Online Appendix for:

In a Small Moment:

Class Size and Moral Hazard in the Italian Mezzogiorno

Joshua D. Angrist

MIT Department of Economics

Erich Battistin

School of Economics and Finance, Queen Mary, University of London

Daniela Vuri

Department of Economics and Finance, University of Rome Tor Vergata

December 2016

Score Manipulation Imputation

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

The first step computes the following four summary statistics.

(1) Within-class average score

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

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

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

(3) Within-class average percent missing

$$MC_i = \frac{\sum_{j=1}^{N_i} M_{ji}}{N_i}, \tag{3}$$

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

(4) Within-class index of answer homogeneity

$$\bar{E}_i = \frac{\sum_{q=1}^{Q} E_{qi}}{Q}, \tag{4}$$

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 manipulators 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 manipulating cluster. The INVALSI manipulation indicator gives, for each class, the membership probability for this cluster. Our hard clustering computations codes a dummy for manipulating classes. This dummy indicates classes whose values of (ψ_{1i}, ψ_{2i}) belong to the manipulating cluster identified by INVALSI.

Manipulation and Class Size

Class size is denoted by s and, in the absence of manipulation, the score on item j is $L_j \in [0, 1]$. Manipulated scores are equal to 1. The manipulated class average score on item j is therefore $y_j = (1 - L_j)p_j + L_j$, where $p_j = \frac{n_j}{s}$ is the fraction of score sheets manipulated for item j and n_j is number of score sheets manipulated for item j. The score gain from manipulation is $\tau_j p_j$, with $\tau_j = 1 - L_j \geq 0$. Large τ_j denotes difficult items, so the returns to manipulation vary with item difficulty. Discovery risk cumulates across items, $\gamma(s) \sum_h n_h$, with $\gamma(s)$ increasing in class size: $\gamma'(s) > 0$. Assuming teachers care about total scores and utility is zero when caught manipulating, the objective is

$$\max_{\mathbf{p_j}} \left(1 - \underbrace{\gamma(s) \sum_h n_h}_{disclosure\ risk}\right) \underbrace{U\left(\sum_h \tau_h p_h\right)}_{utility\ of\ score\ gain} - \underbrace{\beta \sum_h (s - n_h)}_{honest\ grading\ effort}.$$

When utility is linear: $U\left(\sum_h \tau_h p_h\right) = \alpha \sum_h \tau_h p_h$, the FOC for the optimal p_j can be written

$$\frac{\tau_j}{s} + \frac{\beta}{\alpha} - \gamma(s) \sum_h (\tau_j + \tau_h)\, p_h = 0.$$

Comparative statics

$$\left[-\frac{\tau_j}{s^2} - \gamma'(s) \sum_h (\tau_j + \tau_h) p_h \right] ds + \left[-2\tau_j \gamma(s) \right] dp_j = 0,$$

implies equation (3) in the text

$$\frac{dp_j}{ds} = -\frac{1 + \gamma'(s) s^2 \sum_h (1 + \frac{\tau_h}{\tau_j}) p_h}{2\gamma(s) s^2} < 0.$$

More generally, dividing by s and solving for the optimal p_j yields a FOC for item j that can be written:

$$-\gamma(s) + \tau_j \left(\frac{1}{s} - \gamma(s) \sum_h p_h \right) h \left(\sum_h \tau_h p_h \right) + \beta g \left(\sum_h \tau_h p_h \right) = 0,$$

where the equation is multiplied by $g(p) = \frac{1}{U(p)} > 0$ and $h(p) = \frac{U'(p)}{U(p)} > 0$. Using $\frac{g'(p)}{h(p)} = -g(p)$, we obtain

$$\frac{dp_j}{ds} = -\frac{\frac{\gamma'(s)}{\tau_j h\left(\sum_h \tau_h p_h\right)} + \left(\frac{1}{s^2} + \gamma'(s) \sum_h p_h \right)}{\gamma(s) - \tau_j \left(\frac{1}{s} - \gamma(s) \sum_h p_h \right) \frac{h'\left(\sum_h \tau_h p_h\right)}{h\left(\sum_h \tau_h p_h\right)} + \beta g \left(\sum_h \tau_h p_h \right)}.$$

This is negative if the denominator is positive, that is if

$$\frac{h' \left(\sum_h \tau_h p_h \right)}{h \left(\sum_h \tau_h p_h \right)} \leq \frac{\gamma\left(s\right) + \beta g \left(\sum_h \tau_h p_h \right)}{\tau_j \left(\frac{1}{s} - \gamma(s) \sum_h p_h \right)},$$

a sufficient condition for which is $h'(p) < 0$, satisfied by commonly used log-linear preferences.[1] We then have

$$\frac{h' \left(\sum_h \tau_h p_h \right)}{h \left(\sum_h \tau_h p_h \right)} = -\frac{1}{\sum_h \tau_h p_h} < 0,$$

so that $\frac{dp_j}{ds}$ is negative. This implies a negative score gradient in class size

$$\frac{dy_j}{ds} = \frac{d\left[(1 - L_j) p_j + L_j \right]}{ds} = \tau_j \frac{dp_j}{ds}.$$

[1]Note that the quantity at the denominator of the right hand side of the expression above is positive because disclosure risk is presumed to be positive.

Nonlinear disutility of effort

Returning to a model with linear utility of scores, suppose grading effort generates disutility through a nonlinear function, $C\left(s-n_j\right)$, for $j=1,\ldots J$. In this case, what matters to sign $\frac{dp_j}{ds}$ is how the curvature of the cost function compares to the disclosure risk $\gamma(s)$. Assuming additive across-items costs, the maximand can be divided by s to obtain

$$\max_{\mathbf{p_j}} \alpha\left(\frac{1}{s}-\gamma(s)\sum_h p_h\right)\sum_h \tau_h p_h-\frac{1}{s}\sum_h C\left(s\left(1-p_h\right)\right).$$

The optimal p_j satisfies

$$\frac{\tau_j}{s}-\gamma(s)\sum_h\left(\tau_j+\tau_h\right)p_h+\frac{1}{\alpha}\sum_h C'\left(s\left(1-p_h\right)\right)=0,$$

which implies

$$\frac{dp_j}{ds}=-\frac{1+\gamma'(s)s^2\sum_h(1+\frac{\tau_h}{\tau_j})p_h-\frac{s^2}{\alpha\tau_j}\sum_h\left(1-p_h\right)C''\left(s\left(1-p_h\right)\right)}{2\gamma(s)s^2+\frac{s^3}{\alpha\tau_j}C''\left(s\left(1-p_j\right)\right)}.$$

For example, when the cost of honest grading is quadratic and convex

$$C\left(s-n_j\right)=\beta_1\left(s-n_j\right)+\beta_2\left(s-n_j\right)^2,$$

(with positive β_1 and β_2) we have $C'\left(x\right)=\beta_1+2\beta_2 x$ and $C''\left(x\right)=2\beta_2$, so that

$$\frac{dp_j}{ds}=-\frac{1+\gamma'(s)s^2\sum_h(1+\frac{\tau_h}{\tau_j})p_h-2\beta_2\frac{s^2}{\alpha\tau_j}\sum_h\left(1-p_h\right)}{2\gamma(s)s^2+2\beta_2\frac{s^3}{\alpha\tau_j}}.$$

Equation (3) in the text is obtained as special case when $\beta_2=0$. The last expression is negative if

$$\frac{1}{s}+\gamma'(s)\sum_h(1+\frac{\tau_h}{\tau_j})n_h>2\beta_2\left(\frac{\alpha\tau_j}{s}\right)^{-1}\sum_h\left(1-p_h\right).$$

The sign of $\frac{dp_j}{ds}$ is unclear, although for small enough β_2 and large enough $\gamma'(s)$ is most likely negative. A cost function with small curvature with respect to the utility gain ($\alpha\frac{\tau_j}{s}$ ($\beta_2\left(\frac{\alpha\tau_j}{s}\right)^{-1}\approx 0$) is a sufficient condition for $\frac{dp_j}{ds}<0$. Costs might also be concave if honest graders become more efficient when they grade more (negative β_2). In this case, a similar condition can be derived to sign $\frac{dp_j}{ds}$.

References

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

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.

Table A1: Reduced Form Estimates of the Effect of Maimonides' Rule on Class Size, Test Scores, and Score Manipulation

	Math			Language		
	Italy	North/Centre	South	Italy	North/Centre	South
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Class size					
Maimonides' Rule	0.513	0.555	0.433			
	(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. Test Scores					
Maimonides' Rule	-0.0031	-0.0023	-0.0056	-0.0021	-0.0012	-0.0041
	(0.0010)	(0.0009)	(0.0022)	(0.0008)	(0.0008)	(0.0017)
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. Score Manipulation					
Maimonides' Rule	-0.0009	-0.0003	-0.0020	-0.0008	-0.0003	-0.0016
	(0.0004)	(0.0002)	(0.0009)	(0.0003)	(0.0003)	(0.0008)
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

Notes: This table shows the reduced form effect of the Maimonides' Rule on class size (Panel A), test scores (Panel B), score manipulation (Panel C). All models control for a quadratic 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. All regressions include sampling strata controls (grade enrollment at institution, region dummies and their interactions). Other controls are listed in footnote 7 of the paper.

Table A2: First Stage Estimates for Over-Identified Models

	Class size			Score manipulation math			Score manipulation language		
	Italy	North/Centre	South	Italy	North/Centre	South	Italy	North/Centre	South
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Maimonides' Rule (f_{igkt})	0.704	0.753	0.617	-0.0009	-0.0003	-0.0021	-0.0014	-0.0008	-0.0024
	(0.006)	(0.007)	(0.011)	(0.0005)	(0.0003)	(0.0011)	(0.0004)	(0.0003)	(0.0010)
Monitor at institution (M_{igkt})	0.010	0.029	-0.013	-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)
2 students below cutoff	-1.427	-1.154	-1.865	0.002	-0.002	0.008	0.010	0.005	0.018
	(0.083)	(0.101)	(0.138)	(0.005)	(0.003)	(0.012)	(0.005)	(0.004)	(0.011)
1 student below cutoff	-2.258	-2.053	-2.580	0.001	0.001	0.000	0.007	0.009	0.002
	(0.093)	(0.116)	(0.150)	(0.005)	(0.004)	(0.012)	(0.005)	(0.004)	(0.011)
1 student above cutoff	2.411	3.026	1.519	0.000	0.003	-0.004	-0.001	-0.001	-0.001
	(0.097)	(0.132)	(0.138)	(0.006)	(0.005)	(0.013)	(0.005)	(0.004)	(0.012)
2 students above cutoff	1.247	1.546	0.826	0.001	-0.004	0.007	-0.007	-0.005	-0.012
	(0.083)	(0.114)	(0.120)	(0.006)	(0.004)	(0.013)	(0.005)	(0.004)	(0.009)
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 monitor at institution and dummies 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 monitor at institution and dummies for grade enrollment being in a 10 percent window (2 students) above and below each cutoff on score manipulation. All models control for a quadratic 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. All regressions include sampling strata controls (grade enrollment at institution, region dummies and their interactions). Other controls are listed in footnote 7 of the paper.

Table A3: Covariates and Maimonides' Rule with and without External Monitors

	Institutions with Monitor			Institutions without Monitor		
	Italy	North/Centre	South	Italy	North/Centre	South
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Administrative Data on Schools					
% in class sitting the test	0.0001	0.0002	0.0000	0.0000	0.0000	0.0000
	(0.0002)	(0.0002)	(0.0003)	(0.0001)	(0.0001)	(0.0002)
% in school sitting the test	0.0003	0.0003	0.0002	0.0001	0.0001	0.0001
	(0.0002)	(0.0002)	(0.0003)	(0.0001)	(0.0001)	(0.0002)
% in institution sitting the test	-0.0000	-0.0000	0.0001	-0.0001	-0.0002	-0.0000
	(0.0001)	(0.0002)	(0.0003)	(0.0001)	(0.0001)	(0.0001)
	B. Data Provided by School Staff					
Female	-0.0003	-0.0006	0.0001	0.0001	0.0005	-0.0003
	(0.0003)	(0.0004)	(0.0006)	(0.0002)	(0.0002)	(0.0003)
Immigrant	-0.0005	-0.0002	-0.0007	-0.0007	-0.0009	-0.0003
	(0.0003)	(0.0005)	(0.0003)	(0.0002)	(0.0003)	(0.0002)
Father HS	-0.0005	-0.0002	-0.0014	0.0010	0.0003	0.0020
	(0.0005)	(0.0006)	(0.0010)	(0.0003)	(0.0004)	(0.0005)
Mother employed	0.0001	0.0003	-0.0004	0.0015	0.0012	0.0022
	(0.0008)	(0.0010)	(0.0012)	(0.0004)	(0.0006)	(0.0006)
	C. Non-Response Indicators					
Missing data on father's education	0.0014	0.0012	0.0019	0.0000	0.0016	-0.0026
	(0.0011)	(0.0013)	(0.0020)	(0.0007)	(0.0008)	(0.0012)
Missing data on mother's occupation	0.0018	0.0017	0.0020	-0.0002	0.0012	-0.0028
	(0.0011)	(0.0013)	(0.0019)	(0.0007)	(0.0008)	(0.0011)
Missing data on country of origin	0.0006	0.0003	0.0011	-0.0002	-0.0002	-0.0003
	(0.0004)	(0.0004)	(0.0008)	(0.0003)	(0.0003)	(0.0006)
N	34,325	22,174	12,151	105,685	65,324	40,361

Notes: This table reports coefficients from regressions of the variables listed at left on Maimonides' Rule, controlling for a quadratic in grade enrollment, enrollment segment dummies and their interactions, grade and year dummies, and sampling strata controls (grade enrollment at institution, region dummies and their interactions). Columns 1-3 show results for the sample with monitors; columns 4-6 show results for the sample without monitors. Robust standard errors, clustered on school and grade, are shown in parentheses.

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]

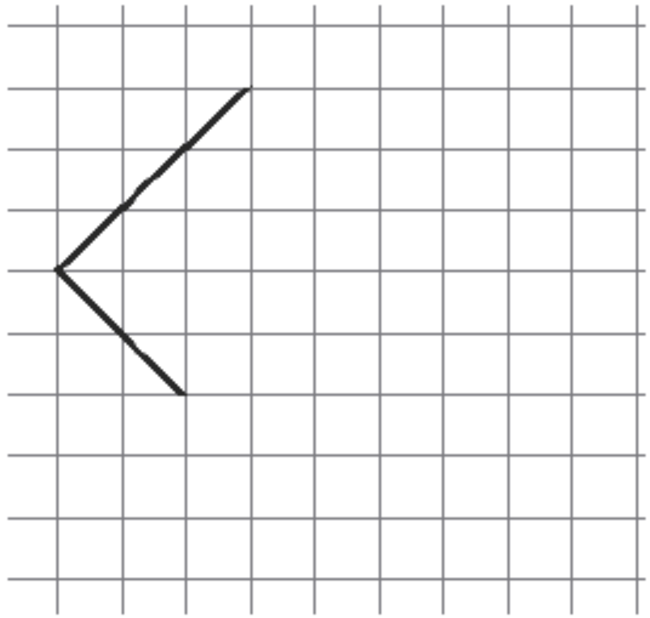
D1_a	☐V	☐F	☐NV		
D1_b	☐V	☐F	☐NV		
D1_c	☐V	☐F	☐NV		
D1_d	☐V	☐F	☐NV		
D2	☐A	☐B	☐C	☐D	☐NV
D3	☐0	☐1	☐NV		
D4_a	☐A	☐B	☐C	☐D	☐NV
D4_b	☐0	☐1	☐NV		
D5	☐A	☐B	☐C	☐D	☐NV
D6	☐A	☐B	☐C	☐D	☐NV
D7	☐A	☐B	☐C	☐D	☐NV
D8	☐A	☐B	☐C	☐D	☐NV
D9	☐0	☐1	☐NV		
D10	☐A	☐B	☐C	☐D	☐NV
D11	☐A	☐B	☐C	☐D	☐NV
D12	☐0	☐1	☐NV		
D13	☐A	☐B	☐C	☐D	☐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 A1: Answer sheet for V grade in 2010/11

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 A2: Example of open-ended question in math test - V grade 2010/11

C4. Nella frase che segue inserisci le parole mancanti scegliendole da questa lista: *così, dove, perché, però, se, siccome.*

……………. non conoscevo la strada, ho chiesto a una signora ………. dovevo andare; …………….. non mi sono perso.

Figure A3: Example of open-ended question in language test - V grade 2010/11