The Causal Effect of Class Size on Pupils' Performance: Evidence from Italian Primary Schools

Larysa Minzyuk [*]and Felice Russo [†‡]

Department of Economics, Mathematics and Statistics. University of Salento.

Abstract

In this study, we attempt to estimate the causal effect of class size on pupils' performance using the data from Italian public primary schools. Since in 2008/09, primary schools in Italy were subject to the maximum class size of 25 pupils, we observe a discontinuous relation between class size and enrolment that we exploit to estimate the class-size effect on pupils' performance in the context of regression discontinuity design (RD). Using a standard fuzzy RD model, we focus on the pupils whose schools are in the small intervals around cutoff values of enrolment. Though we do not find a clear evidence supporting class-size reduction policy, we observe that class size is largely used in Italian public primary schools as a kind of compensatory policy.

JEL classification: J21, J24, H52

Key words: pupils' performance; class size; regression discontinuity design; sorting; Italy.

[*]E-mail address: larissa.minzuk@unisalento.it

[†]E-mail address: felice.russo@unisalento.it

[‡]Special thanks to Gianna Barbieri from the Statistic Office of MIUR and to Patrizia Falzetti from INVALSI for assistance in realization of this study. The usual disclaimer applies.

1 Introduction

The relationship between class-size and education attainment in economic educational research has been widely explored. A number of estimation techniques has been used in order to study the causal effect of class-size on student outcomes, but the existing empirical evidence on the class-size effect is still contrasting and somewhat inconclusive. However, there is a general agreement that practioners should firstly follow an identification strategy able to address potential endogenous selection of students between classes and schools.

In response to the critics of cross-section estimation strategy, in the last decade quasi-experimental approaches dominated the economic research on the class-size effect.[1] Starting with Angrist and Lavy (1999), numerous studies, mostly based on the regression discontinuity (here after, RD) design, originally proposed by Thistlewaite and Campbell (1960), have been conducted using the data from schooling systems in different countries (among others, Hoxby, 2000; Dobbelsteen *et al.*, 2002; Woessman, 2005; Urquiola, 2006; Leuven *et al.*, 2008). Recently, Urquiola and Verhoogen (2009) have analyzed some critical points about the use of RD design when studying the class-size effect. Exploiting the data from Chilean schools, they find that schools, subject to the exogenous maximum-class-size rule and staying in the regression discontinuity intervals, might exercise selection policy on enrolment that generates discontinuities in the relation between enrolment and pupils/schools characteristics other than between enrolment and class size, which are, respectively, "forcing" and "treatment" variables of RD design. Since this condition undermines strongly the validity of this estimation strategy, as in case of a randomized experiment,[2] the authors suggest caution when using RD, especially in application to private schools that may have a better control over enrolment compared to public schools. More recently, Zada *et al.* (2009) have found

[1]Hanushek (1997, 2003), after reviewing a great number of estimates, has concluded that there is no significant evidence about the existence of a positive effect of smaller classes on pupils' achievement. To this, Krueger (2003; see also Hedges *et al.*, 1994) have argued that many studies, rejecting the existence of relation between class size and pupils performance, have been mostly based on cross-section investigations, which are known to be subject to multiple sources of bias (on potential sources of bias see the discussion in Hoxby, 2000).

[2]Lee, 2008; see also Hahn *et al.*, 2001. The presence of jumps in observed family characteristics at the discontinuity point (*i.e.* these characteristics are not balanced to the left and right of the cut-off point) could suggest non-random allocation of pupils around that point.

an evidence of “selection” policy in public schools in Israel as well. Precisely, they observe discontinuities between characteristics of pupils’ households and enrolment only in public secondary schools.

In this paper, we estimate the class size effect in Italian public primary schools, which are now engaged in an important reform movement.[3] A limited research has been done on this issue so far. Among the known studies, Bratti *et al*. (2007) identify the factors that influence education attainment of Italian students in secondary schools (“superiori” schools) in PISA test. They use class size as a control variable that reveals statistically not significant. However, they do not control for the class size that 15 years old students in the first grade of superiori had during five years of primary and three years of media school, though the educational research indicates that class size effect may be particularly important in early schooling producing then a cumulated effect (Nye *et al*., 2000). Brunello and Checchi (2005), using the data of cohorts born in Italy before 1970, find that the lower pupil-teacher ratio calculated at regional level is positively correlated with higher educational attainment and its effect is stronger for those individuals who were born in the regions and cohorts with poorer family background.

Research on the class size effect in Italian primary schools has been previously hampered by data limitations.[4] To our best knowledge, there is no research on this issue using INVALSI data. With this study, we want to fill this gap. We conduct our study using the outcomes of INVALSI test of V grade pupils in 2008/09. Since Italian primary schools in 2008/09 were subject to the maximum-class-size rule of 25 pupils, there exists a strong and discontinuous relation between class size and grade enrolment. We estimate the relation between test results and class size using a *fuzzy* RD model focusing on the pupils whose schools are in the small intervals around cut-off values of enrolment (multiples of 25). Though we do not find any clear evidence which would support

[3]The reform bill, approved in 2008, was motivated by a number of reasons. Namely Italian students did badly in international comparisons as PISA test, PIRLS and TIMSS, though they studied for longer and in smaller classes (see Cipollone and Sestito, 2007). Class size and pupil-teacher ratio exhibited important variations across regions (see Bratti *et al*., 2007). Biagi and Fontana (2008) identify the factors that determine significant variations in class size across Italian regions calling for rationalization of public spending.

[4]Only starting in 2004/05, INVALSI agency, independent body in Italy charged with conducting standard testing, provides tests results of pupils in II and V grades in primary schools.

class-size reduction policy, we observe class size has been largely used as a kind of compensatory policy between Italian public primary schools.

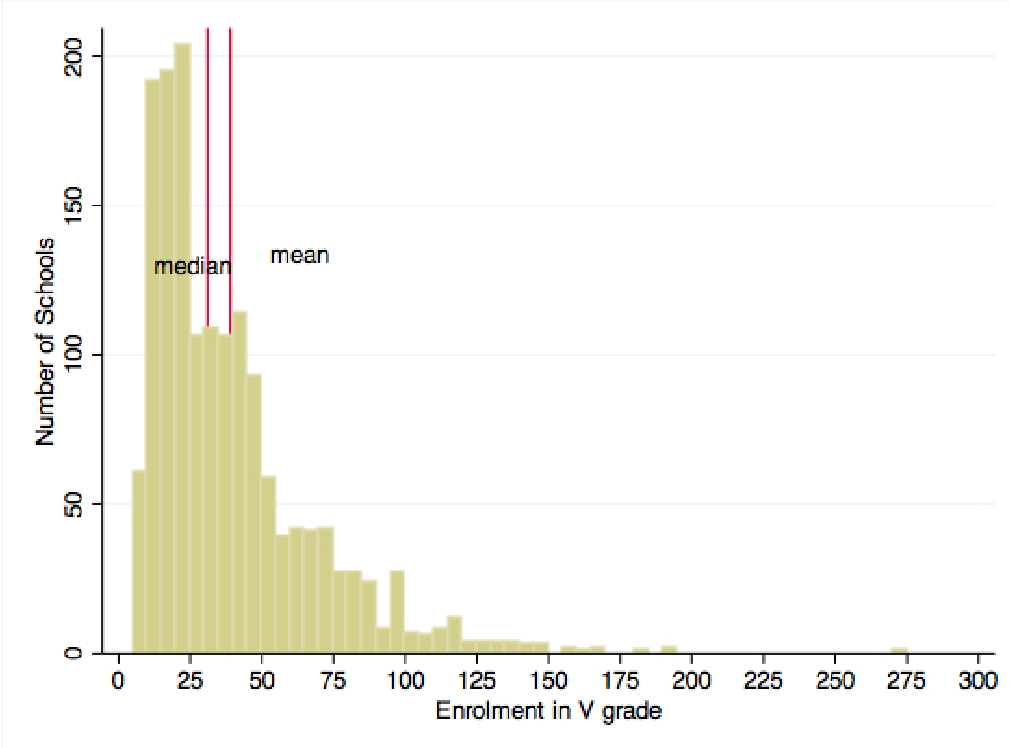
The remainder of the paper is organized as follows. The next section describes the institutional setting of public primary schools in Italy. It is followed by the description of data in Section 3. Section 4 presents the RD design and reports our main results and Section 5 concludes.

2 Italian setting

Primary schools in Italy provide education to pupils from 6 to 10 years old. They are mostly public, free of charge, and the only criteria to be used in deciding admission of a pupil is age. Parents have a right of admission of their children in neighborhood public school, but they can choose another public school unless it is oversubscribed. Alternatively, they may send their children to private schools, which in 2008/09 accounted for about 9.8 percent of all primary schools in Italy.

Though public schools are designed and financed centrally, and the assignment of teachers to them occurs through the central administrative mechanism based upon seniority rules, they exhibit important variations in terms of class size.[5] Table 1 (see in the Appendix) reports the regional distribution of average class size in V grade in primary public schools in 2008/09. As we see, class size varies noticeably across regions (from 15 to 20 pupils per class), the average class at national level is relatively small including about 18 pupils (Column 1). Our dataset includes schools whose average V grade class size is 19.41 (see Column 2). Figure 1 below illustrates the distribution of schools in our dataset according to their enrolment in V grade. The majority of schools is of small dimension: the median enrolment in V grade is around 30 pupils, whereas its mean value is around 40 pupils. More than fifty percent of schools run two classes in V grade (see Figure 2). These facts are coherent with the national data pattern according to which primary public schools in Italy are small and they run relatively few classes per grade.[6]

Figure 1: Number of schools according to enrolment, 2008/09



3 Data

We conduct our study using information from two sources. The first is the INVALSI test results of V grade pupils in primary schools in 2008/09. These results are available for 150,000 pupils coming from 5,303 public and private primary schools ("circoli didattici")[7] from all Italian regions. We restrict our analysis to the public schools whose testing procedure was assisted by INVALSI supervisors. The second source of information is school-level administrative data from the Italian Ministry of Education (MIUR).

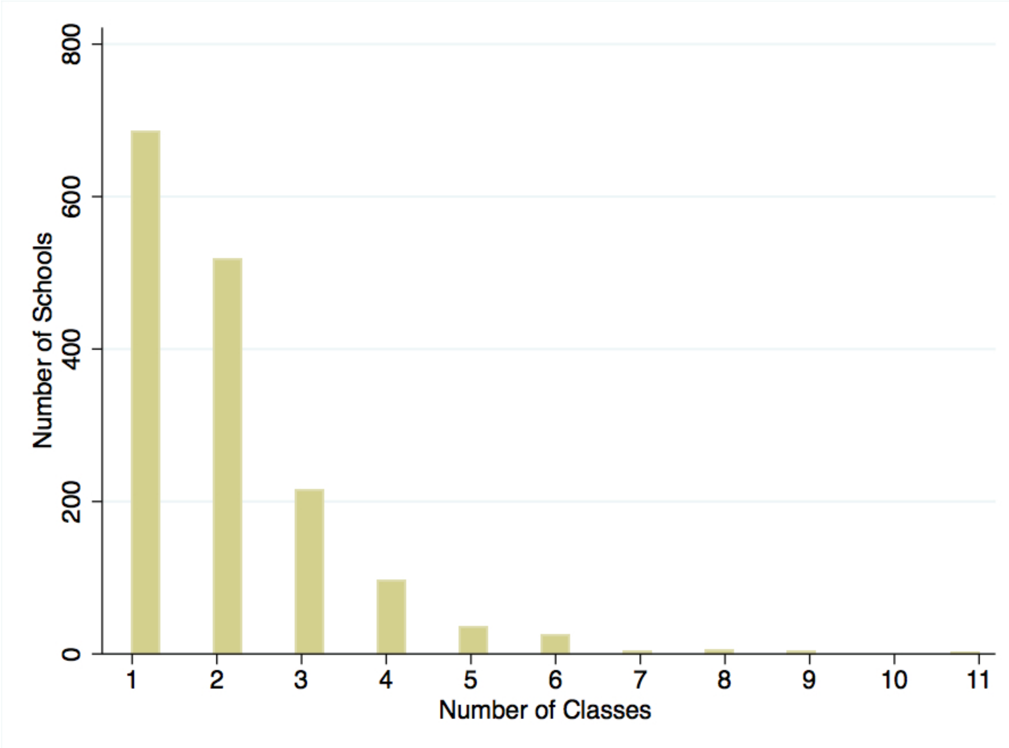
Our analysis begins by matching the sample of INVALSI data with the dataset on school characteristics and class size coming from MIUR. Since the Ministerial data do not contain information about schools in regions with a special statute as "Val d'Aosta" and "Trentino Alto Adige", we lose data about their pupils. In addition, the linked data set decreases because of many missing data

[5]See Biagi and Fontana (2008).

[6]When we say schools, we intend school units ("plessi").

[7]"Circolo didattico" is an administrative aggregation of school units ("plessi") having the same administrative body and situated in the same town. "Circolo didattico" may include "plessi" from more than one neighborhood little towns.

Figure 2: Distribution of schools according to the number of classes in V grade, 2008/09



about pupils characteristics in INVALSI data. At the end, we have 25,407 pupils coming from 1,561 school units ("plessi"), which do not include combine schools, schools for adults, schools run in hospitals and other unusual places.

In our analysis, the unit of observation is pupil. Table 2 presents descriptive statistics for the main variables (personal information, pupils families' characteristics, etc.) in our data. All information used in this study pertains to the same year of the INVALSI test. Since we are concerned with estimating class-size effect on pupils performance, our most important dependent variable is class size. We calculate its average level in V grade at school level.[8] [9]

[8]In Italian primary public schools, the number of classes organized at the moment of subscription remains almost unchanged during all five years due to a particular mechanism through which teachers are assigned to schools. As a result, even if the number of pupils in single classes at grade level may change due to new pupils or pupils moving to other schools, the average grade class size does not vary noticeably. In addition, in Italian primary schools, grade retention almost never occurs. Thus, we may conclude that most of pupils in our data have attended school for the same period of time and have being exposed to the same average grade class size.

[9]In our data, full-time classes are larger than normal-time classes. Their average class sizes are, respectively, 20.33 and 19.26. This result reflects the increasing demand of families for full-time classes.

4 Regression Discontinuity Design and Estimations

It is well known that class size is a potentially endogenous variable because it may be influenced by decisions of schools and/or parents. To examine this prediction in our data we regress class size on number of selected observables describing pupils, schools and towns including regional dummies. Table 3 reports the results. As we see, such variables as pupils' father ISEI index, non-working mother and cohabiting status result strongly correlated with class size in the first specification in Column (1). The statistical significance of these variables, however, decreases or disappears at all as we add to the estimation equation town characteristics and regional dummies (see Columns 2,3). This suggests that the selection problem related with class size may be a result of the aggregation of schools across towns and regions, rather than of behavioral pattern of schools and/or parents.

Until 2008/09, primary schools in Italy were subject to the maximum class-size of 25 pupils, whose application consisted in adding a class each time that grade enrolment exceeded the multiples of 25 pupils.[10] [11] Given the application of the rule, we expect that class size is discontinuously related with grade enrolment in our data. In order to examine it, we calculate the predicted class size in V grade, which is the average class size that schools would have in V grade if they had applied the rule perfectly. The predicted class size is calculated by applying the formula as in Angrist and Lavy (1999):

$$PC_{ist} = \frac{\Phi_{ist}}{int[(\Phi_{ist}-1)/25]+1}. \tag{1}$$

In the above equation Φ_{ist} is the V grade enrolment at school s in town t where the pupil i studied in 2008/09, $int(\bullet)$ is the function that takes the greatest integer less than the given argument.

[10] Actually, the rule was applied in I grade at the moment of subscription because, as it has been already mentioned, in Italian public primary schools the number of classes organized in I grade remains almost unchanged during all five years of elementary school due to the particular procedure through which teachers are assigned to schools.

[11] Schools could deviate from the maximum-class rule in a number of cases. Namely, they could reduce the number of pupils in classes under 25 when classes had pupil/s with special educational needs; when schools were situated in small islands, mounting towns, towns with linguistic minorities and in zones with high rates of adolescent deviant behavior. Finally, the Ministry of Education provided that schools might have smaller and larger than 25 pupils classes depending on the number of tenured teachers. For more on this, see Decreto Ministeriale 24 luglio 1998, n.331. The school reform proposed by Gelmini in 2008 elevated the maximum size of class in primary schools to 28 pupils, starting in the schooling year 2009/2010. In 2009 the reform was approved.

Figure 3: Average and Predicted Class Size (1), 2008/09

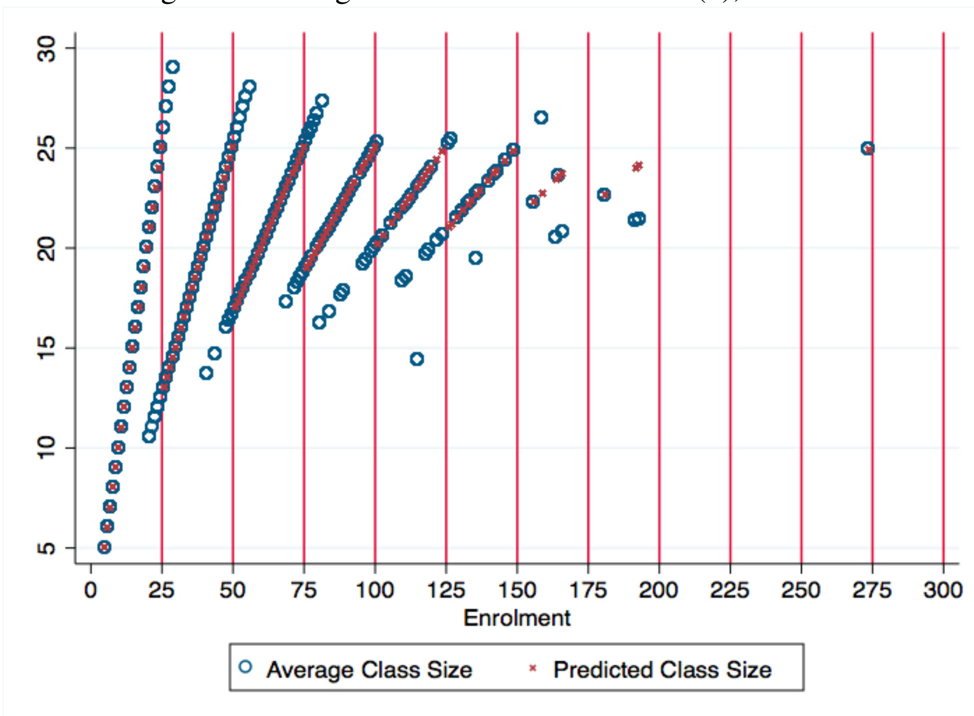
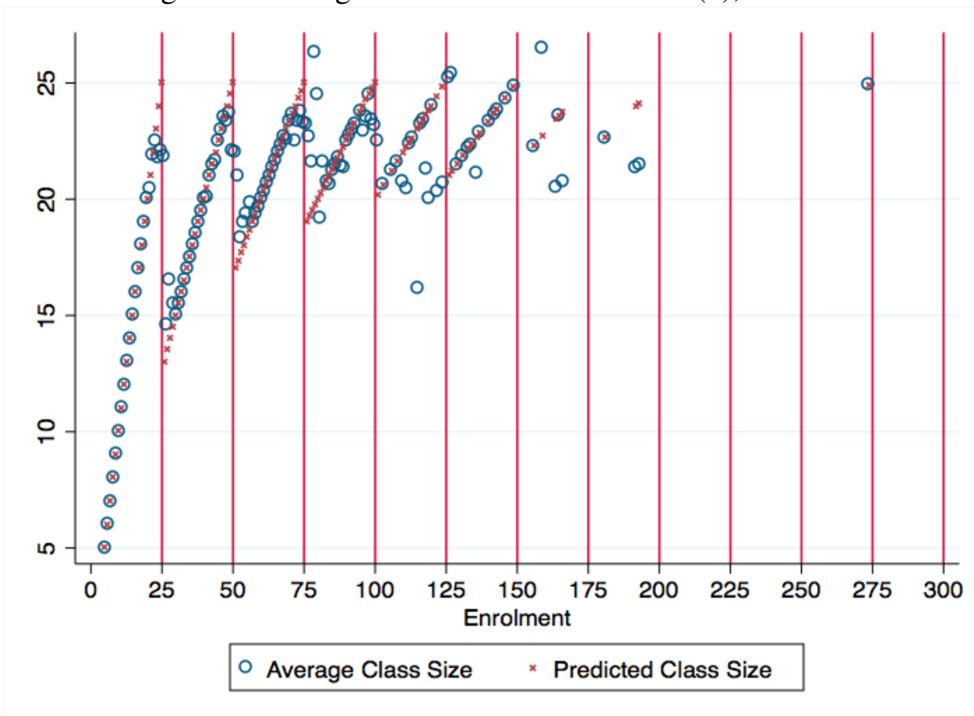


Figure 3 plots the relation between average class size, its predicted level PC_{ist} and enrolment Φ_{ist} at *school* level. Looking at the Figure 3, we observe first that the rule explains a great part of variation in class size in our data. Second, there exists a strong and discontinuous relation between class size and enrolment. Precisely, the average V grade class size exhibits discontinuous jumps around cut-off values (multiples of 25) as it is predicted by the rule.

Since we know the rule explaining class organization in primary schools in 2008/09, and since in our data class size decreases discontinuously at the cut-off values of enrolment, we may use regression discontinuity design to estimate causal relation between class size and pupils performance. Specifically, if schools and parents have imperfect control over the forcing variable (enrolment), the "treatment" effect on pupils performance through variation in class size can be then estimated by confronting test results of pupils in large and small classes on left and right sides of cut-offs.

Regression discontinuity design provides two estimation strategies: *sharp* and *fuzzy*. In a *sharp* RD design, the treatment assignment should be a result of a deterministic function of forcing variable. Opposed to it, *fuzzy* RD does not require that the probability of receiving the treatment

Figure 4: Average and Predicted Class Size (2), 2008/09



Note: The circles scatter plots mean class size for each enrolment cell. The x plot represents the predicted class size that describes the relationship between enrolment and class size that would exist if the maximum class-size rule were applied perfectly.

necessarily changes from 0 to 1 at the cut-off/s, *i.e.* a case of imperfect compliance (Trochim, 1984; Hahn *et al.*, 2001).

Findings from Figure 4, where now circles plots mean values of average class size in each enrolment cell, indicates that actual class size does not obey its predicted level perfectly, yet since the probability of reducing class size contains jumps at cut-offs points, we may adopt a *fuzzy* RD design. A standard model of *fuzzy* RD can be described as follows (van der Klaauw, 2002):

$$P_{ist} \quad = \quad \delta E(CS_{ist}|\Phi_{ist}) + \alpha(\Phi_{ist}) + \varepsilon_i \tag{2}$$

$$E(CS_{ist}|\Phi_{ist}) = \lambda 1(\Phi_{ist} \geq \bar{\Phi}) + \beta(\Phi_{ist}) \tag{3}$$

In the above equations P_{its} is the test score of pupil's i in school s in town t, CS_{ist} is average class size in V grade at school level, Φ_{ist} is the V grade enrolment at school level, $\bar{\Phi}$ indicates the cut-off values of enrolment (multiples of 25), and $\alpha(\bullet)$ and $\beta(\bullet)$ are functions of enrolment since we assume that test results may depend on enrolment other than through class size. Being enrolment a discrete variable, class-size effect can be estimated only parametrically (Lee and Card, 2008). We decide a linear specification for $\alpha(\bullet)$ and $\beta(\bullet)$ choosing the piecewise linear splines whose *kinks* correspond to the values of cut-offs (Urquiola and Verhoogen, 2009, Zada *et al.*, 2009; see also McEwan and Urquiola, 2005).

4.1 Results (Full sample)

Since in RD design the forcing variable, Φ_{ist}, should be related smoothly to pupils and schools characteristics, we start our analysis by looking at the distribution of some selected observables along the cut-off values of enrolment.

Table 4 reports the results of the regressions of selected pupil's characteristics on dummies for cut-offs ($1\Phi \geq 26$, $1\Phi \geq 51$, $1\Phi \geq 76$, $1\Phi \geq 101$, $1\Phi \geq 125$, $1\Phi \geq 150$, $1\Phi \geq 175$) and on linear splines of enrolment. The coefficients on the dummies are the estimates of the average variations in the values of dependent variables at those breaks. As we see, at the first cut-off ($\Phi \geq 26$), all selected variables - ISEI indexes of pupils' mothers and fathers, their mothers' education and the average number of pupils with special educational needs - decrease in their average values. The average decline in ISEI index of pupils' mother is particularly significant and equals 4.3 (the variable's range of variation is 0-68). In addition, the decline in the value of this variable is confirmed by the negative and statistically significant coefficient on the linear spline corresponding to the first cut-off, $(\Phi - 25)1\{\Phi \geq 26\}$. All selected variables in Table 4 result positively related to grade enrolment, which is, basically, a measure of school dimension. This result does not necessarily mean that families with "better" educational and economic status prefer bigger schools, rather it may be a result of geographical aggregation effect, whereas in bigger cities with bigger schools economical and educations status of people generally gets better. Surprisingly, we do not find an evidence that the number of pupils with special educational needs increases in smaller classes though it is explicitly provided by the Ministerial rules. The coefficient on dummies for this variable is significant only on the cut-offs corresponding to 25 and 125 pupils, where they are uniformly negative indicating that the number of such pupils seems to be higher in schools with larger average class size.

Figures 5, 6 and 7 provide an additional evidence of non-smooth distribution of selected variables along enrolment cut-offs. They illustrate the results of the local regressions of ISEI indexes of pupils' father and mother, and of education of pupils' mother in the following intervals: [0,25]; [25,50]; [50,75]; [75,100]; [100,125]; [125,150]; [150,175]; [175;200]. Findings confirm the uniform "selection" pattern along the first four enrolment cut-offs (25, 50, 75 and 100). Specifically, ISEI index of pupils' father and mother, and education of pupils' mother decline at the majority of breaks of the first four cut-off intervals, which include the large majority of observations in our data. The remaining intervals include few schools and present extreme variation in values of selected variables.

Figure 5: Isei index of pupil's father along enrolment (full sample), 2008/09

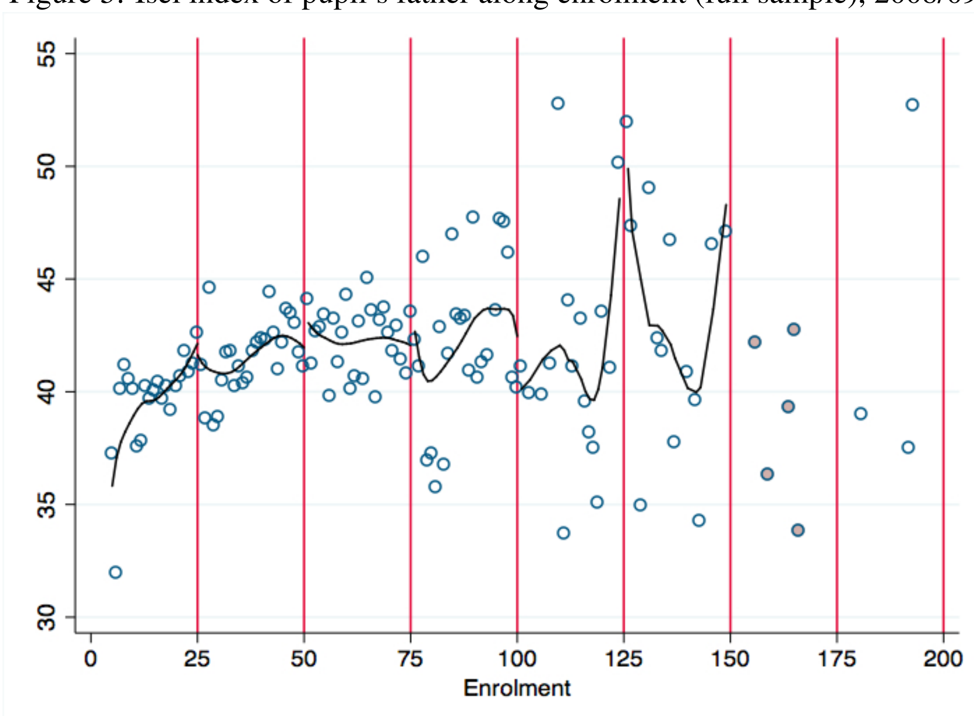


Figure 6: Isei index of pupil's mother along enrolment (ful sample), 2008/09

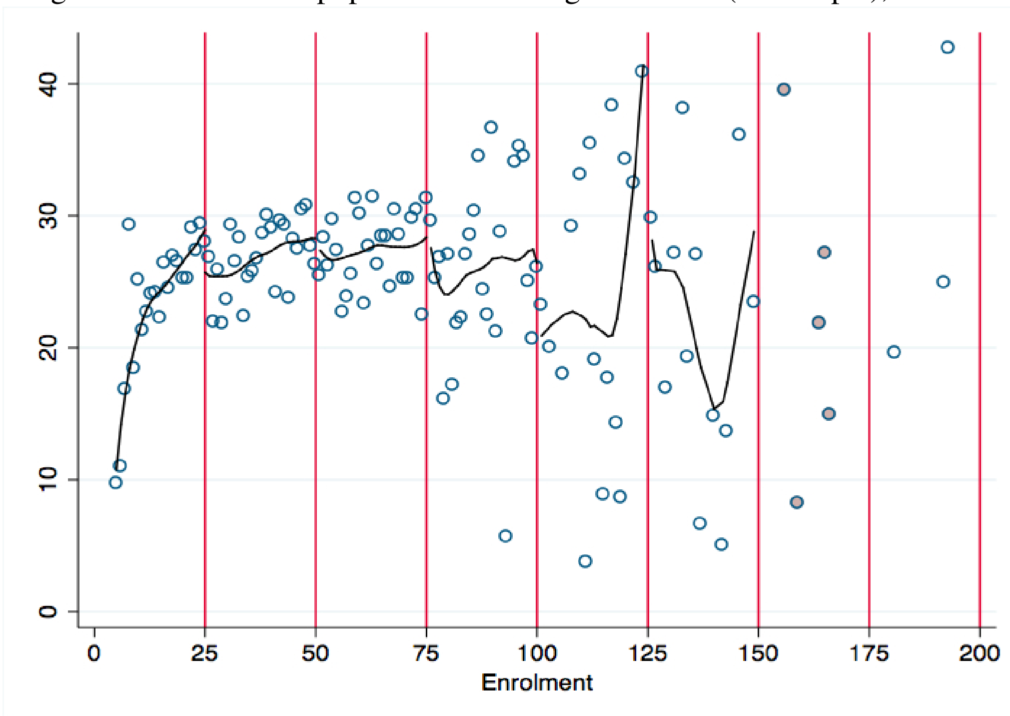
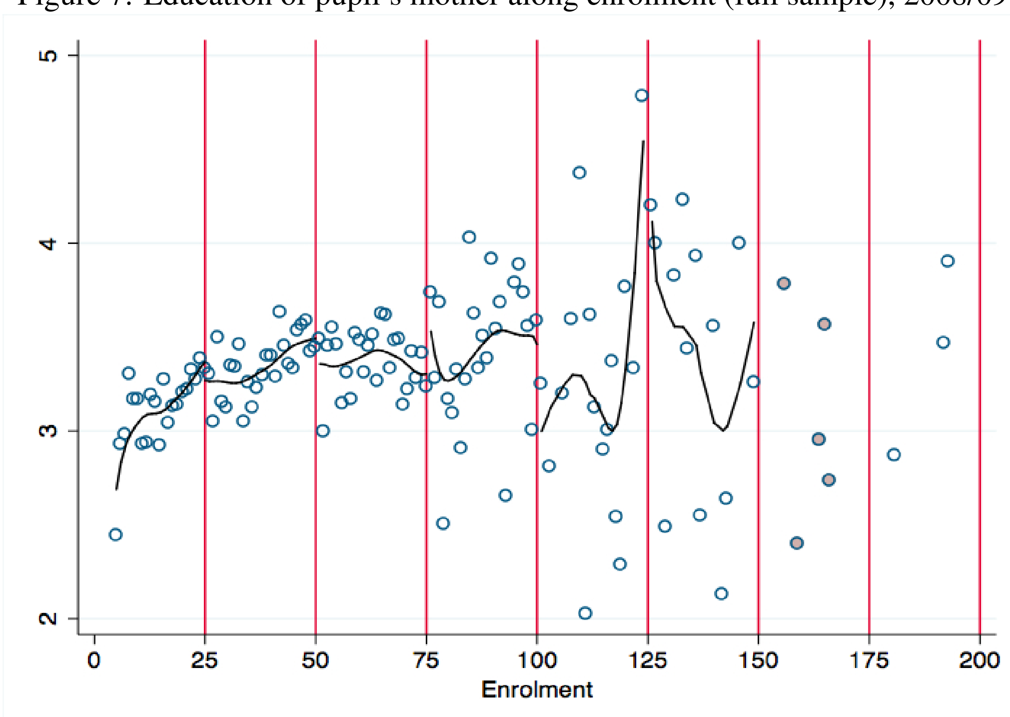
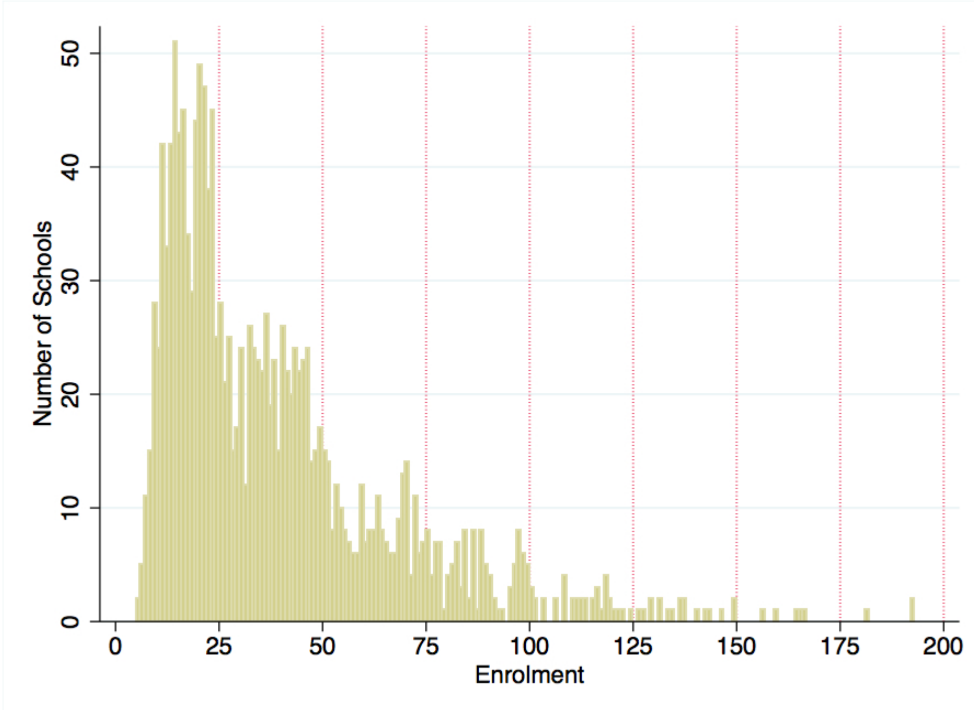


Figure 7: Education of pupil's mother along enrolment (full sample), 2008/09



In addition to the assumption of smooth distribution of pupils and schools characteristics at cut-

Figure 8: Number of schools according to their enrolment in V grade (full sample), 2008/09

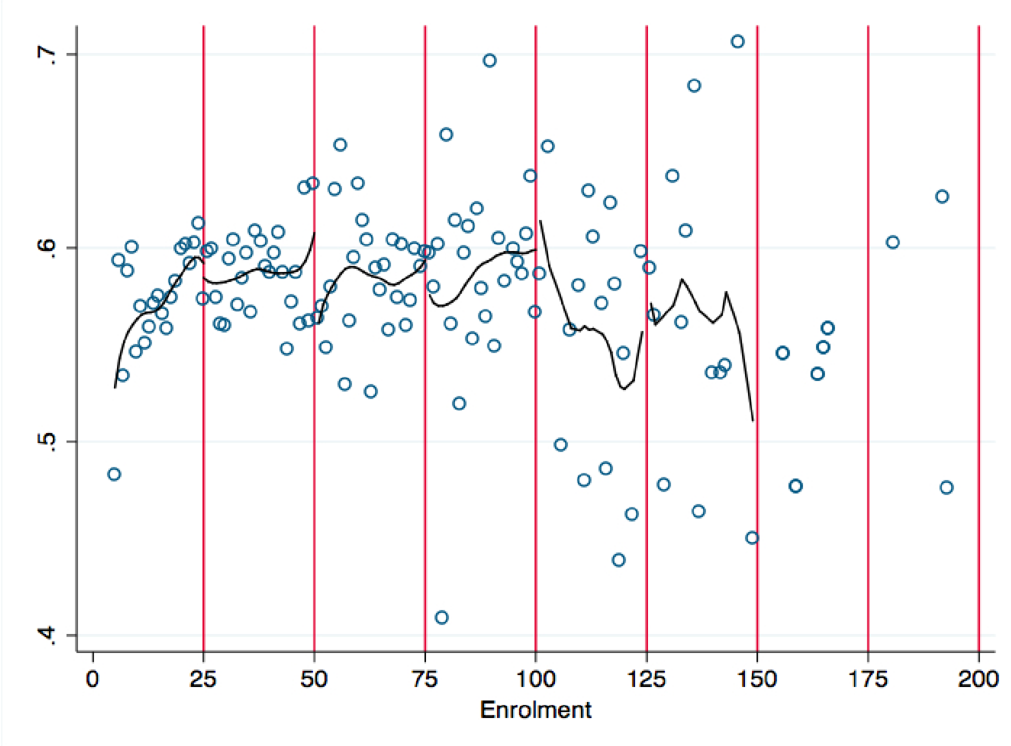


offs, regression discontinuity inference requires that the number of observations nearly to cut-off values on left and right sides are approximately the same. In regard to it, Urquiola and Verhoogen (2009) have constructed a theoretical model of quality differentiation and sorting, according to which schools may stuck at enrolment cut-offs. The authors find the confirm to this prediction by looking at the variations in the number of Chilean private schools around cut-offs of enrolment.

Figure 8 presents the distribution of schools according to their enrolment in V grade in our dataset. Though we consider public primary schools where non-subscribing of pupils seems to be difficult to realize, "stucking" behavior seems to be present at the first threshold. Specifically, if we look at Figure 8, the number of schools with enrolment $25 \leq \Phi < 30$, where an additional class should be added for few pupils, declines compared to the schools with enrolment $20 \leq \Phi < 25$.

Figures 9 and 10 display a graphical analysis visualizing the relationship between test scores and enrolment. They present the results of local weighted regressions of test scores in, respectively, math and Italian language on grade enrolment. Unexpectedly, we see that the average test scores in both math and Italian tend to get worse in the first four intervals when enrolment exceeds its

Figure 9: Distribution of test scores in math along enrolment (full sample), 2008/09

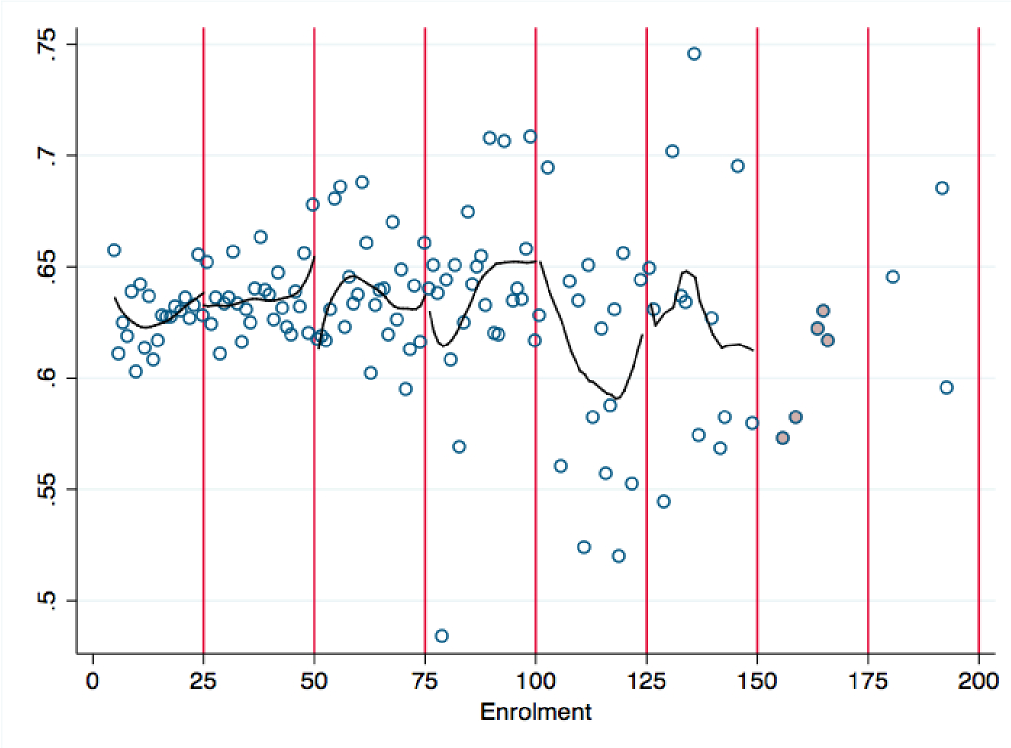


cut-offs values. Only at the threshold of 125 pupils, test results seem to get better on the right side of the enrolment cut-off.

Table 6 shows the results of the regression model in (2) and (3). Column (1) reports the results of the first-stage regression of class size on indicators for whether enrolment exceeds the cut-off values of enrolment along with the piecewise linear splines. In this specification, standard errors are clustered by enrolment.[12] As we can see - the adjusted R-squared is about 0.66 - the rule explains a great part of variation in class size. Its average decline equals to about 9 pupils at the first cut-off gradually decreasing in the following ones as predicted by the rule. The declines are statistically significant in the first four cut-offs. In Column (2) and Column (3), we present the results of the reduced-form regressions of test outcomes in math and Italian language, which show that test scores decline on average in all cut-offs, although the declines are statistically significant only for Italian in 150 and 175 cut-offs. Columns (4) and (5) report IV specifications for math and

[12]Lee and Card (2008) suggest this procedure when the forcing variable is a discrete one.

Figure 10: Distribution of test scores in Italian language along enrolment (full sample), 2008/09



Italian language, respectively, while in Columns (6) and (7) we include selected observables. In both cases, the coefficients on class size are mostly positive, but always not statistically significant. By adding some observables to IV specification as an additional control for possible selection problem (Columns 6-7), it seems that the second stage estimation of pupils performance on class size does not change considerably the results, which suggest that RD design may be appropriate estimation strategy in this case.

4.2 Results (Small Intervals around cut-offs)

In this section, we focus on small intervals around enrolment cut-offs as in Van der Klaauw (2002).[13] We want to estimate class-size effect using the observations from small intervals around

[13]See also Hoxby (2000) and Angrist and Levy (1999, p. 540).

the first four cut-offs (25, 50, 75 and 100).[14] [15]Analysing the observation from intervals of 3 pupils around the first four cut-offs, we do not find empirical evidence of precise sorting pattern of schools and pupils/families in the selected intervals. Table 7 reports the results of the reduced-form regressions of pupils characteristics (ISEI indexes of pupils' father and mother, pupils' mother education) and of the average number of pupils with special education needs on dummies for cut-off values, on enrolment and piecewise linear splines. As we see, ISEI index of both, pupils' mother and father, result significantly related to the class size drop at the second cut-off, but ISEI index of father increases while the index of pupil's mother decreases. Second, though ISEI index of father increases at the second cut-off ($\Phi > 50$), it decreases in average values, being statistically significant, if we consider all right part of the second interval $\Phi \in [51, 53]$. Precisely, the coefficient of the second linear spline is significant and negative. Only in the fourth cut-off interval, all selected observables - ISEI indexes of pupils' father and mother, and mother's education - decline in average in the right part of the cut-off interval $\Phi \in [101, 103]$, yet their average variations at the break $\Phi = 101$ (coefficient for dummy) are uniformly statistically not significant. Analogously, average number of pupils with special educational needs varies through intervals without following any precise pattern.

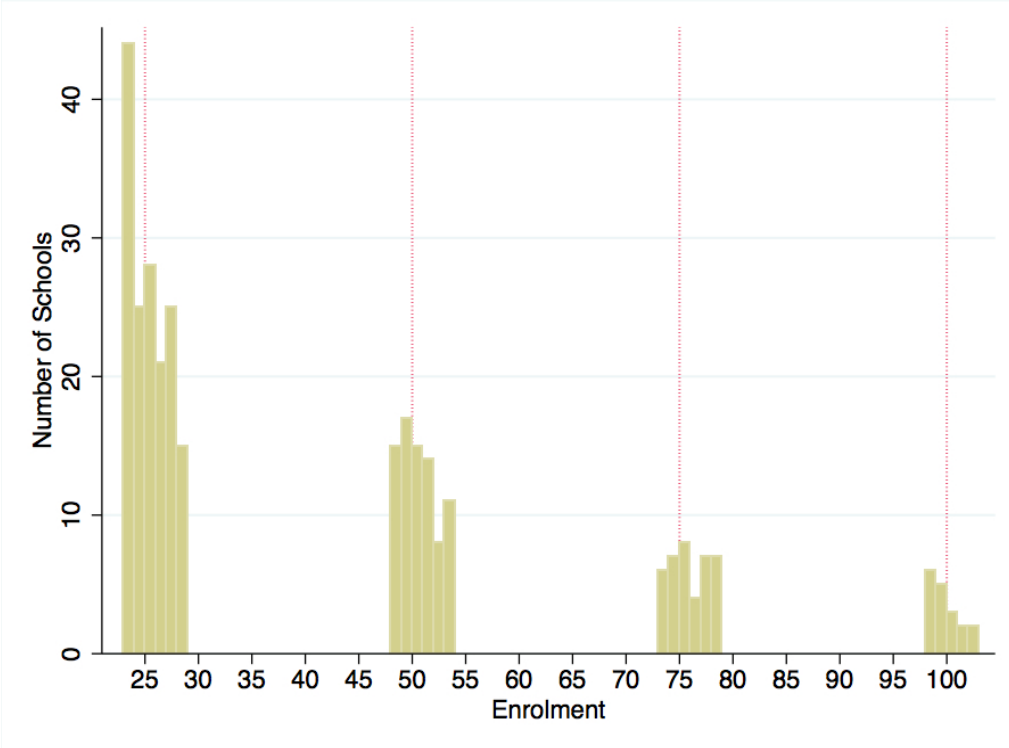
As in the previous section, we control whether the number of schools varies noticeably between left and right sides of cut-off intervals, which would mean that schools in some way have a control over their enrolment by not subscribing to avoid adding classes for few additional pupils. Figure 11 presents an histogram of V grade enrolment in schools belonging to the first four cut-off intervals of ± 3 pupils. Though a precise pattern of schools' distribution remains not clear, it seems that the number of schools on right sides of the majority of the four cut-off intervals tends to decline.

Table 8 reports the results of first stage, reduced-form for test outcomes in math and Italian

[14]We chose the first four cut-offs of enrolment, because more than 90 percent of schools in our dataset have enrolment in V grade $\Phi \leq 100$ pupils.

[15]In addition, at the first four cut-offs, according to the rule, declines in class size are particularly large becoming less noticeable as enrolment grows. For example, as predicted by the rule, class size drops from 25 to 13 pupils at the first cut-off, from 25 to 17 at the second one, from 25 to 19 at the third one, and, finally, from 25 to 20 at the fourth one.

Figure 11: Number of schools in enrolment intervals (25±3; 50±3, 75±3; 100±3), 2008/09



language, and IV stage for math and Italian language without and with inclusion of some observables using the observations from intervals: 25±3; 50±3; 75±3, 100±3. Suprisingly, the first stage estimation shows that average class size, except for the 100±3 interval, in average is positively related with increases in enrolment at cut-offs (see Column (1)). These results agree with the visual evidence: histograms of average class size show that the selected intervals include many non-compliant schools, *i.e.* schools that add an additional class before the enrolment actually exceeds multiples of 25; and schools that do not add it when the rule prescribes so. The latter ones are particularly numerous in the first interval.[16] Columns (2) and (3) of Table 8 report the estimates of reduced form of test scores in math and Italian language. We see that test performance decline in average only at the second cut-off where average class size increases and test results in math and Italian language decline. The evidence is different for the fourth cut-off where average class size decreases together with test results in math and Italian language. The point estimates of IV spec-

[16]For space economy, we omit these Figures for intervals of 3 pupils in the text; they are available from the authors upon request.

ification of class size increasing upon pupils test outcomes are uniformly negative for both math and Italian language, but not statistically significant (columns 4-5). When selected observables are included as controls (Columns 6-7), the coefficients on class size, although they largely fall, are still negative and not statistically significant.[17]

Table 9 provides the same results by enlarging the range of intervals: 25±5; 50±5; 75±5, 100±5. The first stage estimation in Column (1) shows now that only for the 100±5 interval average class size is significantly and negatively related with increases in enrolment at cut-offs, while for all other intervals the coefficients are always not statistically significant. Columns (2) and (3) of Table 9 report the estimates of reduced form of test scores. We observe that test outcomes in math and Italian language decline significantly at the second and fourth of the cut-offs. The estimates of IV specification of the effect of class size on test scores are still negative for both math and Italian language, although not statistically significant in both specifications (without, Columns 4-5, and with inclusion of selected observables, Columns 6-7).

The main results, obtained still focusing on the observations from small intervals around first four cut-offs, are summarized in Table 10. Columns 1-4 report the results of IV stage separately for intervals of ±3 and ±5 pupils using the indicator for whether schools' grade enrolment is above the respective cut-off as an instrument. Columns 5-6 show, respectively, the results of IV stage for pooled cut-off estimation without and with inclusion of observables. For each interval we provide the measure of average class size on both sides of cut-off intervals. In all estimations standard errors are clustered by V grade enrolment. As we see class-size reduction has always a positive and significant (**;***) effect on both test results in math and Italian language in the fourth of cut-off intervals whereas average class-size declines from 23.876 to 21.470 in ±3 pupils interval, and from 23.612 to 21.470 in ±5 pupils interval (Column 4). Moreover, the effect of class-size reduction has a positive and significant (**) effect on test results in math and Italian language in the second of cut-off intervals (Column 2; ±5 pupils). This is the only clear evidence supporting class-size reduction policy that we find in our data. In the remaining cut-offs, most coefficients are

[17]Lee and Card (2008), and Urquiola and Verhoogen (2009) suggest that clustering by enrolment level lowers significance levels.

statistically not significant.

4.3 Class Size Policy

In this section, we focus on schools that in 2008/09 respect the maximum class-size rule perfectly. In the previous section, looking at the observations in small intervals around enrolment cut-offs, we have found that they include a great number of schools that deviate from the rule by adding an additional class when enrolment is not sufficient, or non-adding it as predicted by rule. In this section, according to our data, we want to see whether compliant schools differ from schools that deviate from the rule in a systematic way. In our data there are 1,580 schools units, among them 1,467 have the average class size as predicted by the rule. Clearly, the deviating schools are concentrated in cut-off intervals (see Figure 3). Figure 12 reports the distribution of compliant schools according to enrolment. As we can see, stucking behavior seems to be more evident in this subsample of schools compared to the full sample in Figure 9. Specifically, the number of schools applying the rule declines when enrolment exceeds multiples of 25 pupils of few pupils and grows straight afterwards. The same pattern can be observed at the cut-offs of 50, 75 and 100 pupils.

Focusing on schools in ± 3 pupils intervals around cut-offs, Figure13 shows that the number of compliant schools in cut-off intervals on their left sides is systematically larger than on the right ones (85 and 42 rispectively in the first cut-off interval; 37 and 19 in the second one; 17 and 10 in the third one; and 11 and 3 in the fourth one).

Figure 12: Number of Schools (sample of compliant schools), 2008/09

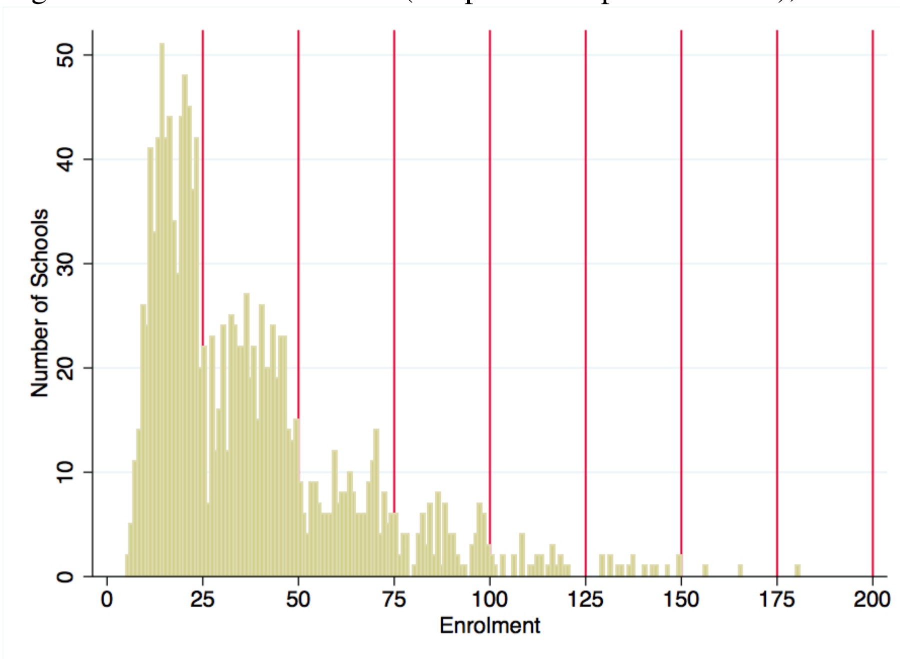
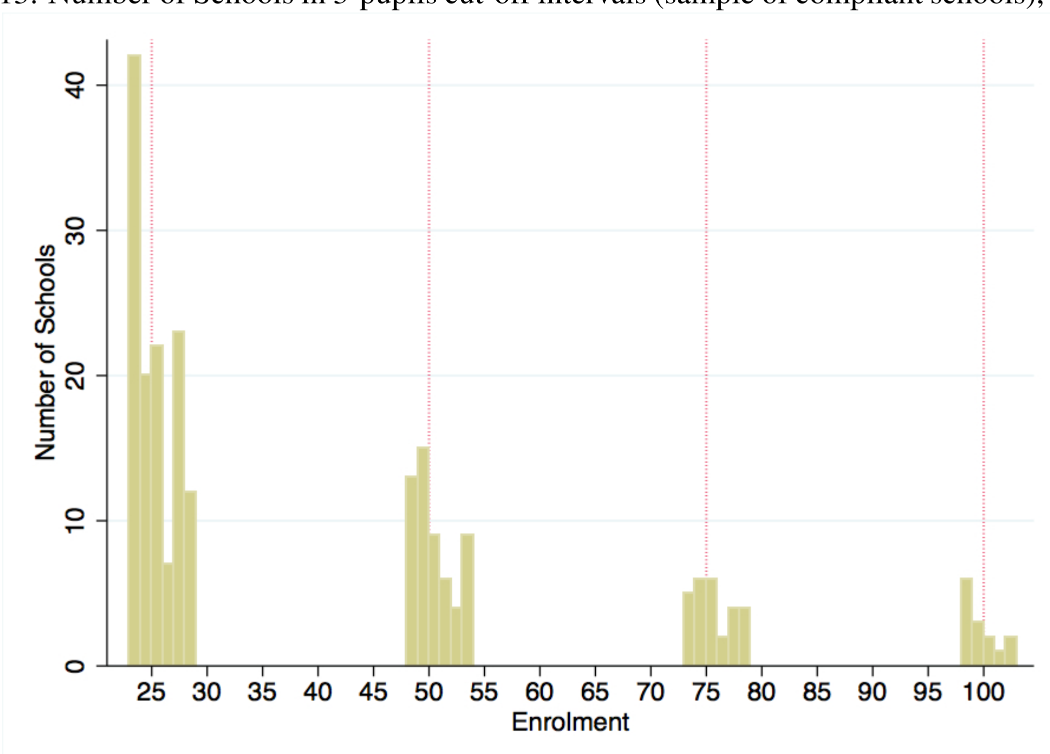


Figure 13: Number of Schools in 3-pupils cut-off intervals (sample of compliant schools), 2008/09



As in the previous analysis (full and±3 and 5 pupils samples), we look at the data of compliant schools to control for the distribution of pupils' characteristics around cut-off values of enrolment. Table 11 reports the results of reduced-form regression of selected pupils' characteristics on cut-off dummies, enrolment and piecewise linear splines. We do not provide the results of reduced-form for average number of pupils with special needs as it results irrelevant in cut-off intervals. As we can see, Table 11 presents a clear evidence of sorting of pupils' characteristics around cut-offs. Specifically, ISEI index of pupils' father is negatively related to class size decline and statistically significant at *all* cut-offs. At the first cut-off, all of the baseline covariates decrease; it is worth noting that the average declines in ISEI indexes of pupils' mother and father are particularly large and equal to 10.3 and 7.8, respectively. The coefficient values for ISEI index of mother and mother's education are negative and significant also, respectively, at the second and the third of cut-offs. These findings make us guess that there may exist a supplementary "rule" to the maximum class-size rule, according to which public primary schools' probability to get resources, mainly teachers, in order to add a class when grade enrolment exceeds the threshold values increases when they have pupils coming from "unfavorite" socio-economic background (*e.g.* less educated mother, lower socioeconomics status of mother and/or father).[18]

5 Conclusions

In this paper we make an attempt to estimate the class-size effect on the pupils' performance using the data from Italian primary schools. We base our estimation strategy on RD design, which has been largely used in the recent economic literature on class-size issue. The application of RD design to estimate class-size effect is appropriate when increases in grade enrolment are linked with jumps in class size as predicted by the threshold rule generating a discontinuous relation between the two variables. Additionally, in order to apply RD estimation strategy, practitioners have to test if the assumptions of RD analyses are not infringed, otherwise it would invalid to infer a "treatment" effect of class size on pupils' test results. The RD main assumptions are that schools

[18]On this see West and Woessmann (2006)

and/or parents should not be able to exactly manipulate the forcing variable (enrolment), and there should not be evidence of sorting of baseline covariates influent for scholastic performance on threshold values of enrolment. It follows then that baseline covariates should be used to test the adequacy of RD design by controling whether their distribution is continuous at cut-offs points and whether there is not a precise sorting of schools according their enrolment there, *i.e.* the number of schools with large and small classes should not vary consistently at cut-offs of enrolment.

Considering our full sample, we find some evidence of discontinuities in the distribution of observed baseline covariates and of numbers of schools along the dimension of enrolment levels. Differently, when focusing on small intervals (± 3 and ± 5 pupils) around selected enrolment cut-offs (25, 50, 75 and 100), selection problem appears to become less evident, as, on the one hand, we still do not observe clear stucking behavior of schools at the enrolment thresholds, and, on the other hand, pupils' characteristics result to be distributed smoothly in the large majority of cut-offs for the ± 3 and 5 pupils samples. Considering these samples, however, we do not find a sufficient evidence supporting class-size reduction policy. In the 2 cut-off interval and, only for the ± 5 pupils sample, the 4 cut-off interval, class-size reduction has a positive and significant effect on test results in math and Italian language. For the rest of our analysis using these samples, class-size is a variable mostly not statistically significant for test results.

We do not find a significant evidence which would strongly support class-size policy, yet our results call for further research on this issue. As we observe in our data, class size is largely used in primary public schools as a kind of compensatory policy. In fact, when we focus on schools that in 2008/09 applied the maximum class-size rule perfectly, we find that they differ from "deviating" from the exogenous maximum-class-size rule schools in a more systematic way. First, the stucking behavior is more evident in the reduced sample of compliant schools. Second, in this sample there is a clearer evidence of sorting of pupils' characteristics around cut-offs points: right sides of cut-off intervals include more pupils with "unfavorite" socio-economic background.[19] A possible issue of future research can be to understand whether the compensatory policy through class-size

[19]Lee and Lemieux note that "The observed sorting may well be evidence of economic agents responding to incentives, and may help identify economical interesting phenomena." (p. 79-80, 2009)

reduction is efficient. The data about Italian schools state that only primary level of schooling manages to compensate for family background of pupils, while in “media” and “superiori” schools education gaps grow depending on students families’ background (Fondazione Giovanni Agnelli, 2011).

References

[1] Biagi, F., and Fontana, A. (2009). “Fabbisogni standard per l’istruzione: Problematiche e stime,” in *Rapporto ISAE: Finanza Pubblica e Istituzioni,* Roma.

[2] Bratti, M., Checchi, D., Filippin, A. (2007). “Territorial differences in Italian students’ mathematical competences: Evidence from PISA,” *Giornale degli Economisti e Annali di Economia*, 66(3), pp. 299-335.

[3] Brunello, G., and Checchi, D. (2005). “School quality and family background in Italy,” *Economics of Education Review*, 24, pp.563-77.

[4] Cipollone, P., and Sestito, P. (2007). “Quanto imparano gli studenti italiani: i divari Nord-Sud,” Unpublished Manuscript, Banca d’Italia.

[5] Dobbelsteen, S., Levin, J., and Oosterbeek, H. (2002). “The causal effect of class size on scholastic achievement: distinguishing the pure class size effect from the effect of changes in class composition,” *Oxford Bulletin of Economics and Statistics*, 64(17), pp.17-38.

[6] Fondazione Giovanni Agnelli (2011). “*Il Rapporto sulla scuola in Italia - 2011*”, Bari: Laterza.

[7] Hahn, J., Todd, P. E., and van der Klaauw, W. (2001). “Identification and estimation of treatment effects with a regression-discontinuity design,” *Econometrica,* 69(1), 201-09.

[8] Hanushek, E. A. (1997). “Assessing the effct of school resources on student performance,” *Educational Evaluation and Policy Analysis*, 19, pp.141-64.

[9] Hanushek, E. A. (2003). “The failure of input-based schooling policies,” *Economic Journal,* 113(485), pp.64-98.

[10] Hedges, L. V., Laine, R., and Greenwald, R. (1994). “Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes,” *Education Researcher*, 23, pp.5-14.

[11] Hoxby, C. (2000). "The effects of class size on student achievement: New evidence from population variation," *The Quarterly Journal of Economics*, 115(4), pp.1239-85.

[12] Krueger, A. (2003). "Economic considerations and class size," *Economic Journal,* 113(485), pp.34-63.

[13] Lee, D.S. (2008). "Randomized experiments from non-random selection in U.S. house elections," *Journal of Econometrics*, 142(2), pp.675-97.

[14] Lee, D. S., and Card, D. (2008). "Regression Discontinuity inference with specification error," *Journal of Econometrics*, 142(2), pp.655-74.

[15] Lee, D. S, and Lemiux, T. (2009). "Regression discontinuity design in economics," *NBER Working Paper*, 14723, February.

[16] Leuven, E., Oosterbeek, H., and Rønning, M. (2008). "Quasi-experimental estimates of the effect of class size achievement in Norway," *The Scandinavian Journal of Economics*, 110(4), pp.663-93.

[17] McEwan, P. J., and Urquiola, M.S. (2005). "Precise sorting around cutoffs in the regression-discontinuity design: Evidence from class size reduction," Unpublished Manuscript, December.

[18] Ney, B., Hedges, L.V., and Konstantopoulos, S. (2000). "Are effects of small classes cumulative?" Evidence from a Tennessee experiment," *The Journal of Educational Research*, 94 (6), pp.226-45.

[19] Thistlethwaite, D., and Campbell, D. (1960). "Regression-Discontinuity analysis: An alternative to the ex post factor experiment," *Journal of Educational Psychology*, 51, pp.309-17.

[20] Trochim, W.K. (1984). *Research design for program evaluation: The regression-discontinuity approach,* Beverly Hills, CA: Sage.

[21] Urquiola, M.S. (2006). "Identifying class size effects in developing countries: Evidence from rural schools in Bolivia," *Review of Economics and Statistics*, 88(1), pp.171-77.

[22] Urquiola, M.S., and Verhoogen, E. (2009). "Class size caps, sorting, and the regression discontinuity design," *American Economic Review,* 99(1), pp.179-215.

[23] van der Klaauw, W. (2002). "Estimating the effect of financial aid offers on college enrolment: A Regression-Discontinuity Approach," *International Economic Review,* 43(4), pp.1249-87.

[24] van Der Klaauw, W. (2008). "Regression-Discontinuity analysis: A survey of recent developments in economics," *Review of Labour Economics*, 22(2), pp. 219-45.

[25] West, M. R, and Woessmann, L. (2006). "Which schools systems sort weaker students into smaller classes? International evidence," *European Journal of Political Economy*, 22, pp.944-68.

[26] Woessmann, L. (2005). "Educational production in Europe," *Economic Policy*, 43, pp.445-93.

[27] Zada, D.C., Gradstein, M., and Reuven, E. (2009). "Class size and the Regression Discontinuity design: the case of public schools," *IZA discussion paper*, 4679, December.

Appendix

Table 1: Average class size in V grade in Italian regions, 2008/09

	Average class size (1)	Average class size (2)
Abruzzo	16.81	19.19
Basilicata	16.42	18.14
Calabria	14.9	18.25
Campania	17.19	19.47
Emilia Romagna	19.61	20.75
Friuli Venezia Giulia	17.34	19
Lazio	18.33	19.94
Liguria	17.77	18.95
Lombardia	18.86	20.49
Marche	18.78	20.23
Molise	14.42	18
Piemonte	17.49	19.27
Puglia	19.48	20.64
Sardegna	16.55	17.86
Sicilia	17.77	20.70
Toscana	18.79	19.61
Umbria	17.08	18.76
Veneto	18.12	18.46
Total	17.94	19.41

Note: In Column 1, regional average class sizes are calculated using the data from all public primary schools in 2008/2009. In Column 2, regional average class size are calculated using the data of schools in our dataset.

Table 2: Summary Statistics, 2008/09

	Mean	SD	Min	Max
Girls	0.49	(0.49)	0	1
Foreign citizenship	0.053	(0.23)	0	1
Foreign country of birth of at least one parent	0.059	(0.23)	0	1
ISEI index father	41.43	(14)	0	68
ISEI index mother	26.25	(23.58)	0	68
Homemaker mother	0.38	(0.49)	0	1
Mother's education (1,2,3,4,5,6)	3.29	(1.36)	1	6
Intact family status (1,2,3,4)	1.17	(0.54)	1	4
Internet at home	0.75	(0.43)	0	1
Family language (1,2,3)	1.22	(0.53)	1	3
Homework assistance at home	0.97	(0.36)	0	1
Schooling time	32.82	(4.25)	27	40
Average V grade class size	19.45	(3.99)	2	29
Average number of pupils with special needs in V grade classes	0.56	(0.60)	0	4
V grade enrolment	44.41	29.45	5	274
School enrolment	217.3	(142.77)	6	1,407
Town population	105,552.07	(313,209.6)	500	2,547000
Mountain municipality (0,1,2)	0.48	(0.76)	0	2
N pupils		25,407		
N schools		1,561		

Table 3: Average class size in V grade on selected observables, 2008/09

	(1)	(2)	(3)
Girls	-0.049454[-1.0670]	-0.026707[-0.8384]	-0.028921[-0.9016]
Foreign Citizenship	0.042926[0.2879]	0.099019[1.0498]	-0.053408[-0.5684]
ISEI Index Father	0.014723***[5.1135]	0.006290***[3.2304]	0.004784**[2.5164]
ISEI Index Mother	0.002900[1.2261]	0.000033[0.0170]	0.000497[0.2649]
Homemaker Mother	-0.198376*[-1.7966]	-0.32523***[-3.8367]	-0.11337[-1.3982]
Mother's Education	0.135299***[4.0131]	0.036768*[1.6843]	0.038560*[1.7244]
Cohabiting Status	0.158667***[2.6345]	0.078944[1.5177]	0.049892[0.9903]
Siblings	-0.065774[-0.4811]	0.192834*[1.8138]	0.138759[1.3675]
Average Number of Pupils with Special Needs		0.029189[0.2080]	0.000567[0.0040]
Dummy Full-time Classes		0.559463***[2.7734]	0.463471**[2.2681]
Dummy Mounting Status			-0.278924*[-1.9397]
Regional Dummies	NO	NO	YES
V grade enrolment		0.269967***[6.1097]	0.260420***[6.1461]
V grade enrolment^2		-0.00218***[-4.9499]	-0.00208***[-4.9408]
V grade enrolment^3		0.000005***[4.2476]	0.000005***[4.2760]
Constant	18.28058***[37.732]	11.52080***[9.4558]	11.62878***[8.9224]
N pupils	25,407	25,407	25,407
N schools	1,580	1,580	1,580
Adjusted R-squared	0.0111	0.3601	0.3888

Note: Robust standard errors in parentheses are clustered for V enrolment. *** p<0.01, ** p<0.05, * p<0.1. The regression in column (3) include regional dummies.

Table 4: Reduced form of selected observables (full sample), 2008/09

	ISEI index father	ISEI index mother	Mother's education	Average number of pupils with special needs
$1\{\Phi \geq 26\}$	-1.442426*(0.7612)	-4.25509***(1.1256)	-0.153009**(0.0662)	-0.21223***(0.0628)
$1\{\Phi \geq 51\}$	-0.479656(0.7402)	-2.195675*(1.2929)	-0.096016(0.0820)	-0.053653(0.0787)
$1\{\Phi \geq 76\}$	-1.424020(1.4181)	-3.621004*(2.0253)	0.002231(0.1291)	0.134445(0.0988)
$1\{\Phi \geq 101\}$	-2.739835(2.8324)	-5.877963(4.8102)	-0.284840(0.3592)	0.238710(0.2001)
$1\{\Phi \geq 125\}$	3.002955(4.2507)	1.255818(6.6750)	0.458503(0.5402)	-0.428608**(0.1811)
$1\{\Phi \geq 150\}$	-1.284081(5.1726)	14.332224(19.4269)	0.328501(0.9411)	0.068589(0.5575)
$1\{\Phi \geq 175\}$	4.514970(6.6951)	17.190307(15.3664)	0.400489(0.7983)	0.309277(0.6437)
Φ	0.175884***(0.0468)	0.447155***(0.0765)	0.019274***(0.0042)	0.023567***(0.0050)
$(\Phi-25)1\{\Phi \geq 26\}$	-0.058561(0.0636)	-0.26349***(0.0986)	-0.005644(0.0055)	-0.014320**(0.0060)
$(\Phi-50)1\{\Phi \geq 51\}$	-0.126090**(0.0552)	-0.124815(0.0965)	-0.01544***(0.0058)	-0.011056**(0.0048)
$(\Phi-75)1\{\Phi \geq 76\}$	0.172814*(0.0993)	0.105862(0.1632)	0.010930(0.0094)	-0.010986(0.0068)
$(\Phi-100)1\{\Phi \geq 101\}$	-0.223417(0.1793)	-0.175793(0.3586)	-0.015374(0.0263)	0.015604(0.0141)
$(\Phi-125)1\{\Phi \geq 126\}$	0.007636(0.2872)	-0.185410(0.4546)	-0.015931(0.0353)	0.015835(0.0151)
$(\Phi-150)1\{\Phi \geq 156\}$	-0.119584(0.4337)	-0.717754(1.3031)	-0.001856(0.0674)	-0.040800(0.0445)
$(\Phi-175)1\{\Phi \geq 176\}$	0.301190(0.3611)	0.980869(1.2663)	0.037525(0.0627)	0.023352(0.0439)
Constant	37.09860***(0.8957)	17.42046***(1.4712)	2.831901***(0.0891)	0.078356(0.0830)
Observations	25407	25407	25407	25407
Adjusted R-squared	0.0078	0.0063	0.0091	0.0227

Note: In all regressions, standard errors are clustered by enrolment levels, see Lee and Card (2008). *** p<0.01, ** p<0.05, * p<0.1.

Table 6: First stage, Reduced-form, and base IV specifications (full sample), 2008/09

	First stage	Reduced-form		IV stage		IV stage with selected observables	
		Math	Italian language	Math	Italian language	Math	Italian language
Class size				0.0019[1.561070]	0.0004[0.482349]	0.0011[0.908194]	-0.0002[-0.257092]
$1\{\Phi \geq 26\}$	-9.1434***[-7.614248]	-0.0159[-1.463022]	-0.0026[-0.342342]				
$1\{\Phi \geq 51\}$	-4.4362***[-4.656766]	-0.0116[-0.759498]	-0.0016[-0.130341]				
$1\{\Phi \geq 76\}$	-2.1556***[-3.116623]	-0.0059[-0.455055]	-0.0077[-0.627582]				
$1\{\Phi \geq 101\}$	-1.9039***[-3.058748]	-0.0027[-0.091130]	-0.01324[-0.476806]				
$1\{\Phi \geq 125\}$	0.7158[0.583753]	0.0610[1.109579]	0.0675[1.464750]				
$1\{\Phi \geq 150\}$	1.6478[0.499686]	-0.0437[-0.546328]	-0.0765**[-2.305240]				
$1\{\Phi \geq 175\}$	1.6493[0.580666]	-0.0059[-0.108556]	-0.04455*[-1.870886]				
Φ	0.8916***[20.608152]	0.0025***[3.145683]	0.00091[1.616019]	0.0010[1.148356]	0.0006[0.864607]	0.0009[1.012493]	0.0004[0.630462]
$(\Phi - 25)1\{\Phi \geq 26\}$	-0.5390***[-6.762968]	-0.0023**[-2.227041]	-0.0007[-0.917171]	-0.0014[-1.266100]	-0.0005[-0.531408]	-0.0015[-1.331527]	-0.0005[-0.564117]
$(\Phi - 50)1\{\Phi \geq 51\}$	-0.1793**[-2.255187]	0.0002[0.272506]	-0.0002[-0.291229]	0.0005[0.582258]	-0.0002[-0.418007]	0.0008[0.973526]	0.00006[0.107036]
$(\Phi - 75)1\{\Phi \geq 76\}$	-0.0918*[-1.666363]	0.0001[0.122735]	0.0010[1.126039]	0.0002[0.284429]	0.0006[0.726411]	-0.00002[-0.027293]	0.0003[0.339204]
$(\Phi - 100)1\{\Phi \geq 101\}$	-0.0799[-1.393899]	-0.0035[-1.563659]	-0.0036*[-1.813448]	-0.0019[-1.384692]	-0.0022[-1.630306]	-0.0015[-1.059379]	-0.002[-1.282714]
$(\Phi - 125)1\{\Phi \geq 151\}$	0.0740[1.013086]	0.001[0.223832]	0.0011[0.384755]	0.0014[0.460239]	0.0025[1.179475]	0.0010[0.362829]	0.0021[1.110111]
$(\Phi - 150)1\{\Phi \geq 151\}$	-0.3087[-1.297835]	0.0051[1.020461]	0.0072***[2.991216]	0.0007[0.146425]	-0.0007[-0.262927]	0.0008[0.167753]	-0.0006[-0.234537]
$(\Phi - 175)1\{\Phi \geq 176\}$	0.2659[1.150893]	-0.0032[-0.913285]	-0.0057***[-7.757145]	-0.0002[-0.057058]	0.0003[0.170784]	-0.0007[-0.203296]	-0.0001[-0.064940]
ISEI index of father						0.0009***[9.436359]	0.0009***[12.484054]
ISEI index of mother						0.0003***[5.456764]	0.0003***[6.629368]
Mother's education						0.0185***[17.534886]	0.0203***[23.468830]
Citizenship						-0.0394***[-8.036349]	-0.0593***[-14.670787]
Pupils with special needs						0.0027[0.562436]	0.0054*[1.684257]
Dummy full-time classes						0.0034[0.581741]	-0.0083**[-2.290604]
Constant	1.5525**[2.394787]	0.5357***[37.936734]	0.6123***[56.280001]	0.5316***[35.356259]	0.611***[53.245598]	0.4836***[30.340452]	0.5786***[45.275868]
Observations	25,407	25,407	25,407	25,407	25,407	25,407	25407
Adjusted R-squared	0.662523	0.003328	0.002116	0.003087	0.001364	0.058442	0.085257

Note: In all regressions, standard errors are clustered by enrolment level, see Lee and Card (2008).*** p<0.01, ** p<0.05, * p<0.1

Table 7: Reduced-form of selected observables on enrolment (+/- 3 pupils intervals), 2008/09

Class size	ISEI Father	ISEI Mother	Mother's education	Pupils with special needs
$1\{\Phi \geq 26\}$	-3.2883(2.5079)	-2.5363(2.7587)	-0.261644(0.2209)	-0.3069***(0.0499)
$1\{\Phi \geq 51\}$	2.5542***(0.8889)	-1.9496*(1.1088)	0.050558(0.1821)	0.4053***(0.0972)
$1\{\Phi \geq 76\}$	-3.46454(2.1830)	2.825578(2.3141)	0.143400(0.3019)	0.2110(0.3235)
$1\{\Phi \geq 101\}$	-1.0707(1.1072)	-1.020663(1.6489)	0.051864(0.0946)	-0.560***(0.0710)
Φ	0.0174(0.0177)	-0.035368(0.0240)	0.000905(0.0015)	-0.00296**(0.0013)
$(\Phi - 25)1\{\Phi \geq 26\}$	1.4980(1.1920)	-0.643432(1.3018)	0.080478(0.1049)	0.067425***(0.0065)
$(\Phi - 50)1\{\Phi \geq 51\}$	-0.8283**(0.3969)	0.470961(0.4469)	-0.039146(0.0871)	-0.1689***(0.0443)
$(\Phi - 75)1\{\Phi \geq 76\}$	2.031**(0.8101)	-1.136631(0.7985)	0.002731(0.1157)	-0.030554(0.1253)
$(\Phi - 100)1\{\Phi \geq 101\}$	-0.6218***(0.0177)	-1.53963***(0.0240)	-0.2237***(0.0015)	0.388616***(0.0013)
Constant	41.03482***(0.7707)	29.28249***(0.9712)	3.330471***(0.0655)	0.690685***(0.0738)
Observations	5,396	5,396	5,396	5,396
Adjusted R-squared	0.0077	0.0064	0.0054	0.0327

Note: In all regressions, standard errors are clustered by enrolment levels, see Lee and Card (2008). *** p<0.01, ** p<0.05, *p<0.1.

Table 8: First stage, Reduced-Form and IV stage (+/- 3 pupils intervals), 2008/09

	First stage	Reduced-form		IV		IV with observables	
	(1)	Math (2)	Italian (3)	Math (4)	Italian (5)	Math (6)	Italian (7)
Class Size				-0.005190(0.0146)	-0.003018(0.0161)	-0.002389(0.0180)	-0.000656(0.0186)
$1\{\Phi \geq 26\}$	0.727449(2.8013)	0.016253(0.0119)	0.015994(0.0137)				
$1\{\Phi \geq 51\}$	1.085096***(0.3597)	-0.027010***(0.0091)	-0.027144***(0.0061)				
$1\{\Phi \geq 76\}$	0.670140**(0.2611)	-0.018062(0.0162)	-0.004612(0.0125)				
$1\{\Phi \geq 101\}$	-0.695183**(0.2975)	-0.051235***(0.0121)	-0.063188***(0.0139)				
Φ	0.024155***(0.0045)	0.000103(0.0002)	0.000270(0.0002)	0.000157(0.0004)	0.000278(0.0005)	0.000049(0.0005)	0.000189(0.0005)
$(\Phi-25)1\{\Phi \geq 26\}$	-2.875964**(1.3461)	-0.011280***(0.0039)	-0.009216(0.0059)	-0.017721(0.0393)	-0.010166(0.0430)	-0.009870(0.0475)	-0.003130(0.0487)
$(\Phi-50)1\{\Phi \geq 51\}$	-1.834880***(0.1560)	-0.007162**(0.0026)	-0.000276(0.0004)	-0.025746(0.0228)	-0.015962(0.0237)	-0.021861(0.0278)	-0.012697(0.0275)
$(\Phi-75)1\{\Phi \geq 76\}$	-0.852245***(0.0599)	0.003576(0.0053)	-0.002082(0.0032)	-0.005844(0.0107)	-0.004736(0.0111)	-0.005556(0.0121)	-0.005015(0.0126)
$(\Phi-100)1\{\Phi \geq 101\}$	-0.960405***(0.0045)	0.032427***(0.0002)	0.032789***(0.0002)	0.007624(0.0194)	0.005876(0.0216)	0.016525(0.0237)	0.013763(0.0247)
Average number of pupils with special needs						-0.003077(0.0092)	0.001518(0.0037)
Mother's education						0.015243***(0.0028)	0.016755***(0.0020)
ISEI index father						0.001056***(0.0003)	0.001029***(0.0002)
ISEI index mother						0.000284**(0.0001)	0.000270**(0.0001)
Constant	21.688432***(0.2628)	0.595040***(0.0120)	0.630779***(0.0095)	0.710567**(0.3228)	0.698901*(0.3510)	0.550023(0.3959)	0.541377(0.4024)
Observations	5,396	5,396	5,396	5396	5396	5396	5396
Adjusted R-squared	0.2756	0.0077	0.0064	0.0066	0.0048	0.0508	0.0679

Note: In all regressions, standard errors are clustered by enrolment level, see Lee and Card (2008).*** p<0.01, ** p<0.05, * p<0.1

Table 9: First stage, Reduced-form and IV stage (+/- 5 pupils intervals), 2008/09

	First stage	Reduced-form		IV stage		IV stage with observables	
	(1)	(2) Math	(3) Italian language	(4) Math	(5) Italian language	(6) Math	(7) Italian language
Class Size				-0.0063[-1.0536]	-0.0049[-0.6483]	-0.0062[-0.8451]	-0.005164[-0.7586]
$1\{\Phi \geq 26\}$	-1.5590[-0.5368]	0.0194**[2.5617]	0.0102[0.8004]				
$1\{\Phi \geq 51\}$	-0.2207[-0.3370]	-0.0523***[-3.5140]	-0.0405***[-3.2612]				
$1\{\Phi \geq 76\}$	-1.3748[-1.6422]	-0.0198[-0.7889]	0.0111[0.7477]				
$1\{\Phi \geq 101\}$	-0.9356**[-2.6952]	-0.0363***[-4.7140]	-0.0496***[-5.8792]				
Φ	0.0359***[4.3538]	-0.00005[-0.5059]	0.00011.2455]	0.0001[0.4466]	0.0003[1.0407]	0.000004[0.0140]	0.0002[0.7256]
$(\Phi-25)1\{\Phi \geq 26\}$	-1.2040*[-1.8100]	-0.0115***[-7.4822]	-0.0041[-1.1362]	-0.0164*[-1.8493]	-0.009[-0.7594]	-0.0150[-1.3750]	-0.0081[-0.7305]
$(\Phi-50)1\{\Phi \geq 51\}$	-0.8509***[-3.9506]	0.0120**[2.1921]	0.0116**[2.3336]	-0.0080[-1.1517]	-0.0040[-0.5177]	-0.0082[-1.0259]	-0.0046[-0.6448]
$(\Phi-75)1\{\Phi \geq 76\}$	0.2617[0.8146]	0.0088[0.7486]	-0.0058[-0.8557]	0.0026[0.3801]	-0.0026[-0.6583]	0.0040[0.5591]	-0.0014[-0.3294]
$(\Phi-100)1\{\Phi \geq 101\}$	-0.9721*** [-118.0334]	0.0326***[302.5693]	0.0329***[280.0795]	0.0118[1.3810]	0.0080[0.7561]	0.0192*[1.8843]	0.0149[1.4966]
Average number of pupils with special needs						-0.0015[-0.2957]	0.0024[0.7857]
Mother's education						0.0163***[9.4293]	0.0177***[12.5080]
ISEI index father						0.001***[6.7266]	0.0010***[8.7763]
ISEI index mother						0.004***[4.1244]	0.0004***[5.0369]
Constant	20.7586***[38.7693]	0.596***[101.7131]	0.6295***[131.5197]	0.728***[5.9698]	0.731***[4.6930]	0.6260***[4.1764]	0.628***[4.4684]
Observations	10,685	10,685	10,685	10,685	10,685	10,685	10,685
Adjusted R-squared	0.3791	0.0052	0.0024	0.0030	0.0008	0.0527	0.0732

Note: In all regressions, standard errors are clustered by enrolment level, see Lee and Card (2008).*** p<0.01, ** p<0.05, * p<0.1

Table 10: Treatment Effect (+/- 3 and 5 pupils intervals), 2008/09

					Pooled with/without observables	
	1 cut-off (25)	2 cut-off (50)	3 cut-off (75)	4 cut-off (100)	(5)	(6)
Three-pupil interval						
Dep.var: Math	0.0214[0.7572]	-0.0355* [-2.0191]	0.0117 [0.5330]	0.1483** [3.5896]	-0.0052[0.0146]	-0.0024[0.0180]
Average class-size (left side of cut-off)	22.174	23.078	23.479	23.876	NO	NO
Average class-size (right side of cut-off)	17.509	20.568	22.413	21.470	NO	NO
Observations	2703	1544	786	363	5396	5396
Three-pupil interval						
Dep.var: Italian	0.0203 [1.1878]	-0.036[-1.4364]	-0.0152 [-0.7091]	0.1137** [3.7467]	-0.0030[0.0161]	-0.000[0.0186]
Average class-size (left side of cut-off)	22.174	23.078	23.479	23.876	NO	NO
Average class-size (right side of cut-off)	17.509	20.568	22.413	21.470	NO	NO
Observations	2703	1544	786	363	5396	5396
Five-pupil interval						
Dep.var: Math	-0.0133* [-2.1755]	0.0728** [2.5789]	0.0151 [0.6742]	0.2193*** [5.6150]	0.0062[-1.0536]	-0.006[-0.8451]
Average class-size (left side of cut-off)	21.355	22.983	23.235	23.612	NO	NO
Average class-size (right side of cut-off)	16.556	20.028	22.839	21.470	NO	NO
Observations	5415	3184	1415	671	10685	10685
Five-pupil interval						
Dep.var: Italian	-0.0081[-0.7705]	0.063** [2.7906]	-0.0166[-1.1553]	0.2847*** [5.6414]	-0.0049[-0.6483]	-0.005[-0.7586]
Average class-size (left side of cut-off)	21.355	22.983	23.235	23.612	NO	NO
Average class-size (right side of cut-off)	16.556	20.028	22.839	21.470	NO	NO
Observations	5415	3184	1415	671	10685	10685

Note: In all regressions, standard errors are clustered by enrolment levels, see Lee and Card (2008). *** p<0.01, ** p<0.05, * p<0.1

Table 11: Reduced-Form Estimates of Selected Observables (sample of compliant schools), 2008/09

	ISEI index father (1)	ISEI index mother (2)	Mothers' education (3)
$1\{\Phi \geq 26\}$	-7.773760*** (2.3482)	-10.318900***(0.9757)	-0.763182***(0.1729)
$1\{\Phi \geq 51\}$	1.930775* (1.0391)	-4.608166***(0.9732)	0.021752(0.2848)
$1\{\Phi \geq 76\}$	-11.345361***(0.7926)	-6.244773(3.8475)	-0.186474*(0.1080)
$1\{\Phi \geq 101\}$	-3.084358**(1.1662)	-0.302702(1.9496)	0.072928(0.0884)
Φ	0.021136(0.0186)	-0.028598(0.0290)	0.000868(0.0014)
$(\Phi - 25)1\{\Phi \geq 26\}$	3.097997***(0.9371)	1.748179***(0.2892)	0.255036***(0.0683)
$(\Phi - 50)1\{\Phi \geq 51\}$	-0.677889*(0.3413)	0.654220**(0.2343)	-0.047910(0.1005)
$(\Phi - 75)1\{\Phi \geq 76\}$	3.965961***(0.1036)	0.570533(1.2766)	0.095906***(0.0324)
$(\Phi - 100)1\{\Phi \geq 101\}$	-0.052756***(0.0186)	-2.039584***(0.0290)	-0.235058***(0.0014)
Constant	40.956966***(0.8348)	29.367030***(1.1970)	3.347153***(0.0697)
Observations	4,111	4,111	4,111
Adjusted R-squared	0.0102	0.0140	0.0093

Note: In all regressions, standard errors are clustered by enrolment, see Lee and Card (2008). *** p<0.01, ** p<0.05, * p<0.1

.