies (20 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

7. Concluding remarks

Multigrade classes allow schools to remain located closer to the families they serve and provide their services at a reasonable cost. However, cost-saving considerations are to be evaluated together with the

effects that multigrade classes may produce on student outcomes. In this paper we provide additional evidence of the effects of multigrade classes both on cognitive and non-cognitive skills. We exploit the discontinuous rules that regulate class composition in Italy as a source of exogenous variation in the probability of attending a multigrade class.

We find a negative effect on students' performance both in literacy and numeracy standardized test scores. The effect is particularly pronounced for numeracy test scores: students placed in multigrade classes obtain a lower score (by half a standard deviation) when compared to students in single grade classes.

Instead, a negative but not statistically insignificant effect is found when looking at marks assigned by teachers. This might depend on the fact that teachers of multigrade classes, aware of the more complex environment faced by their pupils, tend to reward their effort more generously. This is especially true for teachers working in the Southern part of the country, while for schools located in the North it emerges that multigrade classes harm students' skills also when measured by teacher assessment.

In addition, we show that placing students in multigrade classrooms causes an increase in their tendency to attribute successful and failure situations to factors behind their own control.

Since our results pertain to students attending the 5^{th} grade and who are likely to have attended multigrade classes also in previous years of their educational career (although we do not possess this piece of information), the estimated effects can be interpreted as the cumulated effects of having attended a substantial part of the primary school cycle in a multigrade environment.

This evidence suggests that the use of multigrade classes might produce a number of unintended consequences that have to be managed by school administrators. In order to define effective interventions it would be necessary to understand the mechanisms that drive these negative effects. They could be the result of excessive teachers' workload or lack of training on teacher practices within this peculiar classroom environment. It could also be that relationships among peers works in an asymmetric way, as suggested by Leveun et al. (2014): in such a case the negative effect we find could be due to having spent the final year with lower grade peers (which on the contrary could have benefited by being exposed to older peers). Understanding which of these channels slows down the learning process of students placed in multigrade classes is an interesting topic for future research.

Appendix

In Table A.1 are reported descriptive statistics for the whole sample of 5th grade students 2011-2012.

Table A.1. Descriptive statistics. Whole sample

Panel (a). Whole sample					
	Mean	St. Dev.	Min	Max	Obs.
Actual Multigrade1	0.010	0.100	0	1	466,897
Actual Multigrade2	0.019	0.136	0	1	466,897
Predicted Multigrade	0.020	0.139	0	1	473,857
Rasch Literacy Score	202.669	40.588	-35.360	368.097	450,179
Literacy Score	74.448	18.587	0	100	473,725
Teacher grade Literacy	7.623	1.148	1	10	408,339
Rasch Numeracy Score	186.196	42.464	2.281	367.182	445,513
Numeracy Score	52.710	21.229	0	100	473,749
Teacher grade Numeracy	7.701	1.189	1	10	408,185
External Locus of Control	2.202	1.641	0	8	457,148
External Locus of Control Positive	1.310	1.138	0	5	461,126
External Locus of Control Negative	0.895	0.951	0	3	465,192
Female	0.498	0.500	0	1	473,857
Regularly Enrolled	0.949	0.219	0	1	473,579
Early Enrolled	0.017	0.131	0	1	473,579
Late Enrolled	0.033	0.179	0	1	473,579
Pre Primary School	0.182	0.386	0	1	473,857
ESCS index	0.142	1.026	-3.081	2.722	452,782
Italian Father	0.805	0.396	0	1	473,857
Italian Mather	0.797	0.402	0	1	473,857
Full time	0.266	0.442	0	1	453,341
Number of student enrolled in grade 5th	20.272	4.318	1	34	473,857
Class size	20.284	4.281	1	34	473,857
Southern regions	0.384	0.486	0	1	473,857

Notes: The data are drawn from the Invalsi websites (downloaded on 7/2/2016)

In Table A.2, Panel (a), are reported descriptive statistics for the whole sample of 5th grade students attending schools that have no more than a classroom in the 5th grade. Panel (b) of Table A.2 reports descriptive statistics for the sample of students placed in single grade classes, while in Panel (c) reports descriptive statistics for students in a multigrade class (Actual Multigrade1).

Table A.2. Descriptive statistics

	Panel (a) Whole sample			Panel (b) Single Grade			Panel (c) Actual Multigrade1		
	Mean	St. Dev.	Obs.	Mean	St. Dev.	Obs.	Mean	St. Dev.	Obs.
Actual Multigrade1	0.042	0.202	105,442						
Actual Multigrade2	0.078	0.269	105,442						
Predicted Multigrade	0.058	0.233	105,442	0.030	0.169	100,963	0.695	0.461	4,479
Rasch Literacy Score	202.246	41.058	100.003	202.577	41.039	95,953	194.399	40.729	4,050
Literacy Score	74.229	18.649	105,442	74.451	18.469	100,963	69.210	21.732	4,479
Teacher Mark Literacy	7.583	1.123	95,998	7.587	1.124	91,805	7.495	1.085	4,193
Rasch Numeracy Score	182.933	42.393	98,847	183.568	42.263	94,813	168.021	42.708	4,034
Numeracy Score	52.541	21.278	105,442	52.696	21.257	100,963	49.041	21.441	4,479
Teacher Mark Numeracy	7.672	1.171	96,090	7.676	1.173	91,880	7.584	1.124	4,210
External Locus of Control	2.213	1.640	102,128	2.210	1.640	97,790	2.291	1.654	4,338
Ext Locus Control Positive	1.326	1.145	102,927	1.324	1.144	98,558	1.371	1.150	4,369
Ext.Locus Control Negative	0.891	0.948	103,759	0.890	0.947	99,351	0.924	0.973	4,408
Female	0.498	0.500	105,442	0.499	0.500	100,963	0.488	0.500	4,479
Regularly Enrolled	0.950	0.218	105,442	0.950	0.218	100,963	0.953	0.211	4,479
Early Enrolled	0.017	0.130	105,442	0.018	0.131	100,963	0.010	0.101	4,479
Late Enrolled	0.033	0.178	105,442	0.033	0.178	100,963	0.036	0.187	4,479
Pre Primary School	0.162	0.369	105,442	0.165	0.372	100,963	0.090	0.286	4,479
ESCS index	0.152	1.014	105,442	0.162	1.016	100,963	-0.053	0.927	4,479
Italian Father	0.853	0.354	105,442	0.853	0.354	100,963	0.860	0.347	4,479
Italian Mather	0.841	0.366	105,442	0.841	0.366	100,963	0.835	0.371	4,479
Full time	0.207	0.405	105,442	0.209	0.407	100,963	0.156	0.363	4,479
# student enrolled in grade 5	18.907	5.466	105,442	19.392	5.013	100,963	7.954	3.389	4,479
Class size	18.956	5.357	105,442	19.393	5.010	100,963	9.087	2.834	4,479
Southern regions	0.316	0.465	105,442	0.314	0.464	100,963	0.366	0.481	4,479

Notes: The data are drawn from the Invalsi websites (downloaded on 7/2/2016)

The effects of multigrade classes on Literacy and Numeracy Scores: estimates considering the whole sample

In this section we run our regressions considering the whole sample of 5th grade students attending schools that have no more than one classroom in the 5th grade. This choice has been made in order to reduce the measurement error in our indicator of students who are effectively attending a multigrade class (Actual Multigrade 1) and in order to limit endogeneity problems in class size.

In Table A.3 we reported the TSLS estimates of the impact of being placed in a multigrade class both on Literacy and Numeracy standardized test scores. We used the same specifications adopted in Table 3 and Table 4 of the paper. However, now we are able to control not only for class size but also for the number of students enrolled in the 5th grade. Even if this is our preferred specification, results do not change when we only control for class size (results not reported and available upon request).

Table A.3. TSLS Estimates of Multigrade Classes on Literacy and Numeracy Scores. Whole Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Literacy Score	*Literacy Score*	*Rasch Literacy Score*	*Rasch Literacy Score*	*Numeracy Score*	*Numeracy Score*	*Rasch Numeracy Score*	*Rasch Numeracy Score*
Panel A: Two-Stage Least Squares Estimates								
Actual Multigrade1	-11.337***	-7.864***	-10.904***	-6.005**	-9.372***	-7.544***	- 28.649***	-8.089**
	(1.391)	(2.327)	(1.734)	(2.857)	(1.256)	(2.149)	(2.191)	(3.605)
Class Size		-0.499		-1.172**		-0.142		-0.135
		(0.388)		(0.482)		(0.337)		(0.503)
Female	2.365***	2.355***	7.715***	7.701***	-1.963***	-1.968***	-2.745***	-2.805***
	(0.119)	(0.119)	(0.265)	(0.265)	(0.136)	(0.136)	(0.262)	(0.261)
Pre Primary School	0.360	0.331	-0.022	-0.061	0.322	0.305	3.247***	3.019***
	(0.251)	(0.252)	(0.429)	(0.430)	(0.263)	(0.263)	(0.575)	(0.566)
Early Enrolled	-4.213***	-4.199***	-6.230***	-6.221***	-5.295***	-5.287***	5.297***	5.357***
	(0.870)	(0.870)	(1.249)	(1.249)	(0.867)	(0.868)	(1.328)	(1.302)
Late Enrolled	-6.554***	-6.480***	-18.214***	-18.123***	-5.476***	-5.434***	-7.663***	-7.190***
	(0.419)	(0.417)	(0.858)	(0.860)	(0.418)	(0.418)	(0.824)	(0.822)
Full Time	-0.382	-0.438	-0.527	-0.592	0.901**	0.866**	0.810	0.336
	(0.437)	(0.439)	(0.636)	(0.640)	(0.431)	(0.432)	(0.963)	(0.957)
ESCS index	2.482***	2.443***	8.995***	8.947***	3.409***	3.386***	6.600***	6.292***
	(0.106)	(0.104)	(0.177)	(0.177)	(0.109)	(0.108)	(0.217)	(0.210)
Italian Father	2.287***	2.276***	5.910***	5.885***	2.383***	2.378***	2.531***	2.466***
	(0.270)	(0.270)	(0.517)	(0.516)	(0.290)	(0.290)	(0.582)	(0.573)
Italian Mather	1.338***	1.338***	5.086***	5.085***	1.699***	1.697***	3.418***	3.380***
	(0.250)	(0.250)	(0.499)	(0.499)	(0.271)	(0.270)	(0.543)	(0.536)
Students Enrolled 5th Grade		0.611		1.307***		0.208		0.977*
		(0.409)		(0.507)		(0.357)		(0.547)
Observations	105442	105442	100003	100003	105442	105442	98847	98847
Panel B: First Stage								
Predicted Multigrade	0.481***	0.374***	0.485***	0.370***	0.481***	0.362***	0.489***	0.374***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Class Size		0.097***		0.097***		0.097***		0.093***
		(0.000)		(0.000)		(0.000)		(0.000)
Students Enrolled 5th Grade		-0.101***		-0.102***		-0.102***		-0.097***
		(0.000)		(0.000)		(0.000)		(0.000)
First Stage F Statistics	895.283	415.084	803.269	382.63	895.283	415.084	800.267	380.972

Notes: The Table reports TSLS estimates. Standard errors (reported in parentheses) are corrected for heteroscedasticity. In all the regressions we control for regional dummies (20 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

Table **A.4.** reports the TSLS estimates of the impact of being placed in a multigrade class both on student achievement in Literacy and Numeracy as measured by marks assigned by teachers. We estimate the same specifications presented in Table 5 of the paper. However, we are now able to control not only for class size but also for the number of students enrolled in the 5th grade. Even if this is our preferred specification, results do not change when we only control for class size (results not reported and available upon request).

Table A.4. TSLS Estimates of Multigrade Classes on grades assigned by teachers

	(1)	(2)	(3)	(4)
	Teacher Mark Literacy	*Teacher Mark Literacy*	*Teacher Mark Numeracy*	*Teacher Mark Numeracy*
Panel A: Two-Stage Least Squares Estimates				
Actual Multigrade1	-0.020	0.183**	0.007	0.254***
	(0.043)	(0.075)	(0.044)	(0.077)
Class Size		-0.009		-0.016
		(0.012)		(0.012)
Female	0.258***	0.257***	0.059***	0.059***
	(0.007)	(0.007)	(0.008)	(0.008)
Pre Primary School	-0.051***	-0.053***	-0.031***	-0.033***
	(0.011)	(0.011)	(0.012)	(0.012)
Early Enrolled	0.061**	0.062**	0.073**	0.074**
	(0.029)	(0.029)	(0.031)	(0.031)
Late Enrolled	-0.471***	-0.466***	-0.403***	-0.397***
	(0.024)	(0.024)	(0.027)	(0.027)
Full Time	-0.020	-0.023	-0.015	-0.019
	(0.016)	(0.016)	(0.016)	(0.016)
ESCS index	0.315***	0.313***	0.320***	0.317***
	(0.004)	(0.004)	(0.005)	(0.005)
Italian Father	0.181***	0.180***	0.173***	0.172***
	(0.015)	(0.015)	(0.016)	(0.016)
Italian Mather	0.256***	0.256***	0.206***	0.206***
	(0.014)	(0.014)	(0.015)	(0.015)
Students Enrolled 5th Grade		0.017		0.025
		(0.012)		(0.016)
Observations	95998	95998	96090	96090
Panel B: First Stage				
Predicted Multigrade	0.492***	0.366***	0.493***	0.366***
	(0.017)	(0.018)	(0.017)	(0.018)
Class Size		0.097***		0.097***
		(0.003)		(0.003)
Students Enrolled 5th Grade		-0.102***		-0.102***
		(0.003)		(0.003)
First Stage F Statistics	877.901	390.494	877.817	389.313

Notes: The Table reports TSLS estimates. Standard errors (reported in parentheses) are corrected for heteroscedasticity. In all the regressions we control for regional dummies (20 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

In Table **A.5** we report the TSLS estimates of the impact of being placed in a multigrade class on the proposed measures of external locus of control. We replicate the same specification introduced in Table **6** of the main text. However, now we are able to control not only for class size but also for the exact number of students enrolled in the 5th grade. Even if this is our preferred specification, results do not change when we only control for class size (results not reported and available upon request).

Table A.5. TSLS Estimates of Multigrade Classes on External Locus of Control. Whole sample

	(1)	(2)	(3)	(4)	(5)	(6)
	External	*External*	*External Neg*	*Externa Neg*	*Externa Pos*	*External Pos*
Panel A: Two-Stage Least Squares Estimates						
Actual Multigrade2	0.094*	0.014	0.072**	0.060	0.017	-0.055
	(0.055)	(0.091)	(0.032)	(0.052)	(0.036)	(0.060)
Class Size		0.010		-0.011		0.022**
		(0.016)		(0.009)		(0.011)
Female	-0.344***	-0.344***	-0.180***	-0.180***	-0.164***	-0.164***
	(0.011)	(0.011)	(0.006)	(0.006)	(0.007)	(0.007)
Pre Primary School	0.032**	0.033**	0.005	0.006	0.026**	0.026***
	(0.015)	(0.015)	(0.008)	(0.008)	(0.010)	(0.010)
Early Enrolled	0.148***	0.148***	0.054**	0.054**	0.098***	0.098***
	(0.042)	(0.042)	(0.025)	(0.025)	(0.028)	(0.028)
Late Enrolled	0.404***	0.402***	0.208***	0.207***	0.194***	0.193***
	(0.033)	(0.033)	(0.019)	(0.019)	(0.023)	(0.023)
Full Time	-0.041**	-0.040**	-0.027***	-0.027***	-0.014	-0.013
	(0.017)	(0.017)	(0.010)	(0.010)	(0.011)	(0.011)
ESCS index	-0.161***	-0.160***	-0.049***	-0.049***	-0.113***	-0.112***
	(0.006)	(0.006)	(0.003)	(0.003)	(0.004)	(0.004)
Italian Father	-0.126***	-0.125***	-0.064***	-0.064***	-0.061***	-0.060***
	(0.021)	(0.021)	(0.012)	(0.012)	(0.015)	(0.015)
Italian Mather	-0.154***	-0.154***	-0.055***	-0.055***	-0.101***	-0.101***
	(0.020)	(0.020)	(0.011)	(0.011)	(0.014)	(0.014)
Students Enrolled 5th Grade		-0.013		0.010		-0.024**
		(0.016)		(0.009)		(0.011)
Observations	102,128	102,128	103,759	103,759	102,927	102,927
Panel B: First Stage						
Predicted Multigrade	0.482***	0.362***	0.482***	0.362***	0.482***	0.362***
	(0.002)	(0.014)	(0.002)	(0.014)	(0.002)	(0.014)
Class Size		0.097***		0.097***		0.097***
		(0.000)		(0.000)		(0.000)
Students Enrolled 5th Grade		-0.101***		-0.101***		-0.101***
		(0.001)		(0.001)		(0.001)
First Stage F Statistics	889.931	410.655	894.419	413.447	891.059	411.88

Notes: The Table reports TSLS estimates. Standard errors (reported in parentheses) are corrected for heteroscedasticity. In all the regressions we control for regional dummies (20 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

References

Almlund, M., A. Duckworth, J. Heckman and T. Kautz (2011), Personality psychology and economics. IZA Discussion Paper 5500.

Bertoni M., G. Brunello, L. Rocco (2013), When the cat is near, the mice won't play: The effect of external examiners in Italian schools, *Journal of Public Economics*,104, pp. 65-77

Caliendo M. D. Cobb-Clark and A. Uhlendorff (2015), Locus of Control and Job Search Strategies, *Review of Economics and Statistics*, 97(1), pp.88-103

Cebi M. (2007), Locus of Control and Human Capital Investment Revisited, *Journal of Human Resources*, 4, pp. 919-932.

Chiteji N. (2010), Time-preference, non-cognitive skills and well-being across the life course: Do non-cognitive skills encourage healthy behavior? *American Economic Review*, 100(2), pp. 200–204

Coleman M. and DeLeire T. (2003), An Economic Model of Locus of Control and the Human Capital Investment Decision, *Journal of Human Resources*, 38(3), pp. 701-721.

Dardanoni, V., S. Modica and A. Pennisi (2009). Grading across schools. *The B.E. Journal of Economic Analysis & Policy*, 9(1), Article 16.

Gatz M. and M. Karel (1993), Individual Change in Perceived Control over 20 Years, *International Journal of Behavioral Development* 16(2), pp.305-322

Gerhardts, I., U. Sunde and L. Zierow (2016), Denominational Schools and Returns to Education, Gender Socialization in Multigrade Classrooms? Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Schooling and Labor Supply, No. A14-V1

Groves M. (2005), How important is your personality? Labor market returns to personality for women in the US and UK, *Journal of Economic Psychology*, 26(6), pp. 827-841

Hoxby, C.M. (2000). 'Peer effects in the classroom: learning from gender and race variation', Working Paper 7867, NBER.

Johnsson E. and E. Lindahl (2011). The effects of mixed-age classes in Sweden, *Scandinavian Journal of Educational Research*, 55(2), pp.121-144.

Leuven, E and M. Rønning (2014). Classroom Grade Composition and Pupil Achievement. *Economic Journal*, 126(593), pp.1164-1192

Mason, D.A. and R. Burns (1997). 'Reassessing the effects of combination classes', *Educational Research and Evaluation*, vol. 3(1), pp. 1–53.

Mulkeen A. and C. Higgins (2009) Multigrade Teaching in Sub-Saharan Africa: Lessons from Uganda, Senegal, and the Gambia, World Bank Publications.

Ricci R. (2010). The Economic, Social, and Cultural Background: a continuous index for the Italian Students of the fifth grade, mimeo

Sattari R. (2016). The Effect of Multigrade Classrooms on Student Behavior, mimeo.

Schurer, S. (2011), Does attitude matter? Health, labor supply, and locus of control. mimeo.

Thomas, J. L. (2012), Combination classes and educational achievement, *Economics of Education Review* 31(6), pp.1058-1066.

Veenman S. (1995). Cognitive and Noncognitive Effects of Multigrade and Multi-Age Classes: A Best-Evidence Synthesis, *Review of Educational Research*, 65(4), pp. 319-381