In a Small Moment:

Cheating and Class Size in Italian Primary Schools*

Joshua D. Angrist
MIT

Erich Battistin
Queen Mary University of London

Daniela Vuri
Tor Vergata University

October 2013

Abstract

Using a Maimonides-Rule identification strategy based on class-size cutoffs around 25, we document a payoff to smaller classes in Italian primary schools. These gains are driven mainly by schools in Southern Italy, suggesting a substantial return to class size for relatively poor residents of the Mezzogiorno. In addition to low SES, however, the Mezzogiorno is distinguished by pervasive teacher cheating on standardized tests, a fact established by an experiment randomly assigning school monitors. We use Italy's Maimonides Rule to show that small classes facilitate teacher cheating, providing an alternative explanation for the causal effects of class size on test scores in Southern Italy. This motivates a causal model for achievement with two endogenous variables, class size and proportion cheating. The model is identified by a combination of class size cutoffs and randomly assigned classroom monitors. The resulting estimates suggest that the effects of class size on measured achievement in Italian primary schools are driven entirely by the relationship between class size and teacher cheating. Models that estimate class size and cheating effects jointly generate precise zeros for the former.

*Special thanks go to Patrizia Falzetti, Roberto Ricci and Paolo Sestito at INVALSI for providing the achievement data used here and to INVALSI staffers Paola Giangiacomo and Valeria Tortora for advice and guidance in our work with the data. Grateful thanks also go to Gianna Barbieri, Angela Iadecola, and Daniela Di Ascenzo at the Ministry of Education (MIUR) for access to and assistance with administrative schools data. Chiara Perricone provided expert research assistance. This research is supported by the Einaudi Institute of Economics and Finance (EIEF) - Research Grant 2011 and by the Fondazione per la Scuola della Compagnia di San Paolo in Turin. Angrist thanks the Institute for Education Sciences for financial support. The views expressed here are those of the authors alone.

1 Introduction

School improvement efforts often focus on inputs to education production, the most important of which are staffing ratios and class size. Parents, teachers, and policy makers look to small classes to boost achievement and human capital. The question of whether changes in class size have a causal effect on achievement remains controversial, however. Regression estimates often show little gain to class size reductions, with students in larger classes sometimes appearing to do better as a result (Hanushek, 1995). At the same time, a large randomized study, the Tennessee STAR experiment, generated evidence of substantial learning gains in smaller classes (Krueger, 1999). More recently, an investigation of longer-term effects of random assignment to small elementary school classrooms in the STAR experiment shows increases in college attendance for treated students (Chetty et al., 2011).

Randomized evaluations of changes in class size are exceedingly rare. Researchers have therefore turned to quasi-experimental designs in the hope of capturing class size causal effects without the benefit of random assignment. An influential design pioneered by Angrist and Lavy (1999) and Hoxby (2000) exploits the variation in class size generated by rules for classroom assignment in a regime with class size caps. Angrist and Lavy call this source of variation "Maimonides' Rule," after the 12th century Rabbinic sage who proposed a class size cap for bible study. In contemporary Italy, the setting for our study, Maimonides' Rule applies with a cap of 25 (recently increased to 27). Division of classes at integer multiples of such a cutoff generates discontinuities and changes in slope in the relationship between class size and grade enrollment, variation that potentially identify causal effects. Using Maimonides' Rule to estimate the effect of class size on achievement in Israeli elementary schools, Angrist and Lavy (1999) found returns to class on the order of those generated by the STAR experiment, though Hoxby (2000)'s estimates for Connecticut are not significantly different from zero.

Our investigation of class size effects in Italy begins with Maimonides-style estimates for the population of second and fifth graders, most of whom attend much smaller classes than those seen in Israel, where classes as large as 35 are common. The resulting estimates for Italy suggest a modest return to decreases in class size, with a 10 student reduction estimated to boost scores by about 0.05σ. Importantly, however, our estimated returns to class reductions are at least 3 times larger in data from Southern Italy than for the rest of the country; estimates for Northern and Central Italy are small and only marginally significantly different from zero, at best. This paper develops and then substantiate a simple hypothesis for these large regional differences in Italian class size effects.

Consistent with unified Italy's emergence from a diverse collection of small states in the 19th century, modern Italy is marked by wide-ranging regional differences. The South, known as the Mezzogiorno, is distinguished by persistently higher unemployment, lower per-capita income, higher crime rates, and lower educational attainment than are characteristic of Northern and Central Italian provinces.[1] The Mezzogiorno also lags the rest of Italy in financial development (Guiso et al., 2004), political accountability (Nannicini et al., 2013), and workplace productivity (Ichino and Maggi, 2000). Italy's North-South divide, which is larger and more persistent than that seen across America's Mason-Dixon line, has been linked to cultural differences and differences in residents' view of the role of government (Putnam et al., 1993).

Against a backdrop of relative under-development, the Mezzogiorno is also distinguished by widespread cheating on the standardized tests given in Italian primary schools. This can be seen in Figure 1, which reproduces regional estimates of cheating prevalence among elementary school students. These estimates come from the Italian Instituto Nazionale per la Valutazione del Sistema dell'Istruzione (INVALSI), a government agency charged with educational assessment. Cheating classes are identified through a statistical model that looks for surprisingly large average scores, low within-class variability, and suspect patterns of missing values.[2] Measured in this way, the proportion of compromised scores is about 5 percent, much as reported for Chicago elementary schools by Jacob and Levitt (2003). In Southern Italy, however, the proportion of compromised exams averages about 13 percent and reaches 25 percent in some provinces. Further evidence on cheating comes from Bertoni et al. (2013), who analyze data generated by the random assignment of classroom monitors. This experimental analysis also uncovers evidence of a substantial regional gradient in cheating.

We argue below that cheating in Italian primary schools reflects teacher behavior and, for institutional reasons, that the cost of teacher cheating is partly a time cost that increases linearly with class size (Other institutional forces that might link class size with cheating are discussed below.). We then use Maimonides' Rule to show that in regions where cheating is common, larger classes reduce the number of test scores that appear to have been adulterated. A juxtaposition of regional patterns in the causal effects of class size on achievement and the causal effects of class size on cheating uncovers striking parallels in the two empirical relationships. This parallelism leads us to

[1]The Mezzogiorno consists of the administrative regions of Basilicata, Campania, Calabria, Puglia, Abruzzo, Molise, and the islands of Sicily and Sardinia. Italy's 20 Administrative regions are further divided into over 100 provinces.

[2]The INVALSI testing program is described below and in INVALSI (2010). The INVALSI cheating variable identifies classes with substantially anomalous score distributions, coding a probability of cheating for each (see Quintano et al. 2009). Figure 1 uses this variable for the 2009-11 scores of second and fifth graders.

model achievement as a function of two endogenous variables, class size and cheating. The model is identified by a combination of Maimonides' Rule and randomly assigned classroom monitors. The resulting estimates suggest that the relationship between class size and INVALSI assessments is explained entirely by cheating. Double-endogenous variable models produce precise zeros for non-cheating-related class size effects.

Why is the fact that cheating explains measured returns to class size important? The Maimonides Rule research design is motivated by an effort to quantify causal effects of class size in observational studies, without resorting to costly and time consuming random assignment. The design is not guaranteed to work; Urquiola and Verhoogen (2009) show how endogenous sorting by students induces selection bias in comparisons across class size caps in Chilean private schools. By contrast, our analysis uncovers a substantive problem inherent in analyses of the causal effects of class size, regardless of design. We show that even where class size boosts measured achievement, this need not signal increased human capital. Our findings also provide evidence of a surprisingly large behavioral response to what would seem to be weak incentives. Italian teachers work in a highly regulated public sector, with virtually no risk of termination, and are subject to a pay and promotion structure that's largely independent of performance. Even so, and in spite of the fact that cheating is probably not effortless, large numbers of Italian teachers seem to find cheating worthwhile. The question of why this is, and why the effort to cheat is especially likely to be deemed worthwhile by those who teach in the South, is an intriguing topic for further study.

The rest of the paper is organized as follows. The next section presents some institutional background on Italian schools and achievement tests, and explains how the tests are scored. Section 3 describes our data and documents the Maimonides first stage. Following a brief graphical analysis, Section 4 reports Maimonides-style estimates of effects of class size on achievement and cheating, while Section 5 summarizes results from the monitoring experiment. Section 6 reports jointly estimated class size and cheating effects. This section also considers the econometric implications of mismeasured and underestimated cheating rates for our empirical strategy. Section 7 concludes.

2 Background

Italian Schools and Tests

Primary and Secondary schooling in Italy are compulsory from ages 6 to 16, with three stages: 5 years of elementary school (*scuola elementare*), lower secondary school covering grades 6-8 (*scuola media*), and high school (*scuola superiore*), which runs for 3-5 years. Schools are organized into

single- or multi-unit "institutions," much as a single campus might house more than one school in American public systems (a local example is the South Boston Educational Complex, which houses a number of public high schools).

Italian schools have long used matriculation exams for tracking and placement in the transition from elementary to middle school and throughout high school, but standardized testing for evaluation purposes is a recent development. In 2008, INVALSI piloted voluntary assessments in elementary school; in 2009 these became compulsory for all schools and students. INVALSI assessments cover mathematics and Italian language skills in a national administration lasting two days in the Spring.[3] Tests are proctored by local administrators and teachers. Proctors and other teachers are expected to copy students' original responses onto machine-readable answer sheets (*scheda risposta*), which are then sent to INVALSI. Copying is not passive; some of this processing requires teachers to interpret a student's original response as being correct, incorrect, or missing. We give examples in a brief appendix with sample test items and score sheets. This procedure opens the door to score manipulation, as does the fact that proctors are local staff members, who may feel they have an interest in boosting their students' scores.

Related work

Maimonides'-Rule type empirical strategies have been used to identify class size effects in many countries, including the US (Hoxby, 2000), France (Piketty, 2004 and Gary-Bobo and Mahjoub, 2006), Norway (Bonesronning, 2003 and Leuven et al., 2008) and the Netherlands (Dobbelsteen et al., 2002). On balance these results point to modest returns, mostly lower than those reported in Angrist and Lavy (1999) for Israel. A natural explanation for relatively large Israeli findings is the unusually large classes characteristic of Israeli elementary schools. In line with this view, Woessmann (2005)'s finds a mostly weak association between class size and achievement in a cross-country panel covering 17 Western European school systems.

The returns to class size in Italy have received little attention from researchers to date, partly because large representative samples of public school students with achievement data have only recently become available. One of the few existing studies, Bratti et al. (2007), reports regression estimates showing an insignificant class size effect. In an aggregate analysis, Brunello and Checchi (2005) look at the relationship between staffing ratios and educational attainment for cohorts born before 1970; they find that higher pupil-teacher ratio at the regional level are associated with higher average schooling attainment. We have yet to locate other quantitative explorations of Italian class

[3]For a summary of test results and additional background, see http://www.invalsi.it.

size, though Ballatore et al. (2013) recently use a similar sample to estimate the effects of immigrants in the classroom on native achievement.

The monitoring experiment used here to identify the effects of cheating and class size jointly was first analyzed by Bertoni et al. (2013). We replicate some of their earlier findings, while also adjusting for potentially confounding features of the monitoring experiment's sample design that weren't fully accounted for in previous work. Failure to make these adjustments leads to covariate imbalance across treatment and control groups in the monitoring experiment; our analysis corrects this problem. The resulting estimates suggest that the presence of classroom monitors sharply reduce cheating, and that cheating boosts measured scores dramatically. Both of these effects are markedly larger in Southern Italy. Elsewhere in Italy, cheating is a marginal phenomenon.

3 Data and First Stage

Data and descriptive statistics

The assessments used in this study come from INVALSI's testing program in Italian elementary schools, with score data from the years 2009/10, 2010/11 and 2011/12. Raw test scores capture the number of correct answers; for the purposes of regression and two-stage least squares (2SLS) estimation, we standardized these by subject, year of survey and grade to have zero mean and unit variance. We matched INVALSI data on test scores to administrative and survey information describing institutions, schools, classes, and students. Class size can be measured by administrative enrollment counts at the beginning of the school year as well as the number of test-takers (we use the former). Student data include gender, citizenship, and information on parents' employment status and educational background. These data are collected as part of the test administration and meant to be provided by public school staff when scores are submitted. The fewer than 10 percent of Italian primary and secondary school students who attend private schools are omitted from our study.

Our statistical analysis focuses on class-level averages since this is the aggregation level at which the regressor of interest varies. The empirical analysis is restricted to classes with more than the minimum number of students set by law (10 before 2010 and 15 from 2011). This selection eliminates classes in the least populated areas of the country, e.g. mountainous areas and small islands. We also drop schools enrolling more than 160 students in a grade, as these are above the threshold where Maimonides' Rule is likely to matter (this size cutoff trims classes above the 99th percentile of the enrollment-weighted class size distribution).

The resulting matched file includes about 140,000 2nd and 5th grade classes (attended by about

1.3 million students in each grade). Table 1 reports descriptive statistics separately by grade. Statistics are reported at the class level in Panel A, at the school level in Panel B, and at the institution level in Panel C. Class size averages about 20 in both grades, and is slightly lower in the South. The distribution of test scores, shown in Panel A, refers to the class average percent correct. Scores are higher in language than in math and higher in 5th grade than in 2nd grade. The table also shows averages for INVALSI's class-level probability of cheating and for a similar cheating variable that we've constructed (Section 3, below, describes our procedure). Descriptive statistics for the two measures are similar, and indeed, the two cheating variables are highly correlated, with a correlation coefficient above .95.

Maimonides in Italy

Our identification strategy exploits minimum and maximum class sizes for Italy (these rules are a consequence of a regulation known as *Decreto Ministeriale 331/98*). Until the 2008/09 school year, primary school classes were subject to a minimum size of 10 and a maximum of 25. Grade enrollment beyond 25 or a multiple thereof usually prompted the addition of a class. The rule allows exceptions, however. Principals can reduce the size of classes attended by one or more disabled students, and schools in mountainous or remote areas are allowed to retain classes with fewer than 10 students. Finally, the law allows an unrestricted deviation of 10% from the maximum in either direction (that is, the Ministry of Education would have funded an additional class when enrollment exceeds 22 and required a new class when average enrollment otherwise exceeded 28). A 2009/10 reform of the Italian increased the maximum size to 27, with a minimum size of 15, again with a tolerance of 10% (promulgated through *Decreto del Presidente della Repubblica 81/2009*). This reform was rolled out in one grade per year, starting with first grade. In our data, second graders in 2009/10 and fifth graders in any year are subject to the old rule, while second graders in 2010/11 and 2011/12 are subject to the new rule.

Ignoring discretionary deviations near cutoffs, Maimonides' Rule predicts the size of class i in grade g at school k in year t, denoted f_{igkt}, to be the following non-linear and discontinuous function of enrollment:

$$f_{igkt} = \frac{e_{gkt}}{[int\left(\left(e_{gkt}-1\right)/c_t\right)+1]}, \tag{1}$$

where e_{gkt} is beginning-of-the-year grade enrollment at school k, c_t is the cap in effect that year ($c_t =$ 25 or 27), and $int(x)$ is the largest integer smaller than or equal to x. Figures 2 and 3 plot average class size against f_{igkt}, separately for pre- and post-reform periods. The x-axis in these figures is

enrollment in grade, while the solid line shows equation 1. Plotted points show the average class size at each value of enrollment. Actual class size follows predicted class size reasonably closely for enrollments below about 75, especially in the pre-reform period. Theoretical sharp corners in the class size/enrollment relationship are rounded as a result of the fact that the Italian Maimonides' rule allows for discretionary breaks within a 10% band. Many classes are split before reaching the theoretical maximum of 25, and earlier-than-mandated splits occur more often as enrollment increases. In the post-reform period, class size tracks the rule generated by the new cap of 27 poorly once enrollment exceeds about 70.

Cheating Data and Cheating Behavior

The cheating variable used here was constructed as a function of the within-class average and standard deviation of test scores, the number of missing items, and a Herfindahl index of the share of students with similar response patterns. These indicators are used to flag as suspicious those classes with abnormally high performance, an unusually small dispersion of scores, an unusually low proportion of missing items, and a high concentration of response patterns. This procedure yields class-level indicators of compromised scores, separately for math and language. Our procedure is similar to that used by Quintano et al. (2009), which, for the period considered in our empirical exercise, provides the cheating data used in INVALSI publications (INVALSI, 2010). The INVALSI version generates a continuous class-level probability that test scores from a given classroom are compromised by cheating. Our modified procedure generates a dummy variable indicating classrooms where cheating is highly likely; this implicitly becomes a probably in our 2SLS procedure. Methods and formulas used to determine cheating are given in the appendix.[4]

Why does class size affect cheating? First, cheating is facilitated by the need for local proctors and teachers to copy students' original answer sheets onto the machine readable *scheda risposta*. This task, which requires copiers to make judgements about which answers are correct or missing, must be completed by the end of the day on which tests are taken. The fact that some questions are open response also suggests that response adjustment requires some thought and interpretation. The extent of the necessary interpretation is documented with sample questions and a copy of the *scheda risposta* in the Appendix. In larger classes, teachers may not have time to make as many adjustments to scores as when there are fewer students. Second, class size may be related to cheating as a result of the fact that exams are locally proctored. Although it seems unlikely that teachers or

[4]Our procedure follows Jacob and Levitt (2003) in inferring cheating from patterns of answers within and across tests in a classroom. Jacob and Levitt (2003) also compare test scores over time, looking for anomalous changes. Values in the upper tail of the Jacob-Levitt suspicious answer index are highly predictive of their cheating variable.

proctors announce correct answers to an entire classroom at once, they may offer individual students assistance. Finally, a large class increases the risk that cheating behavior will be exposed. The fact that external monitoring reduces cheating suggests those doing it see exposure as undesirable.

4 Effects of Class size on Achievement and Cheating

Graphical Analysis

We begin with a sequence of plots that capture class size effects near enrollment cutoffs without the need for parametric assumptions. The first plot in this sequence, Figure 4, documents the relationship between cutoffs (multiples of 25 or 27) and class size. This figure was constructed from a sample of classes at schools with enrollment that fall in a [-12,12] window around the first four cutoffs shown in Figures 2 and 3. Enrollment values in each window were centered to be zero at the relevant cutoff. The y-axis shows average class size (standardized to be mean zero) conditional on the centered enrollment value shown on the x-axis, and reported as a 3-point moving average. Figure 4 also plots fitted values generated by LLR smoothing for points 3 or more units away from the cutoff on either side. In this context, the LLR smoother uses data on one side of the cutoff only, with an edge kernel and the bandwidth derived by Imbens and Kalyanaraman (2012) used for smoothing.[5]

In view of the 2-3 student tolerance around the stated threshold for opening a new class, enrollment within two points of the cutoff is excluded from the local linear fit. As a result of this tolerance, class size can be expected to decline at enrollment values shortly before the cutoff and to continue to decline thereafter. Consistent with this expectation, the figure shows a clear drop at the cutoff, with the sharpness of the break moderated by values near the cutoff. Class size is minimized at about 3-5 students to the right of the cutoff instead of immediately after, as we would expect were a sharp rule to be tightly enforced. The parametric identification strategy detailed below exploits both the discontinuous and nonlinear variation in class size apparent in Figure 4. Looking only at points close to the cutoff, the change in size generating by moving across a cutoff appears to be on the order 2-3 students.

Using data from the South, math and language scores plotted as a function of enrollment values near Maimonides cutoffs show a jump that mirrors the drop in size seen at Maimonides cutoffs, but there is little evidence of such a jump in schools outside the South. This pattern is documented in Figure 5, which plots math and language scores against enrollment in a format paralleling that used

[5]Specifically, we construct fitted values by applying Stata's lpoly command to class-level data.

to construct Figure 4. The reduced-form achievement drop for schools in Southern Italy is about 0.02 standard deviations (hereafter, σ). Assuming this change in test scores in the neighborhood of Maimonides cutoffs is driven by a causal class size effect, the implied return to a one-student reduction in class size is about $.01\sigma$ in Southern Italy. The absence of a jump in scores at cutoffs in data from schools elsewhere in the country implies that outside the South, class size reduction leave scores unchanged.

Cheating behavior also varies as a function enrollment in the neighborhood of class size cutoffs, with a pattern much like that seen for achievement. This is apparent in Figure 6, which puts the proportion of classes identified as having compromised scores on the y-axis, in a format like that used for Figures 4 and 5. Mirroring the pattern of achievement effects, a discontinuity in cheating behavior emerges only for schools in Southern Italy. This pattern leads us to explore the possibility that the achievement gains that appear to be generated by class size in Figure 5 reflect the cheating behavior captured in Figure 6. Our exploration comes in the form of a multivariate model that allows for both gains in learning and increased cheating when class size shrinks.

Before turning to estimation of a model with multiple causal channels, we consider possible threats to validity in our Maimonides-Rule research design. Changes in student characteristics around the discontinuity threshold represent an important threat to the identification of causal effects in such designs. As discussed by Urquiola and Verhoogen (2009) and Baker and Paserman (2013), parents may be aware of class size caps, choosing to enroll in schools where expected class size is smaller. Likewise, school leaders may try to shift or limit enrollment so as to avoid costs related to changes in class size. Such behavior potentially introduces a type of selection bias. Enrollment-based sorting seems unlikely in Italy, however, since public school students are required to attend local schools and mobility across school districts is modest.[6] Figure 7 supports the view that enrollment variation in Italian schools is untainted by strategic behavior. This figure plots the percentage of female students, the children of immigrants, fathers' education as measured by the proportion of high school graduates, and maternal employment against enrollment, using the same format as used to construct the reduced forms plotted in Figures 5 and 6. Figure 7 offers no evidence of discontinuities in student composition near class size cutoffs.

Empirical Framework

Figures 5 - 7 suggest that variation in class size near the cutoffs or corners induced by Maimonides rule non parametrically identify class size effects. As a practical matter, however, we adopt a

[6]This is similar to the institutional setting described by Angrist and Lavy (1999).

parametric framework. This allows us to efficiently incorporate variation in enrollment a bit farther away from cutoffs, while also exploiting changes in the slope of the relationship between enrollment and class size well away from cutoffs. Finally, a parametric framework easily accommodates models with multiple endogenous variables.

Our parametric framework models y_{igkt}, the average outcome score in class i in grade g at school k in year t, as a polynomial function of the running variable, r_{igkt}, and class size, s_{igkt}. With quadratic running variable controls, the specification pooling grades and years can be written,

$$y_{igkt} = \rho_0(t, g) + \beta s_{igkt} + \rho_1 r_{igkt} + \rho_2 r_{igkt}^2 + \epsilon_{igkt}, \tag{2}$$

where $\rho_0(t, g)$ is shorthand for a full set of year and grade effects. Because their inclusion generates a modest increase in precision, this model also controls for the demographic variables described in Table 2, as well as the stratification variables used in the monitoring experiment described below.[7] Standard errors are clustered by institution, which we reckon to be a conservative strategy in this context.

The instrument for 2SLS estimation of equation (2) is f_{igkt}, as defined in equation (1). To document sensitivity of findings to specification of running variable controls, we also report results from a specification estimated in a sample that includes only schools with enrollment falling in mostly symmetric windows centered around each cutoff. These models - which we refer to as the interacted specification - include a full set of segment (window) main effects while allowing the quadratic control function to differ across segments.[8] The corresponding OLS estimates for models without interacted running variable control are shown as a benchmark.

Parametric Estimates of Class Size Effects

OLS estimates of equation (2) show a marked negative correlation between class size and achievement for schools in the Northern and Central regions, but not in the South. Larger classes are associated with higher language achievement in the South while Southern class size appears unrelated to achievement in math. These findings can be seen seen in columns 1-3 of Table 2.

2SLS estimates using Maimonides' Rule, reported in columns 4-9 of Table 2, suggest larger classes

[7] Control variables include proportion female, the proportion of students behind and ahead of their grade level, the proportion of immigrants, the proportion of students whose father is a high school graduate, have unemployed mothers, mothers not in the labor force, and dummies for missing values for these variables. Stratification controls consist of total enrollment in grade, region dummies, and the interaction between enrollment and region.

[8] Pre-reform segments cover the intervals 10-37, 38-62, 63-87, 88-112, 113-137, and 138-159; post-reform segments cover the intervals 15-40, 41-67, 68-94, 95-121, and 122-159. These segments cover mostly symmetric intervals of width +/- 12 in the pre-reform period and +/-13 in the post-reform period, with modifications at the lower and upper segments to include a few larger and smaller values. A few schools with enrollments beyond the rightmost segment boundary, which marks the upper percentile of the enrollment distribution, were excluded.

reduce achievement in both math and language. The associated first stage estimates, reported in Appendix Table A.1 show predicted class size increases actual class size with a coefficient around one-half when regions are pooled, with a first stage effect of 0.43 in the South and 0.55 elsewhere. 2SLS estimates for Southern schools, implying something on the order of a 0.10σ achievement gain for a 10-student reduction, are 2-3 times larger than the corresponding estimates for schools outside the South. The 2SLS estimates for both regions are reasonably precise; only estimates of the interacted specification for math scores from non-Southern schools fall short of conventional levels of statistical significance. On balance, the results in Table 2 indicate a substantial achievement payoff to class size reductions, though the gains here are not as large as those reported by Angrist and Lavy (1999) for Israel. A substantive explanation for this difference in findings might be concavity in the relationship between class size and achievement combined with Italy's smaller overall class sizes.

The estimates in Table 3 suggest the causal effect of class size on measured achievement in Italy need not reflect more learning in smaller classes. This table reports estimates from specifications identical to those used to construct the estimates in Table 2, with the modification that a class-level cheating indicator replaces achievement as an outcome. The 2SLS estimates in columns 4-9 show a large and precisely-estimated negative effect of class size on cheating rates, with effects on the order 4-5 percentage points for a 10-student class size increase in the South. Estimates for schools outside the South also show a negative relationship between class size and cheating, though here the estimated effects are much smaller and not significantly different from zero. Interestingly, OLS estimates of effect of class size on cheating are largely in line with those generating by 2SLS. This suggests that the relationship between class size and cheating is largely mechanical, and therefore uncontaminated by the sort of selection bias that affects OLS estimates of the corresponding achievement relation. We might therefore put some stock in the OLS estimate of the effect of class size on cheating outside the South. This ranges from -0.0007 to -0.0012 for math and language, effects too small to be found significant when estimated using IV but in the same ballpark as the corresponding IV estimates reported in columns 5 and 8.

5 The Monitoring Experiment

The estimates in Table 3 take cheating to be an outcome variables in a notional class size experiment, but we're ultimately interested in cheating as a causal channel or a control variable in a multivariate model that simultaneously links class size, cheating, and achievement. We therefore turn to INVALSI's monitoring experiment as an independent source of variation in cheating, unrelated to

Maimonides' Rule. In an effort to increase test reliability, INVALSI randomly selected institutions to be observed by an external monitor using a two-stage sampling procedure. Institutions were first sampled with a probability proportional to grade enrollment in the year of the test. Sampling was also stratified by regions. Within sampled institutions, classroom monitors were meant to be randomly assigned to one or two classes per grade, though randomness of within-institution monitoring appears to have been compromised.

Monitors were selected by regional education offices from a pool of mainly retired teachers and principals, who had not worked in towns or at schools they were assigned to monitor for at least two years. Both testing and score sheet transcription were monitored. The presence of an external monitor should therefore have discouraged both student and teacher cheating. Tests without monitors were proctored by local school staff, though the Math teacher for a given grade should not have proctored that grade's test and so on. The effect of monitoring on cheating in this experiment was first reported by INVALSI (2010). We replicate some of these earlier findings, as well as those reported in a related study by Bertoni et al. (2013).[9]

As a preliminary step, Table 4 documents balance across institutions with and without randomly assigned monitors. This table reports regression-adjusted treatment-control differences from models that include sampling strata controls based on the experimental design. Specifically, these specifications include a full set of region dummies and a linear function of institutional grade enrollment that varies by regions. Standard errors are clustered by institution. Administrative variables - collected routinely by the central government - are well-balanced across groups, as can be seen in the small and mostly insignificant coefficient estimates reported in Panel A.

Demographic data and other information provided by school staff, such as instruction time, show evidence of imbalance. This seems likely to reflect a causal data-quality effect, rather than a problem with the experimental design or implementation. The hypothesis that monitors induced more careful reporting by staff is supported by the large treatment-control differential in missing data rates documented at the bottom of the table. Among other salutary effects, randomly assigned monitors reduced item non-response by as much as three percentage points.

The presence of institutional monitors reduced cheating considerably. This can be seen in columns 1-3 of Table 5, which report an estimated effect of the presence of monitors on cheating rates of about 3 percentage points for Italy, with estimates twice as large in the South. These estimates come from models similar to those used to check covariate balance, with a cheating indicator replacing covariates as the dependent variable. Monitoring also reduced language scores

[9]Bertoni et al. (2013) incorrectly analyzed institutions as schools, a mistake that leads to evidence of imbalance between treatment and control groups in their work.

by 0.08σ, while the estimated monitoring effect on math scores is about -0.11σ. Here too, effects of monitoring were much larger in the South, ranging from 0.13σ for math to 0.18σ for language, estimates that appear in column 6 of the table.

The estimates reported in columns 1-3 and 4-6 of Table 5 constitute the first stage and reduced form for a model that uses the assignment of monitors as an instrument for the effects of cheating on test scores. Dividing reduced form estimates by the corresponding first stage estimates produces second stage cheating effects of about 3σ for the Southern, with even larger second stage estimates for the North. These effects seem implausibly large, implying a boost in scores that exceeds the range of the dependent variable in some cases. It also seems likely, however, that the cheating variable used to construct the corresponding first stage effects is substantially mismeasured. If this measurement error attenuates first stage estimates, the resulting second stage estimates are proportionally inflated. The analysis that follows therefore considers the consequences of measurement error in cheating for the empirical strategy that simultaneously captures cheating and class size effects on measured achievement.

6 Effects of Class Size and Cheating on Achievement

The estimates in Tables 2 and 5 motivate a causal model in which academic achievement depends on class size (s_{igkt}) and cheating (c_{igkt}), both treated as endogenous variables to be instrumented. This can be written:

$$y_{igkt} = \rho_0(t,g) + \beta_1 s_{igkt} + \beta_2 c_{igkt} + \rho_1 r_{igkt} + \rho_2 r_{igkt}^2 + \eta_{igkt}, \tag{3}$$

where $\rho_0(t,g)$ is again shorthand for year and grade effects. We interpret equation 3 as describing the average achievement that would be revealed by alternative assignments of class size, s_{igkt}, in an experiment that holds c_{igkt} fixed. This model likewise describes causal effects of changing cheating rates in an experiment that holds class size fixed. In other words, 3 is a model for potential outcomes indexed against two jointly manipulable treatments.

We estimate equation 3 by 2SLS in a setup that includes the same covariates as were included in the models used to construct the estimates reported in Table 2. The instrument list in this case includes Maimonides' Rule (f_{igkt}) and randomly assigned monitors, m_{igkt}. The first-stage equations with these two instruments can be written,

$$s_{igkt} = \lambda_{10}(t,g) + \mu_{11} f_{igkt} + \mu_{12} m_{igkt} + \lambda_{11} r_{ik} + \lambda_{12} r_{ik}^2 + \xi_{ik} \tag{4}$$

$$c_{igkt} = \lambda_{20}(t,g) + \mu_{21} f_{igkt} + \mu_{22} m_{igkt} + \lambda_{21} r_{igkt} + \lambda_{22} r_{igkt}^2 + \upsilon_{ik} \quad (5)$$

where $\lambda_{10}(t,g)$ and $\lambda_{20}(t,g)$ are shorthand for year and grade effects in the two first stages. Estimates of these first stage equations, reported in Table (6), replicate earlier findings in showing both a monitoring and a Maimonides Rule effect on cheating, both of which are considerably more pronounced in the South. The Maimonides first stage for class size remains at around one-half, while the presence of a classroom monitor seems largely unrelated to class size. This is consistent with the hypothesis that monitors were randomly assigned.

OLS estimates of equation 3, reported in columns 1-3 of Table 7, show a strong positive association between cheating and achievement in both math and language, an effect of about the same size in Southern and non-Southern schools. Cheating appears to increase math scores more than language scores. Larger classes are also positively associated with math and language achievement in the Mezzogiorno, though not elsewhere.

IV estimation generates evidence of large effects of cheating, not unlike those discussed in the context of the estimates of Table 5. This can be seen in the 2SLS estimates of β_2,reported in columns 4-9 of Table 7. This table also shows small and mostly statistically insignificant estimates of β_1, the coefficient on class size in the multivariate model. In an effort to boost the precision of these estimates, we estimated over-identified models that add two dummies for values of the running variable that fall within 10% of each cutoff, a specification motivated by the nonparametric first stage captured in Figure4.[10] The most precise of the estimated zeros reported in Table 7, generated by the over-identified specification for Italy as a whole, run no larger than .0024, with an estimated standard error of 0.0016; these appear in column 7. The p-values generated by the over-identification test statics associated with these models offer no evidence against the underlying exclusion restriction.

The most important results in Table 7 are the small and insignificant class size effects for Southern Italy, findings that contrast with the much larger and statistically significant class size effects for the Mezzogiorno reported in Table 2. In column 9 of the latter table, for example, a 10 student reduction in class size is estimates to boost achievement by 0.10 or more. The corresponding multivariate estimates in Table 7 are of the opposite sign, showing that larger classes increase achievement, though not by very much. These estimates come with estimated standard errors ranging from about 0.02 to 0.04, so that the estimate class size effects in Table 2 fall well outside the estimated confidence intervals associated with the multivariate estimates. It seems reasonable, therefore, to interpret the

[10]First stage estimates for the over-identified model appear in Appendix Table A.2.

estimated class effects in Table 7 as precise zeros. This in turn aligns with an interpretation of the return to class size generated by Maimonides Rule-type instruments applied to data from Southern Italy as due entirely to the causal effect of class size on cheating, most likely by teachers.

Cheating with Misclassification

A possible threat to the validity of our interpretation of the results in Table 7 is measurement error in cheating. We show here that as long as misclassification rates are independent of the instruments, mismeasurement of cheating leaves 2SLS estimates of class size effects in the multivariate model unaffected. We show this in the context of a simplified version of the multivariate model, which can be written with a class subscript as

$$y_i = \rho_0 + \beta_1 s_i + \beta_2 c_i^* + \zeta_i, \tag{6}$$

where instruments are assumed to be uncorrelated with the error, ζ_i, as in equation (3). Here, c_i^* is an accurate cheating dummy for class i, while c_i is observed cheating as before.

Let $z_i = [f_i \, m_i]'$ denote the vector of instruments. Assuming classification rates are independent of the instruments conditional on c_i^*, we can write

$$c_i = (1 - \pi_0) + (\pi_0 + \pi_1 - 1)c_i^* + \omega_i, \tag{7}$$

where the residual, ω_i, is defined by

$$\omega_i = c_i - E[c_i | z_i, c_i^*],$$

and the probability that cheating is correctly measured satisfies

$$P[c_i = d | z_i, c_i^* = d] = P[c_i = d | c_i^* = d] = \pi_d, \tag{8}$$

for $d = 0, 1$. Note that $E[z_i \omega_i] = 0$ by definition of ω_i.

Using (7), equation (6) can be rewritten

$$y_i = \left[\rho_0 - \frac{\beta_2(1 - \pi_0)}{\pi_0 + \pi_1 - 1}\right] + \beta_1 s_i + \left[\frac{\beta_2}{\pi_0 + \pi_1 - 1}\right] c_i + \left[\zeta_i - \beta_2 \frac{\omega_i}{\pi_0 + \pi_1 - 1}\right]. \tag{9}$$

We assume that the π_d's are strictly greater than *0.5*, so that reported cheating is a better indicator of actual cheating than a coin toss. This ensures that the coefficient on c_i in (9) is finite and has the same sign as β_2. The 2SLS estimate of the coefficient on reported cheating is therefore upward biased, since $\pi_0 + \pi_1 - 1$ is strictly between 0 and 1 given these assumptions. Most importantly, because the feasible estimating equation, (9), has a residual uncorrelated with the instruments, and the coefficient on class size is unchanged in this model, misclassification of the sort described by (8)

leaves estimates of the class size coefficient, β_1, unchanged.

7 Summary and Directions for Further Work

The causal effects of class size on primary school test scores are identified by quasi-experimental variation arising from Italy's version of the Maimonides Rule enrollment formula first used as a research tool by Angrist and Lavy (1999). The resulting identification strategy shows small classes boosting test scores in Southern provinces, an area known as the Mezzogiorno, but not elsewhere. Analyses of imputed cheating measures and a randomized classroom monitoring experiment reveal substantial cheating in the Mezzogiorno, most likely on the part of teachers. For a variety of institutional and behavioral reasons, teacher cheating is inhibited by larger classes as well as by monitoring. Finally, estimates of a model that jointly captures the causal effects of class size and cheating on measured achievement show the returns to class size in the Mezzogiorno are explained by the causal effects of class size on cheating, with no apparent change in learning.

Interesting and as-yet unanswered questions raised by these findings include the questions of why Italian teachers find cheating worthwhile when their careers and pay are largely divorced from student outcomes, and why teacher cheating is so much more prevalent in the South. On the policy side, it seems worth asking whether teacher cheating can be eliminated by a simple institutional change related to the way score sheets are handled, or whether the underlying problem is behavioral. Recent and pending changes in Italian test administration may provide evidence on this point. No less important, we wonder why there appears to be no improvement in learning due to smaller classes in Italy, while evidence from the US, Israel, and a number of other countries credibly suggests class size reductions facilitate learning. We hope to develop answers to these questions in future work.

Table 1: Descriptive statistics

	2 grade			5 grade		
	Italy	North/Centre	South	Italy	North/Centre	South
	(1)	(2)	(3)	(4)	(5)	(6)
A. Class characteristics						
female	0.47	0.48	0.47	0.49	0.49	0.49
	(0.13)	(0.12)	(0.14)	(0.11)	(0.11)	(0.11)
behind grade level	0.02	0.02	0.02	0.03	0.04	0.02
	(0.04)	(0.04)	(0.03)	(0.05)	(0.05)	(0.04)
ahead grade level	0.01	0.00	0.02	0.01	0.01	0.03
	(0.03)	(0.02)	(0.04)	(0.04)	(0.02)	(0.05)
immigrant 1st generation	0.03	0.04	0.01	0.05	0.07	0.02
	(0.05)	(0.06)	(0.03)	(0.07)	(0.08)	(0.04)
immigrant 2nd generation	0.07	0.09	0.02	0.05	0.07	0.02
	(0.09)	(0.11)	(0.04)	(0.07)	(0.08)	(0.04)
father HS diploma	0.26	0.27	0.25	0.25	0.26	0.24
	(0.17)	(0.16)	(0.18)	(0.17)	(0.16)	(0.17)
mother employed	0.45	0.54	0.31	0.45	0.53	0.31
	(0.27)	(0.26)	(0.22)	(0.26)	(0.26)	(0.21)
pct correct: math	47.90	46.06	51.08	64.16	63.26	65.61
	(14.63)	(12.89)	(16.75)	(12.91)	(10.87)	(15.53)
pct correct: language	69.78	69.18	70.81	74.22	74.27	74.12
	(10.92)	(9.18)	(13.34)	(8.90)	(7.47)	(10.82)
class size	20.12	20.25	19.90	19.66	19.90	19.28
	(3.40)	(3.35)	(3.48)	(3.72)	(3.67)	(3.76)
instruction time <= 30 hours p/w	0.61	0.49	0.81	0.63	0.50	0.85
	(0.49)	(0.50)	(0.39)	(0.48)	(0.50)	(0.36)
instruction time >30 hours p/w	0.36	0.49	0.13	0.33	0.47	0.10
	(0.48)	(0.50)	(0.34)	(0.47)	(0.50)	(0.31)
cheating: math	0.06	0.02	0.14	0.07	0.02	0.14
	(0.24)	(0.13)	(0.35)	(0.25)	(0.15)	(0.34)
cheating: language	0.05	0.02	0.11	0.06	0.02	0.11
	(0.23)	(0.14)	(0.31)	(0.23)	(0.15)	(0.31)
cheating (INVALSI): math	0.07	0.02	0.14	0.06	0.03	0.13
	(0.20)	(0.10)	(0.28)	(0.19)	(0.10)	(0.26)
cheating (INVALSI): language	0.06	0.03	0.11	0.06	0.03	0.10
	(0.17)	(0.10)	(0.24)	(0.18)	(0.11)	(0.24)
N	67,453	42,747	24,706	72,536	44,739	27,797
B. School characteristics						
number of classes	1.95	1.87	2.11	1.94	1.85	2.10
	(1.10)	(1.01)	(1.27)	(1.10)	(0.98)	(1.28)
enrollment	40.52	38.83	43.83	38.87	37.32	41.69
	(25.16)	(22.99)	(28.64)	(25.19)	(22.76)	(28.91)
N	34,591	22,863	11,728	37,476	24,225	13,251

Table 1: Descriptive statistics (cont.)

	2 grade			5 grade		
	Italy	North/Centre	South	Italy	North/Centre	South
	(1)	(2)	(3)	(4)	(5)	(6)
			C. Institution characteristics			
number of schools	2.00	2.32	1.57	2.10	2.42	1.69
	(1.05)	(1.13)	(0.74)	(1.09)	(1.17)	(0.81)
number of classes	3.89	4.33	3.31	4.07	4.48	3.55
	(1.97)	(1.95)	(1.85)	(1.95)	(1.91)	(1.88)
enrollment	86.00	95.33	73.68	85.17	94.02	73.88
	(40.61)	(39.52)	(38.71)	(40.45)	(39.11)	(39.31)
external observer: 2009/10	0.22	0.20	0.23	0.22	0.20	0.23
	(0.41)	(0.40)	(0.42)	(0.41)	(0.40)	(0.42)
external observer: 2010/11	0.22	0.20	0.23	0.21	0.20	0.23
	(0.41)	(0.40)	(0.42)	(0.41)	(0.40)	(0.42)
external observer: 2011/12	0.22	0.20	0.23	0.21	0.20	0.23
	(0.41)	(0.40)	(0.42)	(0.41)	(0.40)	(0.42)
N	17,333	9,866	7,467	17,830	9,997	7,833

"Mean" and "s.d." for class characteristics are computed using one observation per class; "Mean" and "s.d." for school characteristics are computed using one observation per school; "Mean" and "s.d." for institutions are computed using one observation per institution. Categories for educational qualification of father and occupation of mother do not add up to one because of missing values.

Table 2: OLS and IV/2SLS estimates of the effect of class size on test scores

	OLS			IV/2SLS					
	Italy	North/Centre	South	Italy	North/Centre	South	Italy	North/Centre	South
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					A. Math				
Class size	-0.001	-0.002***	0.001	-0.005***	-0.005***	-0.010***	-0.006***	-0.004**	-0.013***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.002)	(0.002)	(0.005)
enrollment	✗	✗	✗	✗	✗	✗	✗	✗	✗
enrollment squared	✗	✗	✗	✗	✗	✗	✗	✗	✗
interactions							✗	✗	✗
N	140,010	87,498	52,512	140,010	87,498	52,512	140,010	87,498	52,512
					B. Language				
Class size	0.0004	-0.002***	0.003**	-0.004***	-0.003***	-0.007**	-0.004**	-0.002	-0.010**
	(0.0006)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.001)	(0.004)
enrollment	✗	✗	✗	✗	✗	✗	✗	✗	✗
enrollment squared	✗	✗	✗	✗	✗	✗	✗	✗	✗
interactions							✗	✗	✗
N	140,010	87,498	52,512	140,010	87,498	52,512	140,010	87,498	52,512

Notes: Columns 1-3 report OLS estimates of the effect of class size on scores. Columns 4-9 report 2SLS estimates using Maimonides' rule as instrument. The unit of observation is the class. Models listed as controling for interactions are estimated in a sample limited to enrollments falling in a +/- 12 window around Maimonides cutoffs, and allow the quadratic control function to vary across windows. Robust standard errors, clustered on school and grade, are shown in parentheses. Control variables include: % female students, % students behind and ahead grade level, % immigrants, % fathers at least high school graduate, % unemployed mothers, % mother NILF, grade and year dummies, and dummies for missing values. All regressions include sampling strata controls (grade enrollment at institution, region dummies and their interactions). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: OLS and IV/2SLS estimates of the effect of class size on cheating

	OLS			IV/2SLS					
	Italy	North/Centre	South	Italy	North/Centre	South	Italy	North/Centre	South
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				A. Math					
Class size	-0.002***	-0.001***	-0.003***	-0.002***	-0.0004	-0.006***	-0.002***	-0.001	-0.005**
	(0.000)	(0.000)	(0.001)	(0.000)	(0.0003)	(0.001)	(0.001)	(0.000)	(0.002)
enrollment	✗	✗	✗	✗	✗	✗	✗	✗	✗
enrollment squared	✗	✗	✗	✗	✗	✗	✗	✗	✗
interactions							✗	✗	✗
N	139,996	87,491	52,505	139,996	87,491	52,505	139,996	87,491	52,505
				B. Language					
Class size	-0.002***	-0.001***	-0.002***	-0.002***	-0.001***	-0.004***	-0.002**	-0.001	-0.004**
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.002)
enrollment	✗	✗	✗	✗	✗	✗	✗	✗	✗
enrollment squared	✗	✗	✗	✗	✗	✗	✗	✗	✗
interactions							✗	✗	✗
N	140,003	87,493	52,510	140,003	87,493	52,510	140,003	87,493	52,510

Notes: The left panel shows OLS estimates of the effect of class size on cheating. The right panel shows 2SLS estimates using Maimonides' rule as instrument. The unit of observation is the class. Robust standard errors, clustered on school and grade, are shown in parentheses. Control variables include: % female students, % students behind and ahead grade level, % immigrants, % fathers at least high school graduate, % unemployed mothers, % mother NILF, grade and year dummies, and dummies for missing values. All regressions include sampling strata controls (grade enrollment at institution, region dummies and their interactions). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Covariate balance in the cheating experiment

	Italy		North/Centre		South	
	Control Mean	Treatment Difference	Control Mean	Treatment Difference	Control Mean	Treatment Difference
	(1)	(2)	(3)	(4)	(5)	(6)
			A. Administrative Data on Schools			
Class size	19.812	0.035	20.031	0.018	19.456	0.062
	[3.574]	(0.030)	[3.511]	(0.037)	[3.646]	(0.051)
Grade enrollment at school	53.119	-0.401	49.804	-0.548	58.483	-0.141
	[30.663]	(0.329)	[27.562]	(0.391)	[34.437]	(0.591)
% in class sitting the test	0.939	0.0001	0.934	0.001	0.947	-0.001
	[0.065]	(0.001)	[0.066]	(0.001)	[0.062]	(0.001)
% in school sitting the test	0.938	-0.0001	0.933	0.001	0.946	-0.001
	[0.054]	(0.001)	[0.055]	(0.001)	[0.051]	(0.001)
% in institution sitting the test	0.937	-0.0001	0.932	0.001	0.945	-0.001
	[0.045]	(0.0004)	[0.043]	(0.001)	[0.045]	(0.001)
% female students	0.482	0.001	0.483	0.0004	0.479	0.003*
	[0.121]	(0.001)	[0.1179]	(0.0011)	[0.126]	(0.002)
			B. Data Provided by School Staff			
instruction time: up to 30 hours per week	0.622	0.010**	0.491	0.008	0.833	0.0140**
	[0.485]	(0.004)	[0.500]	(0.005)	[0.373]	(0.006)
instruction time: more than 30 hours per week	0.339	0.003	0.48	0.002	0.111	0.005
	[0.473]	(0.004)	[0.500]	(0.005)	[0.314]	(0.005)
behind grade level	0.026	0.001**	0.03	0.001*	0.019	0.001
	[0.045]	(0.000)	[0.047]	(0.001)	[0.039]	(0.001)
ahead grade level	0.012	0.0001	0.005	0.0001	0.025	-0.0000
	[0.035]	(0.0002)	[0.017]	(0.0002)	[0.049]	(0.0006)
immigrant 1st generation	0.041	-0.0004	0.057	-0.001	0.015	0.001
	[0.065]	(0.0005)	[0.072]	(0.001)	[0.038]	(0.001)

Table 4: Covariate balance in the cheating experiment (cont.)

	Italy		North/Centre		South	
	Control Mean	Treatment Difference	Control Mean	Treatment Difference	Control Mean	Treatment Difference
	(1)	(2)	(3)	(4)	(5)	(6)
immigrant 2nd generation	0.097	0.001	0.137	0.0004	0.031	0.002***
	[0.120]	(0.001)	[0.13]	(0.0014)	[0.056]	(0.001)
father HS	0.25	0.006***	0.258	0.006***	0.238	0.006**
	[0.168]	(0.002)	[0.163]	(0.002)	[0.176]	(0.003)
mother employed	0.441	0.009***	0.532	0.007**	0.295	0.012***
	[0.267]	(0.002)	[0.258]	(0.003)	[0.210]	(0.004)
			Non-Response Indicators			
Missing data on instruction time	0.039	-0.013***	0.029	-0.010***	0.056	-0.019***
	[0.194]	(0.002)	[0.167]	(0.002)	[0.231]	(0.003)
Missing data on students behind/ahead grade level	0.015	-0.003***	0.012	-0.001	0.021	-0.005***
	[0.109]	(0.001)	[0.096]	(0.001)	[0.127]	(0.002)
Missing data on father's education	0.223	-0.022***	0.225	-0.019***	0.221	-0.027***
	[0.341]	(0.003)	[0.340]	(0.004)	[0.343]	(0.006)
Missing data on mother's occupation	0.195	-0.017***	0.196	-0.008**	0.194	-0.032***
	[0.328]	(0.003)	[0.325]	(0.004)	[0.333]	(0.005)
Missing data on country of origin	0.033	-0.012***	0.025	-0.008***	0.045	-0.018***
	[0.163]	(0.001)	[0.143]	(0.001)	[0.192]	(0.003)
N	140,010		87,498		52,512	

Notes: Columns 1,3 amd 5 show means and standard deviations for variables in the left-most column. Other columns report coefficients from regressions of each variable these variables on a treatment dummy (indicating classroom monitoring), grade and year dummies, and sampling strata controls (grade enrollment at institution, region dummies and their interactions). Standard deviations for the control group are in square brackets, robust standard errors are in parentheses, and p values for F tests are in curly braces. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: First stage and reduced form monitor effects

	First stage			Reduced form		
	Italy	North/Centre	South	Italy	North/Centre	South
	(1)	(2)	(3)	(4)	(5)	(6)
			A. Math			
Observer at institution	-0.029***	-0.010***	-0.062***	-0.112***	-0.075***	-0.180***
	(0.002)	(0.001)	(0.004)	(0.006)	(0.005)	(0.012)
Means	0.064	0.020	0.139	0.007	-0.074	0.141
(sd)	(0.246)	(0.139)	(0.346)	(0.637)	(0.502)	(0.796)
N	139,996	87,491	52,505	140,010	87,498	52,512
			B. Language			
Observer at institution	-0.025***	-0.012***	-0.047***	-0.080***	-0.054***	-0.130***
	(0.002)	(0.001)	(0.004)	(0.004)	(0.004)	(0.009)
Means	0.055	0.023	0.110	0.01	-0.005	0.035
(sd)	(0.229)	(0.149)	(0.313)	(0.523)	(0.428)	(0.649)
N	140,003	87,493	52,510	140,010	87,498	52,512

Notes: Columns 1-3 report first stage estimates of the effect of a classroom observer on cheating. Columns 4-6 show the reduced form effect of an observer on test scores. All models control for a quadratic polynomial in grade enrollment, segment dummies and their interactions. The unit of observation is the class. Robust standard errors, clustered on school and grade, are shown in parentheses. Control variables include: % female students, % students behind and ahead grade level, % immigrants, % fathers at least high school graduate, % unemployed mothers, % mother NILF grade and year dummies, and dummies for missing values in these variables. All regressions include sampling strata controls (grade enrollment at institution, region dummies and their interactions). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Twin first stages

	A. Cheating					
	Math			Language		
	Italy	North/Centre	South	Italy	North/Centre	South
	(1)	(2)	(3)	(4)	(5)	(6)
Maimonides' rule	-0.001**	-0.0003	-0.002**	-0.001**	-0.0003	-0.002**
	(0.000)	(0.0002)	(0.001)	(0.000)	(0.0003)	(0.001)
Observer at institution	-0.029***	-0.010***	-0.062***	-0.025***	-0.012***	-0.047***
	(0.002)	(0.001)	(0.004)	(0.002)	(0.001)	(0.004)
N	139,996	87,491	52,505	140,003	87,493	52,510
	B. Class size					
	Italy	North/Centre	South			
	(1)	(2)	(3)			
Maimonides' rule	0.511***	0.552***	0.432***			
	(0.006)	(0.008)	(0.011)			
Observer at institution	0.014	0.032	-0.007			
	(0.024)	(0.027)	(0.045)			
N	140,010	87,498	52,512			

Notes: Panel A report first stage estimates of the effect of the Maimonides' rule and a classroom observer on cheating. Panel B report first stage estimates of the effect of the Maimonides' rule and a classroom observer on class size. All models control for a quadratic polynomial in grade enrollment, segment dummies and their interactions. The unit of observation is the class. Robust standard errors, clustered on school and grade, are shown in parentheses. Control variables include: % female students, % students behind and ahead grade level, % immigrants, % fathers at least high school graduate, % unemployed mothers, % mother NILF grade and year dummies, and dummies for missing values in these variables. All regressions include sampling strata controls (grade enrollment at institution, region dummies and their interactions). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: OLS and IV/ 2SLS estimates of the effect of class size and cheating on test scores

	OLS			IV/2SLS			IV/2SLS (overidentified)		
	Italy	North/Centre	South	Italy	North/Centre	South	Italy	North/Centre	South
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				A. Math					
Class size	0.002***	-0.0004	0.005***	0.001	-0.0004	0.0006	0.0000	-0.0004	0.0004
	(0.001)	(0.0007)	(0.001)	(0.002)	(0.0030)	(0.0044)	(0.0020)	(0.0027)	(0.0038)
Cheating	1.413***	1.403***	1.413***	3.824***	7.315***	2.884***	3.812***	7.265***	2.886***
	(0.006)	(0.009)	(0.007)	(0.189)	(0.786)	(0.159)	(0.189)	(0.780)	(0.158)
Overid test [P-value]							[0.640]	[0.664]	[0.847]
N	139,996	87,491	52,505	139,996	87,491	52,505	139,996	87,491	52,505
				B. Language					
Class size	0.003***	0.0002	0.006***	0.001	0.001	0.001	0.002	0.001	0.004
	(0.001)	(0.0005)	(0.001)	(0.002)	(0.002)	(0.004)	(0.002)	(0.002)	(0.003)
Cheating	1.179***	1.084***	1.213***	3.279***	4.469***	2.786***	3.239***	4.229***	2.786***
	(0.005)	(0.007)	(0.006)	(0.182)	(0.448)	(0.179)	(0.179)	(0.414)	(0.179)
Overid test (P-value)							[0.167]	[0.166]	[0.328]
N	140,003	87,493	52,510	140,003	87,493	52,510	140,003	87,493	52,510

Notes: The left panel shows OLS estimates of the effect of class size and cheating on scores. The center panel shows 2SLS estimates using Maimonides' rule and classroom observer as instruments.The right panel shows overidentified 2SLS estimates which also use a dummy for grade enrollment being in a 10 percent window below and above each cutoff as instrument. All models control for a quadratic polynomial in grade enrollment, segment dummies and their interactions. The unit of observation is the class. Robust standard errors, clustered on school and grade, are shown in parentheses. Control variables include: % female students, % students behind and ahead grade level, % immigrants, % fathers at least high school graduate, % unemployed mothers, % mother NILF grade and year dummies, and dummies for missing values in these variables. All regressions include sampling strata controls (grade enrollment at institution, region dummies and their interactions). * significant at 10%; ** significant at 5%; *** significant at 1%.

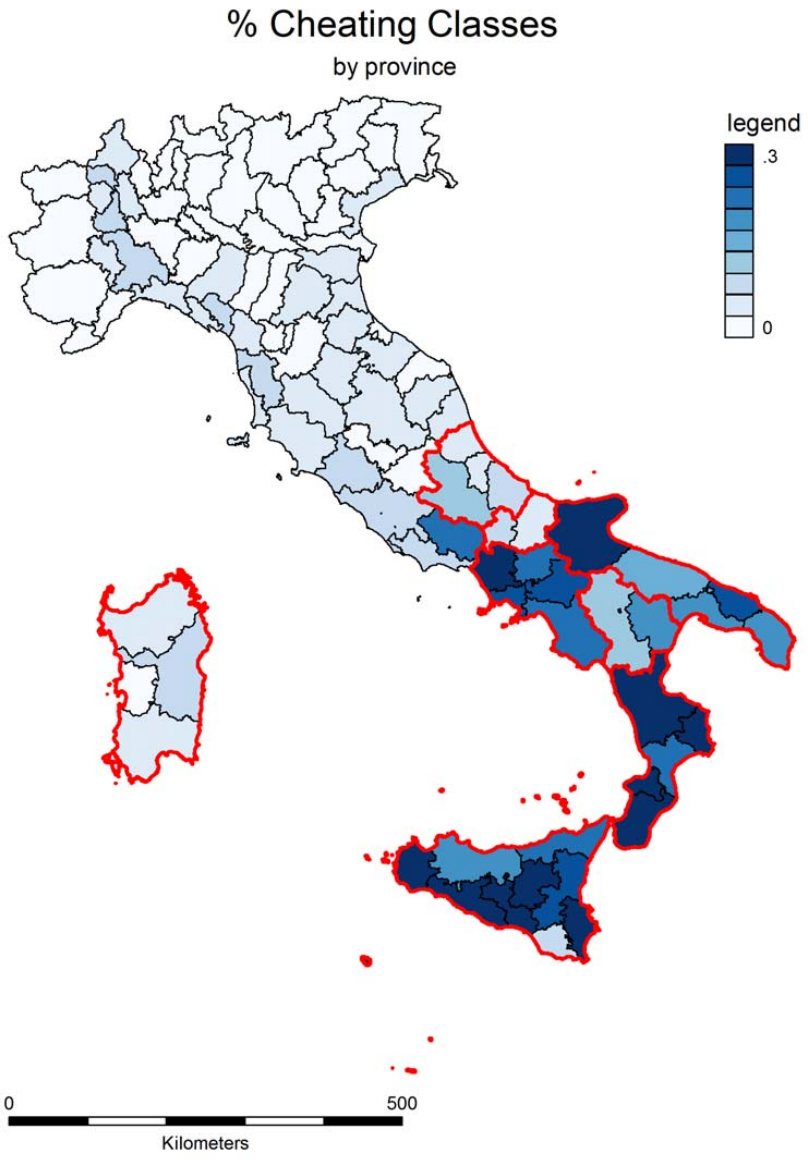
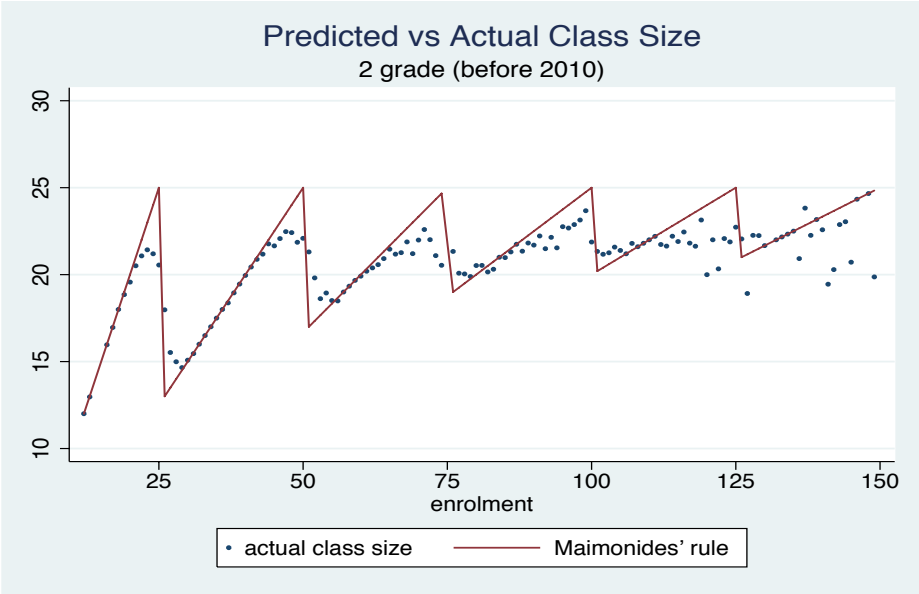


Figure 1: Cheating map



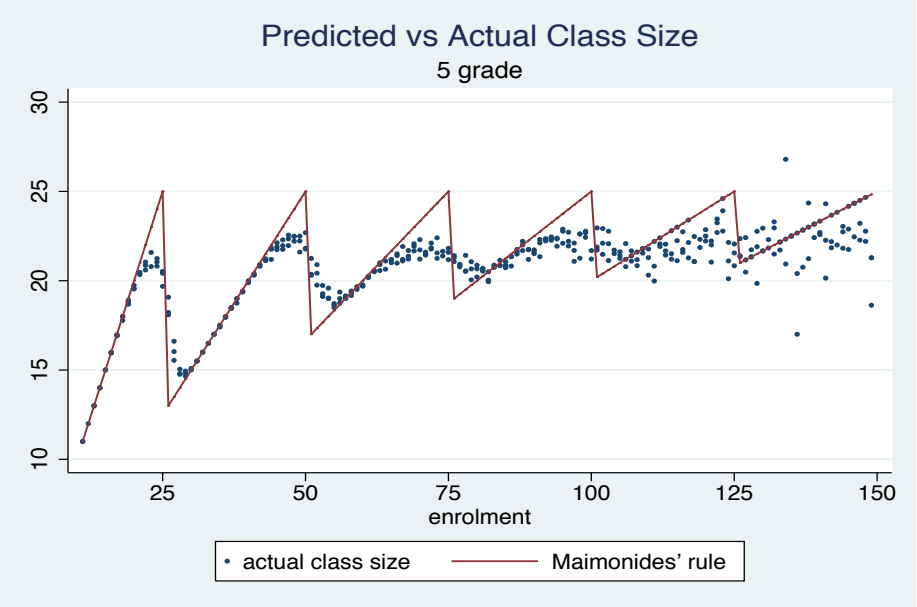


Figure 2: Class size by enrollment. The figure shows actual class size and as predicted by Maimonides' rule in pre-reform years

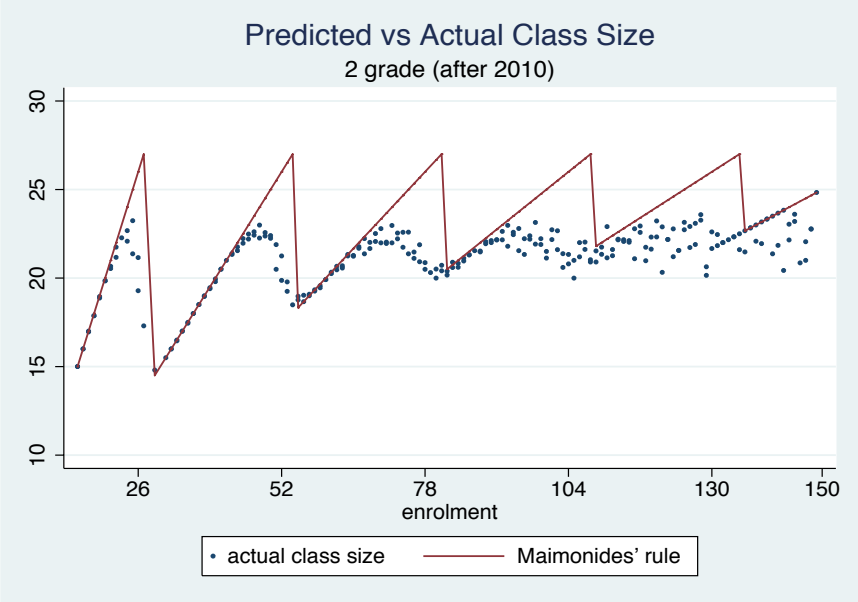
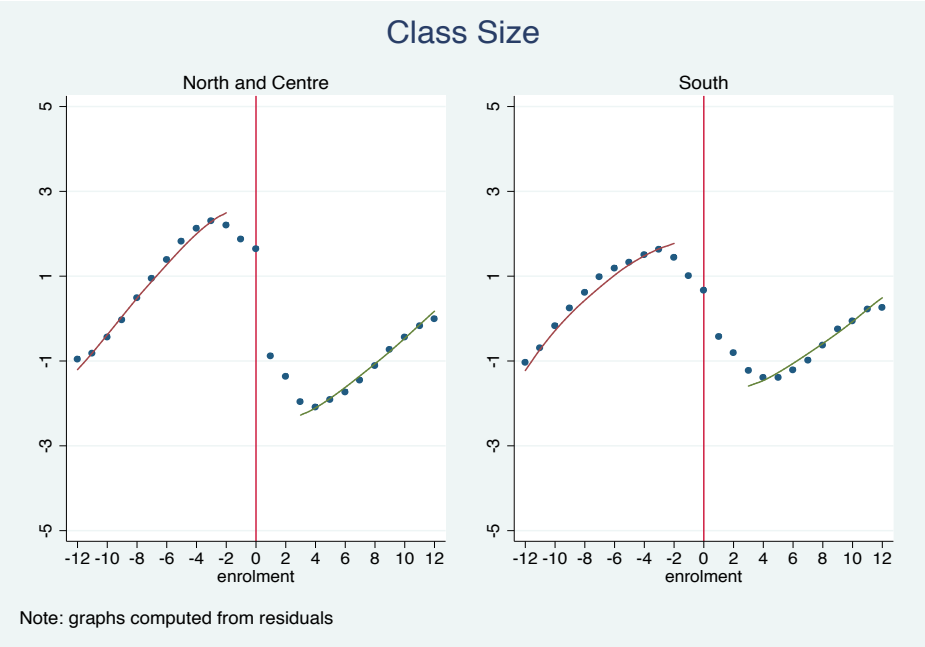
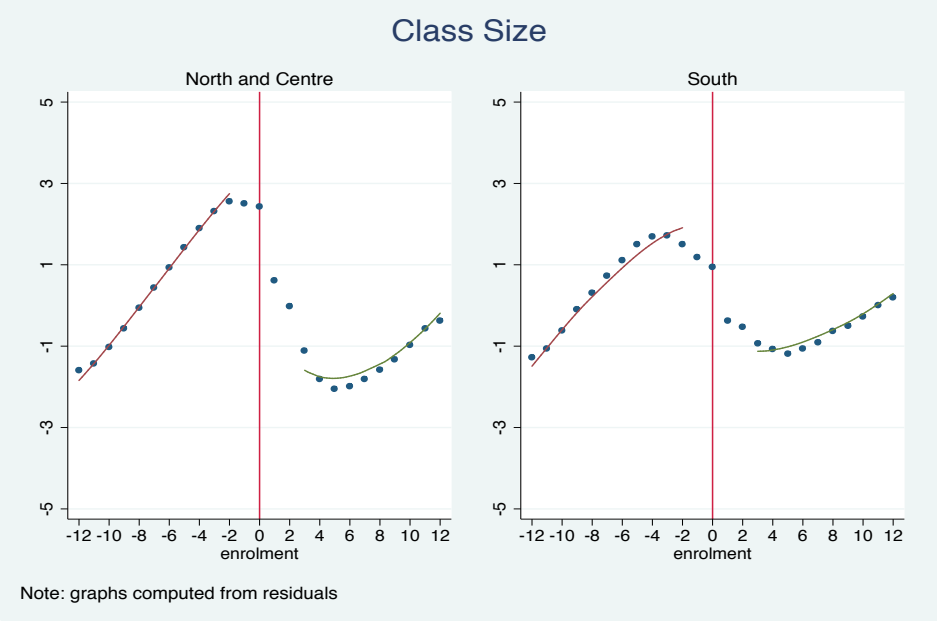


Figure 3: Class size by enrollment. The figure shows actual class size and as predicted by Maimonides' rule in reform years

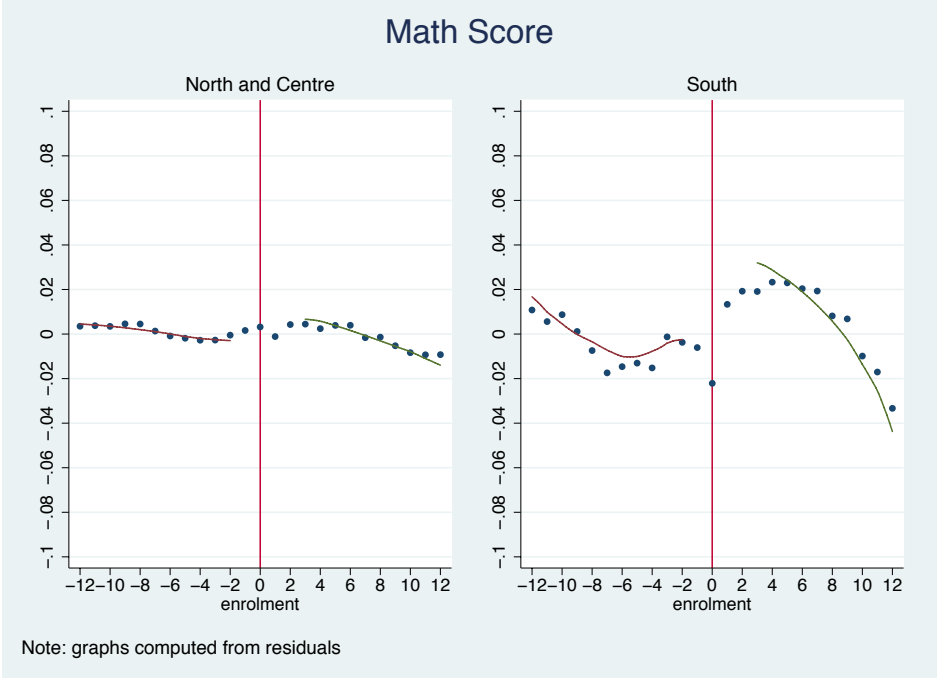


Panel A: Class Size Around Cutoffs for Grade 2



Panel B: Class Size Around Cutoffs for Grade 5

Figure 4: Class size and enrollment, centered at Maimonides cutoffs. The solid line shows a one-sided LLR fit.



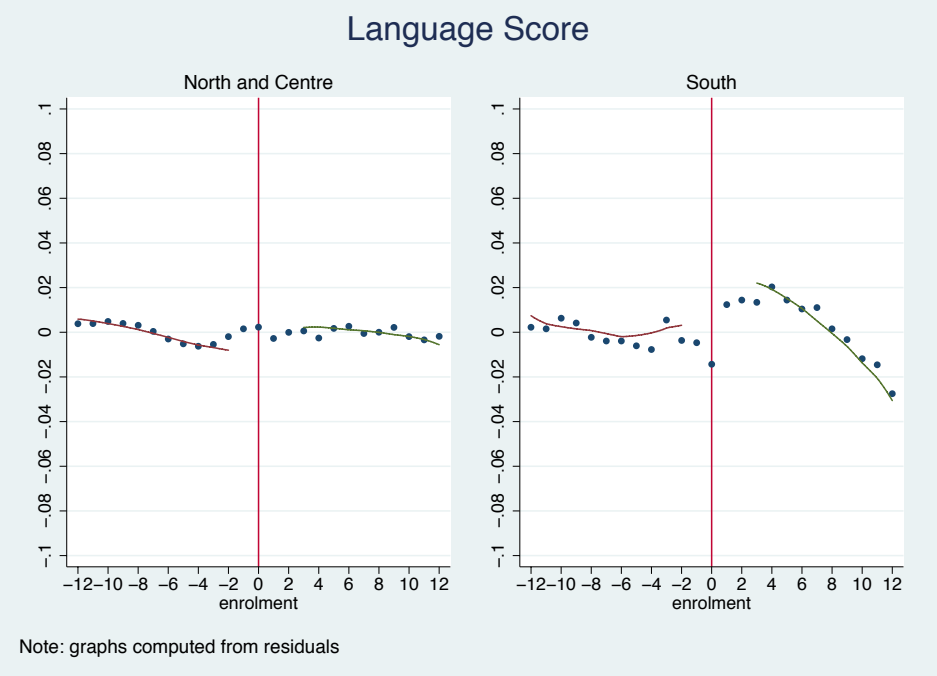
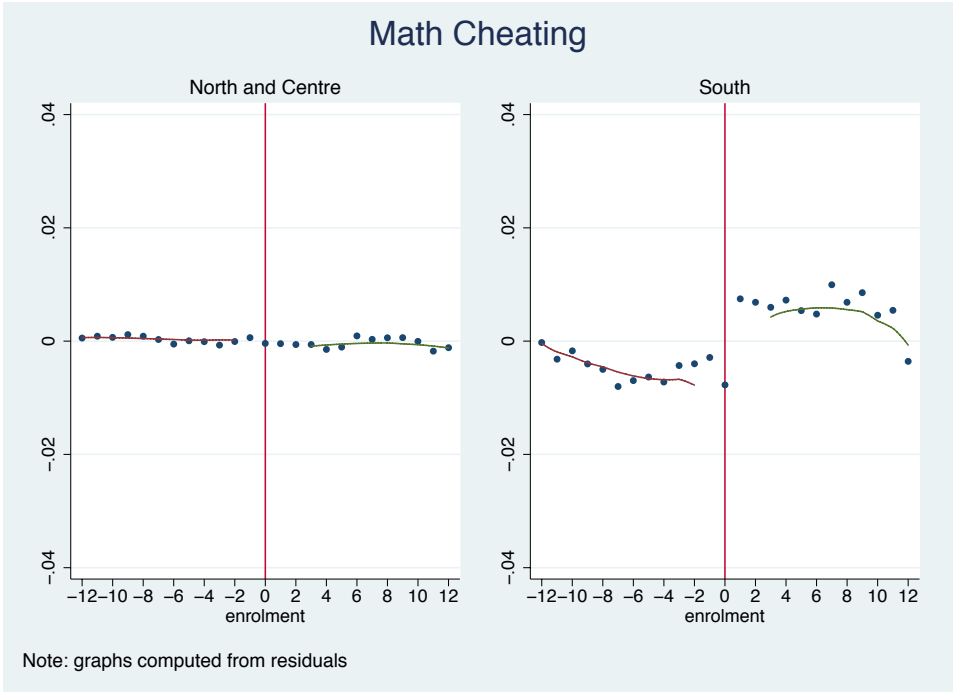


Figure 5: Test scores and enrollment, centered at Maimonides cutoffs. The solid line shows a one-sided LLR fit.



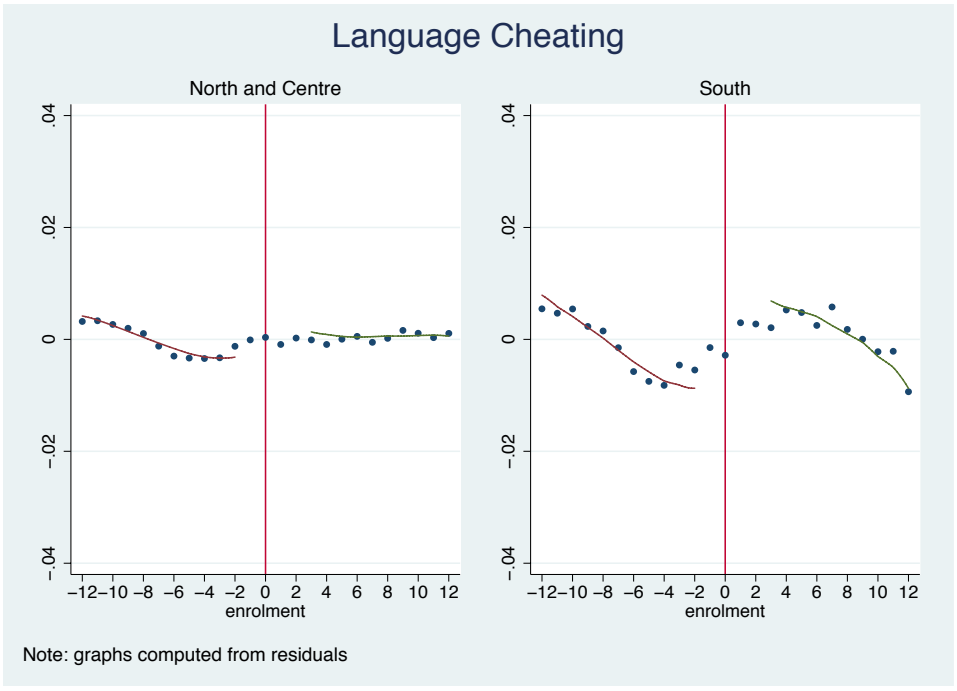


Figure 6: Cheating and enrollment, centered at Maimonides cutoffs. The solid line shows a one-sided LLR fit.

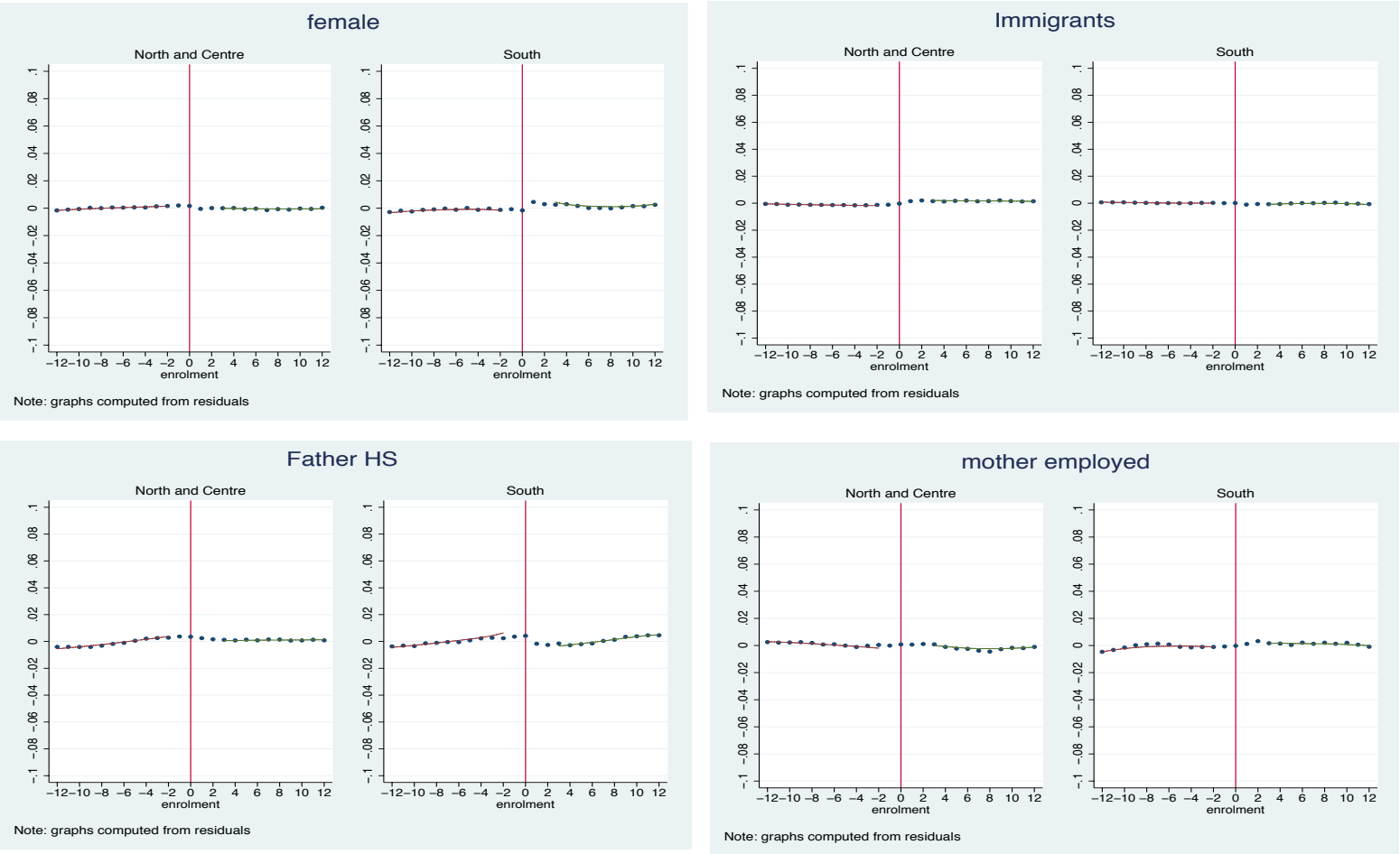


Figure 7: Covariates and enrollment, centered at Maimonides cutoffs. The solid line shows a one-sided LLR fit.

References

Angrist, J. D., and V. Lavy (1999): "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics*, 114(2), 533–575.

Baker, O., and D. Paserman (2013): "Grade Enrollment Sorting under an Incentives-Based Class Size Reduction Program," Unpublished mimeo.

Ballatore, R., M. Fort, and A. Ichino (2013): "The Tower of Babel in the classroom: immigrants and natives in Italian schools," Unpublished mimeo.

Bertoni, M., G. Brunello, and 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*, forthcoming.

Bezdek, J. (1981): *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York.

Bonesronning, H. (2003): "Class size effects on student achievement in Norway: Patterns and explanations," *Southern Economic Journal*.

Bratti, M., D. Checchi, and A. Filippin (2007): "Territorial differences in Italian students' mathematical competences: Evidence from PISA," *Giornale degli Economisti e Annali di Economia*, 66(3), 299–335.

Brunello, G., and D. Checchi (2005): "School quality and family background in Italy," *Economics of Education Review*, 24, 563–577.

Chetty, R., J. Friedman, N. Hilger, E. Saez, D. Schanzenbach, and D. Yagan (2011): "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR," *Quarterly Journal of Economics*, 126(4), 1593–1660.

Dobbelsteen, S., J. Levin, and H. Oosterbeek (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(1), 17–38.

Gary-Bobo, R. J., and M.-B. Mahjoub (2006): "Estimation of class-size effects, using Maimonides' rule: the case of French junior high schools," CEPR Discussion Papers 5754.

Guiso, L., P. Sapienza, and L. Zingales (2004): "The Role of Social Capital in Financial Development," *American Economic Review*, 94(3), 526–556.

Hanushek, E. A. (1995): "Interpreting recent research on schooling in developing countries," *The World Bank Research Observer*, X, 227–246.

Hoxby, C. (2000): "Peer Effects in the Classroom: Learning from Gender and Race Variation," NBER Working paper 7867.

Ichino, A., and G. Maggi (2000): "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm," *Quarterly Journal of Economics,*, 115(3), 933–959.

Imbens, G., and K. Kalyanaraman (2012): "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *Review of Economic Studies*, 79(3), 933–959.

INVALSI (2010): "Sistema Nazionale di Valutazione - A.S. 2009/2010, Rilevazione degli apprendimenti," *Technical Report*.

Jacob, B., and S. Levitt (2003): "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *Quarterly Journal of Economics*, 118(3), 843–77.

Krueger, A. (1999): "Experimental estimates of education production functions," *Quarterly Journal of Economics*, 114, 497–532.

Leuven, E., H. Oosterbeek, and M. Ronning (2008): "Quasi-experimental estimates of the effect of class size achievement in Norway," *The Scandinavian Journal of Economics*, 110(4), 663–693.

Nannicini, T., A. Stella, G. Tabellini, and U. Troiano (2013): "Social Capital and Political Accountability," *American Economic Journal: Economic Policy*, 5, 1957–1969.

Piketty, T. (2004): "Should we reduce class size or school segregation? Theory and evidence from France," presentation at the Roy Seminars, Association pour le dévelopement de la recherche en économie et en statistique (ADRES), 22 November, available at: http://www.adres.polytechnique.fr/SEMINAIRE/221104b.pdf.

Putnam, R., R. Leonardi, and R. Nanetti (1993): *Making Democracy Work*. Princeton University Press, Princeton.

Quintano, C., R. Castellano, and S. Longobardi (2009): "A Fuzzy Clustering Approach to Improve the Accuracy of Italian Student Data. An Experimental Procedure to Correct the Impact of the Outliers on Assessment Test Scores," *Statistica & Applicazioni*, Vol.VII(2), 149–171.

Urquiola, M., and E. Verhoogen (2009): "Class size caps, sorting, and the regression discontinuity design," *American Economic Review*, 99(1), 179–215.

Woessmann, L. (2005): "Educational production in Europe," *Economic Policy*, 43, 445–493.

Appendix

Cheating Imputation

Our imputation is closely related to that used by INVALSI and described in Quintano et al. (2009). INVALSI assigns a cheating probability to each class in three steps.

The first step computes the following four summary statistics.

(1) Within-class average score:

$$\bar{p}_i = \frac{\sum_{j=1}^{N_i} p_{ji}}{N_i}, \tag{10}$$

where p_{ji} denotes the score of student j in class i; N_i denotes the number of test-takers in class i.

(2) Within-class standard deviation of scores:

$$\sigma_i = \sqrt{\frac{\sum_{j=1}^{N_i} (p_{ji} - \bar{p}_i)^2}{N_i}}. \tag{11}$$

(3) Within-class average percent missing

$$MC_i = \frac{\sum_{j=1}^{N_i} M_{ji}}{N_i}, \tag{12}$$

where M_{ji} is the fraction of test items skipped by student j in class i.

(4) Within-class index of answer homogeneity:

$$\bar{E}_i = \frac{\sum_{q=1}^{Q} E_{qi}}{Q}, \tag{13}$$

where $q = 1, .., Q$ indexes test items and E_{qi} is a Gini measure of homogeneity that equals value zero if all students in class i provide the same answer to item q. This can be interpreted as the Herfindahl index of the share of students with similar response patterns in the class.

In the second step, the first two principal components are extracted from the 4×4 correlation matrix determined by these indicators, yielding a percentage of explained variance which is - across years, subjects and grades - well above 90%. Denote these principal components by ψ_{1i} and ψ_{2i}. The third step consists of a cluster analysis that creates G groups from the distribution of (ψ_{1i}, ψ_{2i}). INVALSI sets $G = 8$, yielding a matrix whose elements are, for each class, eight group membership probabilities. This procedure is known as "fuzzy clustering" (see Bezdek, 1981), since data elements

(classes, in our setting) can be assigned to one or more groups. With "hard clustering", data elements belong to exactly one cluster.

INVALSI identifies likely cheaters as those in the group with values of (ψ_{1i}, ψ_{2i}) that are most extreme (see Figure 8 in Quintano et al. 2009). In practice, the suspicious group is characterized by (i) abnormally large values of $\bar{p}_i$, and (ii) small values of σ_i, MC_i and $\bar{E}_i$, relative to the population average of these indicators. This group is flagged as the "outlier" or cheating cluster. The INVALSI cheating indicator gives, for each class, the membership probability for this cluster. Our hard clustering computations codes a dummy for for cheating classes. This dummy indicates classes whose values of (ψ_{1i}, ψ_{2i}) belong to the cheating cluster identified by INVALSI.

Table A.1: Reduced form estimates of the effect of Maimonides' rule on test scores, cheating and class size

	Math			Language		
	Italy	North/Centre	South	Italy	North/Centre	South
	(1)	(2)	(3)	(4)	(5)	(6)
			A. Class size			
Maimonides rule	0.511***	0.552***	0.432***			
	(0.006)	(0.008)	(0.011)			
Means	19.88	20.07	19.58			
(sd)	(3.58)	(3.52)	(3.64)			
N	140,010	87,498	52,512			
			B. Scores			
Maimonides' rule	-0.003***	-0.002**	-0.006***	-0.002**	-0.001	-0.004**
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Means	0.007	-0.074	0.141	0.01	-0.005	0.035
(sd)	(0.637)	(0.502)	(0.796)	(0.523)	(0.428)	(0.649)
N	140,010	87,498	52,512	140,010	87,498	52,512
			C. Cheating			
Maimonides' rule	-0.001***	-0.0003	-0.002**	-0.001**	-0.0003	-0.002**
	(0.000)	(0.0002)	(0.001)	(0.000)	(0.0003)	(0.001)
Means	0.065	0.02	0.139	0.055	0.023	0.110
(sd)	(0.246)	(0.139)	(0.346)	(0.229)	(0.149)	(0.313)
N	139,996	87,491	52,505	140,003	87,493	52,510
			D. Cheating (INVALSI)			
Maimonides' rule	-0.001**	-0.000*	-0.001*	-0.001***	-0.0003*	-0.001**
	(0.000)	(0.000)	(0.001)	(0.000)	(0.0002)	(0.001)
Means	0.066	0.025	0.133	0.057	0.028	0.106
(sd)	(0.191)	(0.099)	(0.272)	(0.175)	(0.104)	(0.244)
N	140,003	87,494	52,509	140,004	87,493	52,511

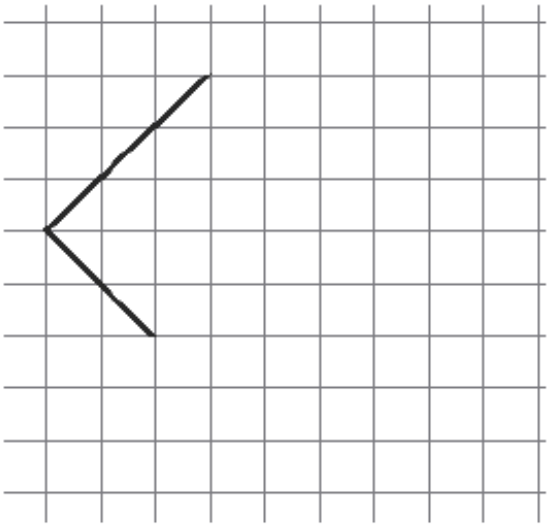
Notes: This table shows the reduced form effect of the Maimonides' rule on test scores (Panel A), cheating (Panel B), cheating (INVALSI) (Panel C) and class size (Panel D). All models control for a quadratic polynomial in grade enrollment, segment dummies and their interactions. The unit of observation is the class. Robust standard errors, clustered on school and grade, are shown in parentheses. Control variables include: % female students, % students behind and ahead grade level, % immigrants, % fathers at least high school graduate, % unemployed mothers, % mother NILF grade and year dummies, and dummies for missing values in these variables. All regressions include sampling strata controls (grade enrollment at institution, region dummies and their interactions). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.2: First stage regressions for overidentified IV/2SLS

	Class size			Cheating math			Cheating language		
VARIABLES	Italy	North/Centre	South	Italy	North/Centre	South	Italy	North/Centre	South
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Maimonides	0.735***	0.776***	0.658***	-0.001**	-0.001	-0.002	-0.001***	-0.001**	-0.002**
	(0.006)	(0.007)	(0.012)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
observer	0.014	0.030	-0.003	-0.029***	-0.010***	-0.062***	-0.025***	-0.012***	-0.047***
	(0.023)	(0.026)	(0.044)	(0.002)	(0.001)	(0.004)	(0.002)	(0.001)	(0.004)
10% above cutoff	3.026***	3.098***	2.957***	-0.002	-0.004	-0.001	-0.009*	-0.007*	-0.012
	(0.075)	(0.094)	(0.121)	(0.006)	(0.004)	(0.013)	(0.005)	(0.004)	(0.011)
within 10% from cutoff	-1.575***	-1.356***	-1.945***	-0.0001	0.001	-0.002	0.007**	0.007***	0.007
	(0.054)	(0.065)	(0.094)	(0.0035)	(0.002)	(0.008)	(0.003)	(0.002)	(0.008)
N	140,010	87,498	52,512	139,996	87,491	52,505	140,003	87,493	52,510

Notes: Columns 1-3 report first stage estimates of the effect of the Maimonides' rule, a classroom observer and a dummy for grade enrollment being in a 10 percent window below and above each cutoff on class size. Columns 4-9 show first stage estimates of the effect of the Maimonides' rule, a classroom observer and a dummy for grade enrollment being in a 10 percent window below and above each cutoff on cheating. All models control for a quadratic polynomial in grade enrollment, segment dummies and their interactions. The unit of observation is the class. Robust standard errors, clustered on school and grade, are shown in parentheses. Control variables include: % female students, % students behind and ahead grade level, % immigrants, % fathers at least high school graduate, % unemployed mothers, % mother NILF grade and year dummies, and dummies for missing values in these variables. All regressions include sampling strata controls (grade enrollment at institution, region dummies and their interactions). * significant at 10%; ** significant at 5%; *** significant at 1%.

D23. Osserva la seguente figura.



a. Completa la figura in modo da ottenere un quadrato.

b. Spiega come hai fatto per disegnare il quadrato.

.....

.....

.....

Figure 8: Example of open question in math test - V grade 2010/11

C3. **Leggi questa frase: "L'uccello notturno emise un suono talmente acuto che fece molta paura agli abitanti del bosco".**
Indica nella tabella quali parole sono nomi e quali non lo sono. Metti una crocetta per ogni riga della tabella.

		È un nome	Non è un nome
a.	L'	☐	☐
b.	uccello	☐	☐
c.	notturno	☐	☐
d.	emise	☐	☐
e.	un	☐	☐
f.	suono	☐	☐
g.	talmente	☐	☐
h.	acuto	☐	☐
i.	che	☐	☐
l.	fece	☐	☐
m.	molta	☐	☐
n.	paura	☐	☐
o.	agli	☐	☐
p.	abitanti	☐	☐
q.	del	☐	☐
r.	bosco	☐	☐

Figure 9: Example of open question in language test - V grade 2010/11

Servizio Nazionale di Valutazione a.s. 2010/11

CLASSE:

Scheda Risposte Studente n°

Risultati delle prove

Codice istituto: Codice Scuola:

Codice plesso: Livello:

Codice Classe: **NON CAMPIONE**

Codice studente: Numero progressivo studente:

PROVA ITALIANO (1)

A1	☐A	☐B	☐C	☐D	☐NV	**C1_a1**	☐0	☐1	☐NV		
A2	☐A	☐B	☐C	☐D	☐NV	**C1_a2**	☐0	☐1	☐NV		
A3	☐A	☐B	☐C	☐D	☐NV	**C1_b1**	☐0	☐1	☐NV		
A4	☐A	☐B	☐C	☐D	☐NV	**C1_b2**	☐0	☐1	☐NV		
A5	☐A	☐B	☐C	☐D	☐NV	**C1_b3**	☐0	☐1	☐NV		
A6	☐A	☐B	☐C	☐D	☐NV	**C2**	☐A	☐B	☐C	☐D	☐NV
A7	☐A	☐B	☐C	☐D	☐NV	**C3_a**	☐Nome		☐Non_Nome		☐NV
A8	☐A	☐B	☐C	☐D	☐NV	**C3_b**	☐Nome		☐Non_Nome		☐NV
A9	☐A	☐B	☐C	☐D	☐NV	**C3_c**	☐Nome		☐Non_Nome		☐NV
A10	☐A	☐B	☐C	☐D	☐NV	**C3_d**	☐Nome		☐Non_Nome		☐NV
A11	☐A	☐B	☐C	☐D	☐NV	**C3_e**	☐Nome		☐Non_Nome		☐NV
A12	☐A	☐B	☐C	☐D	☐NV	**C3_f**	☐Nome		☐Non_Nome		☐NV
A13	☐A	☐B	☐C	☐D	☐NV	**C3_g**	☐Nome		☐Non_Nome		☐NV
A14	☐A	☐B	☐C	☐D	☐NV	**C3_h**	☐Nome		☐Non_Nome		☐NV
A15	☐A	☐B	☐C	☐D	☐NV	**C3_i**	☐Nome		☐Non_Nome		☐NV
A16	☐A	☐B	☐C	☐D	☐NV	**C3_l**	☐Nome		☐Non_Nome		☐NV
A17	☐A	☐B	☐C	☐D	☐NV	**C3_m**	☐Nome		☐Non_Nome		☐NV
B1	☐A	☐B	☐C	☐D	☐NV	**C3_n**	☐Nome		☐Non_Nome		☐NV
B2	☐A	☐B	☐C	☐D	☐NV	**C3_o**	☐Nome		☐Non_Nome		☐NV
B3	☐A	☐B	☐C	☐D	☐NV	**C3_p**	☐Nome		☐Non_Nome		☐NV
B4	☐A	☐B	☐C	☐D	☐NV	**C3_q**	☐Nome		☐Non_Nome		☐NV
B5	☐A	☐B	☐C	☐D	☐NV	**C3_r**	☐Nome		☐Non_Nome		☐NV
B6	☐A	☐B	☐C	☐D	☐NV	**C4**	☐0	☐1	☐NV		
B7	☐A	☐B	☐C	☐D	☐NV	**C5**	☐0	☐1	☐NV		
B8	☐A	☐B	☐C	☐D	☐NV	**C6**	☐A	☐B	☐C	☐D	☐NV
B9	☐A	☐B	☐C	☐D	☐NV	**C7**	☐0	☐1	☐NV		
B10	☐A	☐B	☐C	☐D	☐NV	**C8**	☐A	☐B	☐C	☐D	☐NV
B11	☐A	☐B	☐C	☐D	☐NV	**C9**	☐A	☐B	☐C	☐D	☐NV
B12	☐A	☐B	☐C	☐D	☐NV	**C10**	☐0	☐1	☐NV		
B13	☐A	☐B	☐C	☐D	☐NV						
B14	☐A	☐B	☐C	☐D	☐NV						
B15	☐A	☐B	☐C	☐D	☐NV						

PROVA MATEMATICA (1)

D1_a	☐V	☐F	☐NV		
D1_b	☐V	☐F	☐NV		
D1_c	☐V	☐F	☐NV		
D1_d	☐V	☐F	☐NV		
D2	☐A	☐B	☐C	☐D	☐NV
D3	☐0	☐1	☐NV		
D4_a	☐A	☐B	☐C	☐D	☐NV
D4_b	☐0	☐1	☐NV		
D5	☐A	☐B	☐C	☐D	☐NV
D6	☐A	☐B	☐C	☐D	☐NV
D7	☐A	☐B	☐C	☐D	☐NV
D8	☐A	☐B	☐C	☐D	☐NV
D9	☐0	☐1	☐NV		
D10	☐A	☐B	☐C	☐D	☐NV
D11	☐A	☐B	☐C	☐D	☐NV
D12	☐0	☐1	☐NV		
D13	☐A	☐B	☐C	☐D	☐NV
D14	☐A	☐B	☐C	☐D	☐NV
D15	☐A	☐B	☐C	☐D	☐NV
D16_a	☐0	☐1	☐NV		
D16_b	☐0	☐1	☐NV		
D17_a	☐V	☐F	☐NV		
D17_b	☐V	☐F	☐NV		
D17_c	☐V	☐F	☐NV		
D17_d	☐V	☐F	☐NV		
D18	☐A	☐B	☐C	☐D	☐NV
D19	☐0	☐1	☐NV		
D20	☐A	☐B	☐C	☐D	☐NV
D21_a	☐0	☐1	☐NV		
D21_b	☐0	☐1	☐NV		
D22	☐A	☐B	☐C	☐D	☐NV
D23_a	☐0	☐1	☐NV		
D23_b	☐0	☐1	☐NV		
D24_a	☐0	☐1	☐NV		
D24_b	☐0	☐1	☐NV		
D24_c	☐0	☐1	☐NV		
D25	☐0	☐1	☐NV		
D26	☐A	☐B	☐C	☐D	☐NV
D27	☐A	☐B	☐C	☐D	☐NV
D28_a	☐km	☐m	☐cm	☐mm	☐NV
D28_b	☐km	☐m	☐cm	☐mm	☐NV
D28_c	☐km	☐m	☐cm	☐mm	☐NV
D29_a	☐V	☐F	☐NV		
D29_b	☐V	☐F	☐NV		
D29_c	☐V	☐F	☐NV		
D29_d	☐V	☐F	☐NV		
D30	☐A	☐B	☐C	☐D	☐NV

(1) Barrare NV per risposta non valida (2 risposte o risposta incomprensibile) e non barrare nulla in caso di risposta omessa

(**ATTENZIONE Non spillare, non modificare per nessun motivo i dati precompilati della scheda)**

Figure 10: Answer sheet for V grade in 2010/11