PROGRAMMA EDUCATION
FGA WORKING PAPER

N. **50** (12/2013)

Rationale and incentives for cheating in the standardised tests of the Italian assessment system

Gerard Ferrer-Esteban

Fondazione Giovanni Agnelli

Fondazione
Giovanni Agnelli

www.fondazione-agnelli.it **segreteria@fga.it**

Rationale and incentives for cheating in the standardised tests of the Italian assessment system [1]

Gerard Ferrer-Esteban [a b]

[a] Fondazione Giovanni Agnelli

[b] Interdisciplinary Group on Educational Policies

(IGEP), University of Barcelona

ABSTRACT

Academic achievement levels are overestimated in some Italian schools and areas due to cheating in the Invalsi standardised tests. This phenomenon has already been highlighted by several studies which report on significant predictors, such as the presence of controllers in classrooms and schools, the geographical location of the schools, and the high-stakes accountability system with which the tests are linked. Some research has also pointed out how school cheating may reveal a lack of social and civic capital of students or teachers. While the volume of existing research on student cheating is significant, research on the determinants of teacher cheating is limited. This paper proposes a comprehensive framework to identify the rationale and incentives behind the illicit behaviours of those who should respond to deontological professional standards in the compulsory education system. Logistic regression models at classroom level have been carried out for teachers in the 5th year of primary education, the 1st and 3rd years of lower secondary education, and the 2nd year of upper secondary. The results are presented within the two main sections of this framework: *teacher cheating as a result of the bonding form of network-based civic capital*, and *teacher cheating as a lack of culture-based civic capital*. Firstly, our findings suggest that teacher cheating is more likely in socially homogeneous classrooms, where strong ties are more likely. In these homogeneous learning environments, cheating may be understood as a form of community-based support, since it is found to help, to a large extent, the more disadvantaged students. Secondly, our results reveal that teacher cheating is consistently associated with non-civic-minded practices undertaken by schools and administrators, which do not match legal requirements, or recommendations, such as social tracking of students between classes and exclusion of students from tests. Finally, the findings confirm a strong association between teacher cheating and the cultural context in which schools are located: the more deviant behaviours – illicit behaviours against laws and regulations – are found in the location of the school, the more cheating behaviours are found among teachers. The main results are preceded by confirmatory general findings on the geographical location of schools, the effect of high-stakes testing and external controllers during the test taking.

[1] Fondazione Giovanni Agnelli (Turin, Italy) promoted this research in the framework of the book entitled 'La valutazione della scuola' (2014), on the Italian school evaluation system. The author would like to especially thank Martino Bernardi for his excellent research assistance, Gianfranco De Simone for helpful suggestions on econometric modelling, and the rest of the Foundation's staff for helpful general comments and discussions. He is grateful to Stefano Molina (Fondazione G. Agnelli) and Jorge Rodríguez Menés (Universitat Pompeu Fabra) for providing helpful theoretical insights to further understand cheating behaviour. He is also grateful to participants in workshops at Irvapp (Trento) and Invalsi (Rome) for comments, to Patrizia Falzetti (Invalsi) for help with the data, and Sergio Longobardi (Università degli Studi di Napoli 'Parthenope') for providing information on the cheating detection method used by the Invalsi.

INTRODUCTION: A GLANCE AT THE PHENOMENON OF SCHOOL CHEATING

Standardised assessment systems may be altered by many illicit actions aimed at increasing the overall level of student performance. As stated in an OECD report on current practices of student standardised testing, "although the evidence is unclear as to whether standardised tests lead to improved student outcomes, there is more certainty that standardised tests lead to increased strategic behaviours on the part of schools and teachers" (Morris, 2011). This is particularly true when the tests are framed in high-stakes accountability systems, which have extrinsic incentives that encourage teachers and schools to raise aggregated student scores. Among different actions to raise the test scores are those of 'teaching to the test' and 'narrowing the curriculum', which, although not illicit, may have negative effects on the quality of teaching and instruction. In contrast, we focus our attention on illicit actions undertaken by teachers, aimed at altering the realisation and the outcomes of the tests: cheating in the test and the exclusion of certain students from taking the test.

There are many forms of cheating, and these may be undertaken by both students and educators: copying, suggestion of correct answers, 'adjustments' while checking or filling the answer sheets, etc. Other strategies, not focused on the test, can be alteration of the composition of the students who are being tested. Students may be wrongly classified as students with special needs (so excluded from aggregated scores), or simply advised not to go to class on the day of the test. The consequences resulting from dishonest behaviour are multiple, but it is worth mentioning those of great importance, such as invalidation of the whole system of assessment and accountability (for both the authorities and the families), alteration of the system of student promotion and retention, or the unreliability of policies of teacher incentives.

Several cases of teachers cheating have been identified in many countries, published mostly in reports and newspaper articles (Amrein & Berliner, 2002; Nichols & Berliner, 2002). Cases of cheating by teachers have been periodically reported to the public attention many times in the United States, triggering a broad public debate on the reliability of the high-stakes testing system. An authentic scandal across the nation arose, for example, after widespread cheating in the majority of public schools in the district of Atlanta was uncovered in 2009. In the seven previous years, the students of this district achieved results significantly above average in the Criterion Referenced Competency Tests. The standardised tests are administered by the state of Georgia, and their outcome determines federal funding of the No Child Left Behind programme. After an article in a local newspaper denouncing the anomaly of the tests results, in 2011, there was an official inquiry that revealed systematic cheating in 44 of the 56 schools taking part in the test. The investigation showed that at least 178 teachers and school administrators had altered the test results, mainly through 'correction' of wrong answers. Attributing their unlawful conduct to the 'enormous pressure' for achieving better results, 82 of them admitted their responsibility. The teachers attributed the pressures to the social environment (parents,

media), but particularly to the educational authorities and the accountability system based on explicit threats of dismissal in case of failure. Although this is the most notorious cheating scandal in the United States, it is not an exception. A 2013 report sent by the U.S. Government Accountability Office to the Secretary of the U.S. Department of Education described other incidents in Illinois, Maryland, Pennsylvania, Texas, Washington DC, and in California, where the results of 23 schools were invalidated for cheating by school administrators and teachers in 2012.

As regards the practice of excluding a certain profile of students from the test, numerous cases have been identified internationally. Students with more learning difficulties are excluded in order to increase the aggregate average of classes and schools, especially in countries with high-stakes tests such as Canada, the United States and the Netherlands. In the Netherlands, for example, the inspectorate has reported how, in some cases, students who were more likely to be sent to less prestigious school tracks did not take part in the high-stakes tests (Mons, 2009). In the case of Ontario, in Canada, teachers reported that some schools had gone too far in reducing the number of students so as to bring up the average (Bélair, 2005, cited by Mons, 2009). In the United States, numerous cases have been identified in several states such as California, Texas and Alabama, where students were excluded to raise the average scores (Haladyna et al., 1991; Madaus et al., 1992). Finally, in Italy, anecdotal evidence was found of teachers openly encouraging students to stay home as a form of opposition to the eventual future use of the test as a control of teacher quality – although this use has always been rejected by evaluation authorities (Paccagnella & Sestito, 2013).

To date, there have been few research studies aimed at examining the extent of cheating in the stages of compulsory education (primary and lower education), and even fewer dealing specifically with cheating led by teachers. In the United States, where more research is carried out, studies are mainly focused on the relationship between the incentive systems (rewards and sanctions) and dishonest behaviours. In Chicago schools, for example, it has been estimated that, every year, there is a minimum of 4-5% teacher cheating in elementary schools, and that this phenomenon is associated with minimal changes in incentive schemes, which lead to significant distortions in conducts (Jacob & Levitt, 2003). In other countries analysed, such as Hungary, similar levels of teacher cheating to that in American schools have been reported (Horn, 2012).

In the Italian case, recent publications have revealed the presence of the phenomenon of cheating. The first researchers who found evidence of cheating in the Italian standardised tests were Quintano, Castellano and Longobardi (2009), who developed a two-stage method for detecting it. Though their primary aim was to propose a method to identify cheating, the authors also pointed out certain factors associated with it, such as the geographical distribution of outlier units, highlighting how the phenomenon is non-randomly distributed across Italian regions. The authors pointed out, for the first time, that

there were higher levels of cheating in southern regions, hypothesising that teacher support to the students depended on the school location. Other studies went further and tried to identify *motivations*, such as social capital based on particularistic values (Paccagnella & Sestito, 2014), *deterrents*, such as external control during tests or social capital based on universalistic values (Bertoni et al., 2013; Paccagnella & Sestito, 2014), and *incentives*, such as student peer effects (Lucifora & Tonello, 2012).

UNDERSTANDING CHEATING: AN INTERDISCIPLINARY ROUTE TO COPE WITH DEVIANCE

The rationale and motivations behind turning to illicit behaviours obviously depend on the goals of those who cheat, which are strongly conditioned by the characteristics of the accountability and standardised assessment systems. For instance, it has been widely demonstrated in the research literature that *high-stakes* tests tend to increase the likelihood of students and teachers cheating. The more the test results have formal consequences for students, teachers or schools, the more opportunistic behaviours will be encouraged. Not surprisingly, the negative consequences of high-stakes accountability systems are often explained by the adage of Donald T. Campbell (1976), which has become a principle in social science: "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor."

Nevertheless, depending on the motivations, and even independently from the existence of schemes of explicit incentives, illicit behaviour may be seen differently alongside the historic and cultural context. Social acceptation and/or relativisation of cheating behaviour may reinforce its social legitimisation, so contributing to its dissemination. This section deals with a theoretical framework that may help to understand the incentives or goals of teachers who cheat.

Behind cheating: understanding deviance and moral disengagement

Cheating has been largely understood as a form of deviant behaviour, and while *deviance* has been mainly approached in both sociological and psychological theories, it has also been approached in economic studies. Depending on the sociological framework, particularly in *culturalist* theories, cultural factors play an important role in explaining behaviours lying outside established institutional norms, which in turn do not always coincide with the norms of certain social sub-culture groups. People learn to perform illicit actions through socialisation with others, by assuming or rejecting values, norms and beliefs. In this context, an individual does not always perceive their actions as significant deviations because they are consistent with the values and norms of the social reference

group. This means that deviant actions are perceived in accordance with the values of a group. And therefore justified.

Within the framework of these theories, it is worth mentioning the *structuralist approach*, which considers it important to recognize that actions considered illicit may be a manifestation of the cultural system and be strongly correlated with the social structure. Deviance is socially constructed and linked to rules of behaviour created in a particular society. This idea is further developed through *social learning theories*, which have been widely developed from the *interactionist approach*. From this perspective, illicit behaviour is learned through social interaction. By means of social learning, people interiorise a definition of deviation, which may have positive or negative connotations. When the act undertaken has a socially negative connotation it is justified by the context, in response to certain situations, exogenously determined.

One of the most influential and controversial learning theories of the sociology of deviance is *differential association*, which has focused on the direct learning of specific normative definitions (Akers & Jennings, 2009; Akers et al., 1979; Lanza-Kaduce & Klug, 1986; Sutherland, 1947). From this perspective, behaviour is predicted by the *differential association* with groups that individuals deal with, to the extent that social groups provide the context within which learning occurs: whether individuals adopt deviant behaviour depends on the frequency and intensity with which they are exposed to it. In the context of social learning theory, deviant behaviour is learnt in personal interaction, individuals learn by modelling others, and the primary learning mechanism is the *operant conditioning* linked to differential reinforcements, negative or positive (Akers & Jennings, 2009; Lanza-Kaduce & Klug, 1986).

Social learning is linked with *social cognitive* perspectives based on the behaviourist tradition and, although to a lesser extent, with the theories of *moral cognitive development*. Actually, social learning theory also developed from psychology, understood here as *social cognitive theory* (Bandura, 1986), whereby behaviours are modelled by means of both intrinsic and external, social, reinforcement. Moral standards, which serve as guides and deterrents for action, are constructed from evaluative social reactions to one's conduct, as well as exposure to the self-evaluative standards modelled by others in the course of socialisation (Bandura et al., 1996). Since moral agency is grounded in a self-regulatory system, according to which individuals self-monitor, judge and self-react, types of conducts may be substantially different, even if moral standards are the same. We can talk about *moral disengagement* when people do not self-react and do not activate this internal control, behaving in a detrimental way. People engage in reprehensible conduct to the extent to which they have made socially acceptable and justified the morality of their actions (Bandura, 2002; Bandura et al., 1996).

From social learning theories it is understood that moral development refers to the conformity of individuals to moral rules – in terms of behaviour and affection – rather than

a structural cognitive change. However, several studies from social constructivism have argued how moral development – *reasoning about justice* – responds to cognitive development stages, culturally universal, which can be stimulated to a greater or a lesser extent, depending on the cultural and social context. These cognitive-developmental theories were born from the work of Jean Piaget (1932) and developed with the theories and research of several authors. One of these, Lawrence Kohlberg, developed a model of moral stages (Kohlberg, 1976, 1981), with three levels of six stages, according to cognitive changes occurring in a sequential pattern, from less to more rationality, to evaluate situations of moral relevance. According to this model, the relationship between moral development and deviant behaviour is indirect and interacts with situational forces (Lanza-Kaduce & Klug, 1986). It is through experiences of social interaction, rather than internalization of norms, which structure the basic rules and principles (Kohlberg, 1976).

This model is able to categorise behaviours, which may respond to different moral developmental stages. Following phases 1 (following the rules to avoid punishment) and 2 (following the rules in response to their own needs and interests) of moral development (up to 9-11 years), which are in the *pre-conventional level*, there are the two stages of the *conventional level*. This level includes most adolescents and adults, in which moral development is related to the concern about social approval, loyalty to persons, groups, and authority, and the concern about the welfare of others and society (Kohlberg, 1976). At this level, individuals understand, accept and support the values and norms of both their reference groups and society as a whole. There is a significant difference between the two stages at this level: in stage 3 the focus is on orientation towards relevant close groups such as family and friends, while in stage 4 it is on institutional authorities, such as normative framework and the law. In line with this framework, deviant behaviours would be seen as a result of potential conflicts between stages. Early moral stages, which give priority to own needs and interests of individuals or close local groups, would be detrimental to the advanced stages, in which importance is given to collective regulations and the priority is focused on the concern for the welfare of the whole society.

Approaching cheating associated factors: social and civic capital

It is worth noting how the two stages of *moral reasoning* of Kohlberg agree with, to some extent, the main types of social capital understood as network-based civic engagement based on reciprocity and trust (Gittell & Vidal, 1998; Putnam, 1995, 2000; Woolcock & Narayan, 2000). The third stage relates to an individual's moral prerequisite for developing the strong ties of *bonding social capital* established between members of a community with a homogeneous composition. In contrast, the fourth stage relates to the moral precondition for developing the weak ties of *bridging social capital* amongst different social groups or socially heterogeneous groups. Here, social capital is understood as a collective *stock*, and refers to the norms and networks that enable people to act collectively (Woolcock &

Narayan, 2000) in order to benefit collectively (Portes, 2000). This feature of collective stock – social capital as a feature of communities – and collective benefits means this conception diverges from the sociological conception, according to which ties are established to yield benefits to individuals (Coleman, 1988; Portes, 2000). Although social capital is a network-based conception, it has a strong cultural base, since it is basically understood as a civic culture that can be collectively used (Trigilia, 2011).

In both forms of social capital, the networks and the associated norms of reciprocity are valuable in terms of trust, solidarity and mutual support. Both society-based and community-based networks may provide resources to support the most disadvantaged members or groups. Nonetheless, network-based social capital may have negative consequences as well. For instance, *bonding* social capital is potentially *exclusive*, since socially homogeneous groups have self-referenced interests which may diverge from those held by other groups of the same society, or from those of society as a whole. In these cases, this form of social capital may be detrimental to the *bridging* form, which by definition is concerned with solidarity, mutual respect and cooperation, values related to the welfare of the society as a whole. Indeed, as pointed out by Portes (1998), the strong ties which bring benefits to members of a group generally restrict access to outsiders.

Nevertheless, according to Solow (1995), as these concepts do not comply with certain criteria, from an economic point of view, they cannot be considered as *social capital*: as a stock of capital, it should be measurable; it should have a non-negative economic *payoff*; and mechanisms through which social capital can be accumulated and depreciated should be defined. To meet these criteria, we should consider the definition of social capital proposed by Guiso, Sapienza and Zingales (2006, 2010) as *civic capital*: "those persistent and shared beliefs and values that help a group overcome the *free rider* problem in the pursuit of socially valuable activities". The authors do not refer to a network-based civic capital, but to a culture-based one, where social capital is about values and beliefs which are shared by a community and persist over time. Moreover, this civic capital is related to all types of economic interactions and not restricted to political participation (Guiso et al., 2010).

The authors point out that relevant direct measures of civic capital may identify values that induce people to be against actions that give private benefits at high social costs (Guiso et al., 2010). Specifically, they refer to opinions about *free riding* and other behaviours which deviate from the public good (e.g. tax evasion or avoidance, littering). In fact, as mentioned earlier, no form of social capital has a negative economic pay-off. Nevertheless, it is worth noting that this definition of social capital does not deal with the dichotomy between *universalistic* and *particularistic* values (de Blasio et al., 2014), which to some extent runs parallel with *bridging* and *bonding* types of network-based social capital. Paccagnella and Sestito (2014), in their work on cheating and social capital, suggest that certain particularistic social values, approached as social capital, may lead to negative externalities.

RESEARCH QUESTIONS

The main objective of this paper is to deal with the incentives and rationale that explain teacher cheating in the standardised tests of the compulsory education system. The theoretical perspectives of social learning and moral cognitive development highlight the importance of social interaction between individuals and their institutional, social and cultural context, on behaviour. In addition, they frame the rationale whereby cheating behaviour may occur: while social learning and social cognitive theory emphasise how external reinforcement may drive self-justified deviant behaviour, the cognitive model of moral development aims to place deviant behaviour in stages when moral thought develops. We use approaches to moral disengagement or development as a general interpretative framework to reasonably understand the findings, but also to link these morality theories with the culture-based civic capital approach.

As regards the empirical work, we focus on the two perspectives approached from the *social capital framework*: the *network-based civic capital* and *culture-based civic capital*, which provide further insights on how cheating may be understood and explained. The *network-based approach* of civic capital is understood as a collective stock and refers to the norms and networks that enable people to act collectively. The norms and networks of reciprocity are valuable in terms of trust and solidarity, and can provide resources to support the most disadvantaged members or groups – even if, as in the case of teacher cheating, it is detrimental to the society as a whole. In contrast, the *culture-based approach* of civic capital refers to collectively shared beliefs and values that help a group overcome *free rider* problems (Guiso et al., 2010). Here cheating would predict the lack of civic capital, since it is potentially associated with shared behaviours that deviate from the public good.

Along with these approaches, we aim to find the aggregated factors associated with teacher cheating, which are distributed in three levels of analysis: classroom, school and province (Table 1). Firstly, in line with the framework of *network-based civic capital*, we wanted to observe to what extent cheating behaviours are more likely in classrooms where there are higher stocks of *bonding* social capital (see table 1, columns *b* and *c*). Furthermore, we also wanted to confirm whether teacher cheating, understood as a form of community-based support, is addressed to help, to a large extent, the more disadvantaged students (e.g. low SES or grade-retained students). We based this on the hypothesis that strong ties established in socially homogeneous groups may benefit their members, in this case in the form of teacher help for students who have difficulties with the test, while being detrimental to society at large – since teacher cheating undermines the monitoring and accountability objectives of the testing systems.

Secondly, in line with the *culture-based civic capital*, we wanted to observe to what extent cheating behaviours of teachers are associated with both *school-based and context-based deviant behaviours*. For this we used indicators of behaviours which are collectively shared

and deviate from the public good (of the whole society), that is, actions that provide private benefits at high social cost. As shown in Table 1 (columns *d* and *e*), as regards school culture, we analysed to what extent cheating behaviours are more likely in schools which undertake practices that do not match legal requirements, or recommendations, such as social tracking of students between classes, and exclusion of students from tests. We hypothesise that teachers in schools with lower levels of civic-minded capital are more likely to cheat. Although deterrents such as external controllers during the test taking are effective, they may be less effective in preventing cheating in schools where non-civic-minded values are deeply rooted (schools with the lowest levels of civic capital). On the other hand, as regards the culture in which schools are located, we analysed whether teacher cheating was more likely in contexts with higher indices of illicit behaviours, against laws and regulations. Here we start from the hypothesis that teachers in schools located within contexts which have a lack of civic capital – in which there are high rates of teacher absenteeism or high levels of illicit economic activities – are more prone to cheat in standardised tests.

Two theoretical assumptions were made to justify why we use deviant behaviours as a proxy of (lack of) civic social capital. First, as in Guiso et al. (2010), when defining a cultural approach of civic capital, we assume that aggregated measures of deviant behaviour reflect non-civic-minded beliefs and values, which are shared by a community. This assumption has its theoretical roots in the above-mentioned *interactionist approach* of *culturalist* theories: deviance may be socially constructed and linked to rules of behaviour created in a particular social sub-group, in which illicit behaviours are interiorised and justified to respond to exogenously determined situations. Secondly, evaluating moral-related behaviours may be an asset, compared to evaluating self-reported moral-related values and beliefs, especially when we refer to deviant culture. Here a clarification is needed: we argue that culture-based civic capital is strongly linked to moral standards as defined by the *social cognitive theory*. As stated earlier, according to this theory, people avoid behaving in a detrimental way when they self-regulate and activate internal controls in accordance with those moral standards (Bandura, 2002; Bandura et al., 1996). Moral standards, which are socially modelled, are translated into actions and behaviours through which moral agency is exercised (Bandura et al., 1996). However, a self-regulatory system is not invariant, so conducts may differ significantly, even if moral standards remain constant. The agreement, but also the lack of accordance, between moral standards and effective behaviours may support the use of deviant behaviours as precise approximations of non-civic-minded values with consequences on the public good. Potential problems of omitted variables related to using behaviours as indirect measures of civic capital are dealt with in the section on the econometric strategy.

Table 1. A comprehensive framework of cheating-related factors

			Social capital			
			Network-based civic capital		Culture-based civic capital	
			Bonding social capital		*Culture-based deviant behaviours*	
		Incentives to cheat	Conditions for cheating	Rationale for cheating (compensation function)	School-related factors	Context-related factors
		[a]	[b]	[c]	[d]	[e]
system level	High-stakes tests	✓				
classroom level	Opportunity (lack of control in classrooms)	✓				
	Homogeneity of social composition		✓			
	Classrooms with low SES average			✓		
	Classrooms with grade-retained students			✓		
school level	Exclusion of students from tests				✓	
	Social tracking between classrooms				✓	
province level	Teacher absenteeism				✓	✓
	Tax evasion (Italian public radio-television)					✓
	Estimated risk of tax evasion					✓
	Estimated rate of undeclared work					✓

These two blocks are preceded by two general – and minor – research questions that refer to incentives to cheat associated with the institutional and physical conditions of test-taking, previously highlighted in other research papers (see table 1, column *a*). Firstly, we wanted to confirm the results of Jacob & Levitt (2003), who found that teachers are more likely to fall into illicit behaviours when they are within high-stakes testing systems. The Italian testing system is high-stakes for 8th grade students, not for teachers, since the test results partially contribute to the final mark when leaving lower secondary school. However, effects of an increase in teacher cheating are expected, especially in contexts where bonding-related support for disadvantaged students is observed. Secondly, we

analysed whether teacher cheating is mitigated when tests are directly controlled in the classroom, and whether there are spill-over effects when there are external controllers in other classrooms in the same school. As confirmed elsewhere in numerous studies (Bertoni et al., 2013; Ferrer-Esteban, 2013; Paccagnella & Sestito, 2014), we expect both direct and spill-over effects of external controllers to reduce cheating.

DATA AND METHODS

The INVALSI student performance dataset

In this paper the dataset of standardised tests administrated by the Italian National Institute for the Evaluation of the Education System (Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione, Invalsi), in the 2009/10 and 2011/12 academic years, was used. Although the main analyses focused on the 2011/12 year, we used the 2009/10 data as it included that for 8th graders. This study dealt with the surveys on 5th, 6th, 8th and 10th grade student performance. In the Italian education system structure, these grades correspond, respectively, to the 5th year of primary education, the 1st and 3rd years of lower secondary education, and the 2nd year of upper secondary. While participation in the survey was voluntary for the schools in the 2008-2009 academic year, the national survey was obligatory for all schools from year 2009-2010. For this research, in the 2009/10 wave the estimation models constructed covered approximately 1,500,000 students and 82,500 classrooms, from the 5th, 6th, and 8th grades (on average, 10, 11 and 13 years old). The models in the 2011/12 wave included approximately 1,426,000 students and 79,100 classrooms from the 5th, 6th and 10th grades (on average, 10, 11 and 15 years old), spread over five macro-areas, 20 regions and 103 provinces. The test covered mathematics and Italian language, and were administered by teachers following a protocol set by Invalsi. This protocol suggested that the presence of teachers not specialised in the subject under test would be appropriate. External inspectors were sent to a sample of classrooms, in schools randomly selected across the regions, to control the realisation of tests, to check the answer sheets, and return the results to Invalsi. As stated earlier, only the tests carried out by 8th graders were high-stake for students, since they partially contributed to the student mark at the end of the lower secondary.

A cheating indicator based on suspicious answer strings

The binary dependent variable was whether a classroom was suspected of cheating in maths and reading tests. Many methods to identify cheating in schools have been developed, most designed to detect cheating by students, with much less attention to developing methods to identify cheating promoted by teachers. In the literature, research reports and articles propose or compare a number of different methods to identify student cheating in multiple

choice tests (Angoff, 1974; Belleza & Belleza, 1989; Frary, 1993; Sotaridona, 2003; Sotaridona & van der Linden, 2006; Wesolowsky, 1999). In contrast, there are very few methods to identify the teacher cheating. The most relevant is that elaborated by Jacob and Levitt (2003), based on two indicators: *unexpected test score fluctuations* and *suspicious answer strings*. This is the method that we tested in the present study.

In Italy, as mentioned earlier, since 2012 the system's evaluation authorities have used an algorithm developed by Quintano *et al.* (2009), although today with and additional indicator. This focuses on the presence of outlier units that may introduce an upward bias in the distribution of average scores by class. The first stage in this method consists of computing four indexes of response behaviour, such as the class mean score, standard deviation of the mean score, the class non-response rate, and the index of homogeneity of answers. The second stage weights every class based on the probability of belonging to the set of outlier units, which is calculated by a fuzzy clustering algorithm (Quintano *et al.*, 2009). This method is compatible with the identification of teacher cheating, but may also identify cheating by students when individual illicit behaviour is conducted by a large section of the class.

However, this method proved to be unsatisfactory in both theory and practice, because to a large extent, with the indicators chosen, classrooms that 'cheat' and truly excellent classrooms could not be distinguished. This algorithm gave both types of classrooms a high probability of cheating, which implied 'adjusting' their average score downwards. This high risk of detecting 'false positives' may obviously result in a lack of confidence by school operators, not just in the reliability of the testing system, but also in the evaluation authorities. Nevertheless, in response to the complaints received from the schools unfairly penalised and to the methodological criticisms from the scientific community, the Invalsi is now testing a new correction algorithm, in which a fifth indicator has been introduced. This new indicator takes into account the students' background, and also compares the results of Invalsi standardised tests with those of summative assessment by teachers at the end of the first term of the school year. Using this new indicator, a previously suspicious class is unlikely to be 'corrected' if the models carried out with both 'previous performance' and 'student background' confirm that the outlier pattern is plausible.

As regards to this research, and since we are mainly interested on teacher cheating, we partially replicate a method proposed by Jacob and Levitt to detect illicit behaviours adopted by teachers (Jacob & Levitt, 2003). Their method combines two indicators. The first is the *Unexpected Test Score Fluctuations*, which basically refers to unexpected scores gains that can be explained by cheating. A classroom will be suspected of cheating if unexpectedly large gains are followed by lower than usual test score gains for the same students the following year. When test scores are monitored over time, student gains due to talented teacher or rich educational programs are likely to be permanent. Unfortunately, this first indicator could not be calculated due to the lack of longitudinal data in the Invalsi

dataset. Instead, we focused on the second, a composite indicator based on the detection of *Suspicious Answer Strings*. With this second indicator, different and complementary ways for a teacher to cheat could be detected – not only the easiest, but also more sophisticated actions. Specifically, it is a combination of four measures of *Suspicious Answer Strings* (Jacob & Levitt, 2003): the unlikely block of identical answers given to consecutive questions in the classroom, the classroom average variance across all test items, the variance (as opposed to the mean) in the degree of correlation across questions within a classroom, and the extent to which a student's response pattern differs from other students with the same aggregate score that year. The overall measure of cheating is constructed, within a given subject and grade, ranking classrooms on each of the four indicators, and taking the sum of squared ranks across the four measures. In the empirical work we use the 90^{th} and the 95^{th} cut-offs to identify the classrooms suspected to cheat. An extended description of the construction of these measures is provided in the appendix.

Cheating-related factors

To analyse the phenomenon of cheating, we considered variables at three levels: classroom, school, and province (table 2). As noted in the section on research questions, the selection depended not only on the availability of data but also on the theoretical framework. As regards the variables related to *network-based civic capital*, calculated at the classroom level, we deal with both the conditions and the rationale for cheating. The first variable, which responds to the question of the extent to which *bonding social capital* is a good predictor of cheating behaviour, is a measure of social composition dispersion within classrooms. Since *bonding* is characterised by strong ties established between members of a community with a homogeneous social composition, we expect that a measure of social dispersion will give information on the degree of classroom social homogeneity. The measure chosen was the standard deviation of the student economic, social and cultural status (ESCS) index within classrooms, which indicates the dispersion of the socioeconomic index from the overall classroom average. The formula is expressed as:

$$\sigma_{ESCS_{cs}} = \sqrt{\frac{\sum_{i=1}^{n}(ESCS_i - \overline{ESCS}_c)^2}{n-1}} \quad (1)$$

where $\sigma_{ESCS_{cs}}$ is the standard deviation of the ESCS index in classroom *c* in school *s*, $ESCS_s$ is the socioeconomic status of student *i*, and $\overline{ESCS}_c$ is the classroom average of the student ESCS index.

To test whether teacher cheating is used to help socially and academically disadvantaged students, we used variables for student composition, aggregated at the classroom level. First, we included the fraction of grade-retained students, then, we included the classroom

socioeconomic status using dummies of the quarters of the ESCS index. We also used the ESCS classroom average, but the sign and significance of the results was not affected. Working with quarters of the index provided more information as it allowed us to confront the propensity to cheat in low, mid and high SES classrooms.

As regards the variables related to *culture-based civic capital*, calculated at school and province level, we used factors of lack of civicness and expressing deviant behaviours. For the potential predictors of deviant behaviours at school level, we used two measures. The first was the fraction of students who missed the test, as a proxy of the percentage of students excluded from the test taking. Since we assume that students who missed the test for justified reasons are randomly distributed across schools, which can be understood as a lower bound, we expect that significant amounts of students intentionally excluded from the test will be clearly appreciated.

Second, as a proxy of deviant behaviour linked to a certain school culture, we used a measure of social tracking between classrooms within schools. Since student socioeconomic background is very unlikely to be an explicit criterion for classroom composition, we hypothesise that social segregation between classrooms is a product of family pressures, teacher expectations or pedagogical decisions. All these practices may aim to aggregate students at a homogeneous level of academic abilities in the same classrooms, and considering that student performance is highly related to social origin, we may reasonably expect that it leads, to some extent, to social segregation within schools. It is responsibility of school administrators and the board to avoid such situations: beside ethical considerations, it also contravenes the recommendations of education authorities, as it disrespects the principle of 'equity-heterogeneity' in classroom student composition. To the extent that these practices directly depend on the school autonomy, and that they may effectively lead to unequal educational opportunities for students (Agasisti & Falzetti, 2013; Ferrer-Esteban, 2011), we argue that social dissimilarity between classrooms may mirror the non-civic-minded decisions and actions of the school agents. As in the case of homogeneity of the classroom social composition, we used the standard deviation, from the school average, of the socioeconomic index across classrooms. This is expressed as:

$$\sigma_{ESCS_S} = \sqrt{\frac{\sum_{s=1}^{n}(ESCS_c - \overline{ESCS}_s)^2}{n-1}} \qquad (2)$$

where σ_{ESCS_S} is the standard deviation of the classroom ESCS index in school *s*, $ESCS_c$ is the average socioeconomic status of classroom *c*, and $\overline{ESCS}_s$ is the school average of socioeconomic status. Robustness checks have been undertaken with alternative measures of social dispersion as proxies of social tracking practices within schools, such as the variance between classrooms within schools (Agasisti & Falzetti, 2013; Ferrer-Esteban, 2011).

For the predictors of deviant behaviours at the province level, we used the percentage of teacher absenteeism. In addition to mobility and retirement, this could be considered one of the main factors with negative consequences for didactic and teaching continuity. When this absenteeism is quantitatively relevant and differs significantly across provinces, we can talk of a phenomenon related to the civic capital of both the social context and the local school system. Finally, with regard to the context-related cultural factors potentially associated with school cheating, we used three measures of deviant behaviours, at the province level, which refer to illicit economic activities existing alongside the Italian official economy. Firstly, we used a tax evasion indicator related to public radio and television fees, as provided by the economic newspaper "Il Sole 24 Ore" (23 January 2014). Since anyone who has a broadcasting signal reception device (terrestrial or satellite) for radio or television must pay the tax (except for people over 75, and very low income households), which means most Italian households, it becomes a good predictor of how widespread the lack of civic capital is across the country. Secondly, we used the risk of tax evasion, as provided by "Centro Studi Sintesi" and published in "Il Sole 24 Ore" (27 August 2012). The rate of evasion risk was calculated as the gap between the level of welfare and disposable household income. Low scores indicate a potential risk of tax evasion, while high scores indicate that the levels of well-being are, on average, lower or in line with the household income. Finally, the last indicator is related to the submerged economy, and expressed as the ratio between the number of missing income tax payers and the population, over 15 years old, not unemployed or inactive because of being a housewife or student (year 2011). This indicator related to undeclared work was also provided by the "Centro Studi Sintesi" based on data from the Italian Finance Department and the Italian National Institute for Statistics (Istat), and published in "Il Sole 24 Ore" (17 June 2013).

Finally, different dummy variables were used for the research questions concerning the incentives to cheat. First, with regard to the presence of external controllers, we introduced two dummy variables to indicate, respectively, whether test taking was directly monitored by an external controller, and whether it was not directly, but indirectly, controlled by external controllers in other classrooms of the same school. As regards the high-stakes nature of the test, whether the class was the last year of primary education, or the first year or third year of lower secondary school was taken into consideration. According to the research hypothesis, teachers of the third grade of middle school should most likely be prone to cheat, as they would tend to help more in tests with formal consequences for students. To deal with differences between grades we ran regression models, separated from the main models of the empirical work. Since the cheating measure is based on the 90^{th} percentile cut-off of a ranking measure of cheating, we had to pool the datasets of the three grades and then calculate cheating across the grades to determine how cheating is distributed across the grades.

Table 2. Descriptive statistics

	Primary education					Lower secondary education					Upper secondary education				
	mean	sd	min	max	freq	mean	sd	min	max	freq	mean	sd	min	max	freq
Suspected classroom of cheating (90th percentile cut-off)	0.11	0.31	0.0	1		0.11	0.31	0.0	1		0.11	0.31	0.0	1	
Suspected classroom of cheating (95th percentile cut-off)	0.06	0.24	0.0	1		0.06	0.24	0.0	1		0.06	0.24	0.0	1	
Unmonitored classrooms in unmonitored schools	0.76	0.42	0.0	1	19453	0.70	0.46	0.0	1	16567	0.57	0.49	0.0	1	11827
Monitored classrooms	0.07	0.25	0.0	1	1676	0.08	0.27	0.0	1	1925	0.10	0.30	0.0	1	2132
Unmonitored classrooms in monitored schools	0.17	0.37	0.0	1	4305	0.22	0.41	0.0	1	5124	0.32	0.47	0.0	1	6702
Social student heterogeneity in classroom	0.86	0.26	0.1	2		0.88	0.24	0.1	2		0.86	0.16	0.3	2	
Low-SES classrooms	0.33	0.47	0.0	1	8478	0.33	0.47	0.0	1	7872	0.33	0.47	0.0	1	6921
Mid-SES classrooms	0.33	0.47	0.0	1	8478	0.33	0.47	0.0	1	7872	0.33	0.47	0.0	1	6864
High-SES classrooms	0.33	0.47	0.0	1	8478	0.33	0.47	0.0	1	7872	0.33	0.47	0.0	1	6876
Fraction of grade-retained students in classroom	0.03	0.05	0.0	1		0.07	0.08	0.0	1		0.20	0.18	0.0	1	
Students who missed the test	10.92	5.12	0.0	48		11.37	5.24	0.0	52		14.53	8.48	0.0	60	
Social tracking between classrooms	0.28	0.15	0.0	2		0.27	0.14	0.0	1		0.28	0.11	0.0	1	
Fraction of first-generation immigrants in classroom	0.05	0.07	0.0	1		0.06	0.07	0.0	1		0.05	0.07	0.0	1	
Fraction of second-generation immigrants in classroom	0.05	0.09	0.0	1		0.05	0.08	0.0	1		0.07	0.08	0.0	1	
Fraction of female students in classroom	0.50	0.11	0.0	1		0.49	0.11	0.0	1		0.50	0.29	0.0	1	
Classroom size	17.76	3.91	10.0	32		19.99	3.97	10.0	34		19.12	4.56	10.0	35	
School size	5.11	2.08	1.0	14		6.13	3.06	1.0	17		7.97	3.27	1.0	20	
Teacher absenteeism (province)	7.22	1.63	4.1	13		7.28	1.67	4.1	13		7.29	1.67	4.1	13	
Illicit economic activity - Evasion of public TV fees (province)	72.67	10.71	47.9	93		72.65	10.85	47.9	93		72.36	10.75	47.9	93	
Illicit economic activity - Estimated risk of tax evasion (province)	96.50	25.05	48.0	141		95.54	24.84	48.0	141		95.77	24.79	48.0	141	
Illicit economic activity - Estimated rate of undeclared work (province)	21.94	8.82	8.0	44		22.36	8.91	8.0	44		22.64	9.03	8.0	44	

Table 2. Descriptive statistics (continued)

	Primary education					Lower secondary education					Upper secondary education				
	mean	sd	min	max	freq	mean	sd	min	max	freq	mean	sd	min	max	freq
Unemployment rate	8.44	3.99	2.1	19		8.54	4.01	2.1	19		8.65	4.00	2.1	19	
Province with metropolitan area	0.39	0.49	0.0	1		0.36	0.48	0.0	1		0.36	0.48	0.0	1	
Index of teacher precariousness	-0.20	1.09	-4.9	1		-0.22	1.10	-4.9	1		-0.27	1.14	-4.9	1	
Share of early school leavers	0.20	0.07	0.1	0		0.20	0.07	0.1	0		0.20	0.07	0.1	0	
Share of teacher turnover	0.10	0.02	0.1	0		0.10	0.02	0.1	0		0.10	0.02	0.1	0	
Share of adult population participating in education	0.06	0.02	0.0	0		0.06	0.02	0.0	0		0.06	0.01	0.0	0	
Density of province	550.7	717.3	39.0	2630		512.1	691.4	39.0	2630		521.8	705.0	39.0	2630	
Share of adult population with a low level of education	0.49	0.08	0.3	1		0.49	0.08	0.3	1		0.49	0.07	0.3	1	
North West	0.26	0.44	0.0	1	6543	0.23	0.42	0.0	1	5452	0.24	0.43	0.0	1	4998
North East	0.18	0.39	0.0	1	4685	0.19	0.39	0.0	1	4406	0.18	0.38	0.0	1	3619
Centre	0.18	0.39	0.0	1	4606	0.19	0.39	0.0	1	4377	0.17	0.37	0.0	1	3469
South	0.26	0.44	0.0	1	6689	0.28	0.45	0.0	1	6519	0.30	0.46	0.0	1	6104
Islands	0.11	0.32	0.0	1	2911	0.12	0.33	0.0	1	2862	0.12	0.32	0.0	1	2471
Observations (classrooms)	25434					23616					20661				
Schools	6972					5435					3733				
Provinces	103					103					103				
Regions	20					20					20				
Macro-areas	5					5					5				

Source: Invalsi dataset 2011-12

Econometric strategy: determinants, incentives, and deterrents of cheating

In this paper we estimated logistic regression models. We wanted to explore to what extent the explanatory factors included in the equation make the likelihood of cheating increase or decrease. The baseline equation can be expressed as follow:

$$\begin{aligned}\text{logit}\{\Pr(cheat_{csp} = 1 \,|\mathbf{X}_{csp})\} \\ = \alpha + \gamma_1 O_{csp} + \gamma_2 H_{csp} + \gamma_3 F_{csp} + \gamma_4 E_{sp} + \gamma_5 T_{sp} + \gamma_6 A_p + \gamma_7 D_p \\ + \gamma_8 C_{csp} + \gamma_9 P_p + \gamma_{10}(O_{csp} * T_{sp}) + \varepsilon_{csp}\end{aligned} \quad (3)$$

where $cheat_{csp}$ is whether a classroom c in school s in province p is suspected of cheating. O_{csp} covers two dummies that refer to the opportunity to cheat, that is, whether a classroom c was monitored by an external controller, and whether it was indirectly controlled. As regards to the factors related to the network-based civic capital, H_{csp} is the measure of classroom social homogeneity, while F_{csp} is a vector that refers to the factors that would indicate the compensation function of teacher cheating: the classroom portion of grade-retained students and a dummy for whether a classroom has a low SES average.

As to the culture-based civic capital, we included three factors associated with the school culture: the school-level measure E_{sp}, which represents the fraction of students who missed the test (as a proxy of the percentage of students excluded from the test taking); the school-level measure T_{sp}, which is a variable denoting the extent of school practices that result in social tracking between classrooms; and the province-level measure A_p, which is the share of teacher absenteeism. Here we were also interested in analysing the extent to which the mentioned external deterrents of cheating, during the test taking, are effective in schools with high levels of non-civic-minded culture, or lack of civic capital. For this we included the vector $O_{csp} * T_{sp}$, which represents the interaction term between a continuous measure of lack of civic capital at the school level (social tracking between classrooms within schools) and the binary predictors of both direct and indirect external control during the test.

When talking about social dissimilarity between classrooms, we should take into account that, in Italy, especially in urban centres, many schools occupy several buildings which are often physically separated – possibly by more than 1 km. This may lead to social homogeneity within, and social heterogeneity between, classrooms, not as a result of practices of student tracking, but as a reflection of the social composition of the school's surroundings. Here we assume that a school is more likely to be spread across different buildings when there are more classrooms. In order to account for this phenomenon, and bearing in mind that no information is available on the urban or rural location of schools, we included the measure Z_{sp} of school size, which reflects the number of classrooms in each school. Finally, we also included, in separate specifications, three factors associated

with the deviant culture of the school context, all of them representing illicit economic activities alongside the official economy (D_p): evasion of public television fees, a measure of the risk of tax evasion, and the estimated rate of undeclared work.

Here we deal with potential problems of omitted variables, since behaviours denoting law acceptance or conformity (e.g. tax compliance vs. tax evasion) may be effectively driven by factors not necessarily related to the civic capital, such as economic payoffs or legal enforcement. In fact, as pointed by Guiso et al. (2010), this is the reason why outcome-based measures such as behaviours, as indirect measures of civic capital, are difficult to interpret. We assume, in fact, that measures of deviant behaviours may only partially inform on the lack of civic-minded capital. However, we still argue that, to the extent that the following two assumptions are accepted, they may still be important predictors. Firstly, when dealing with school and province-level aggregated factors, such as social tracking of students, tax evasion or undeclared work, we assume that the legal framework and the existing measures of law enforcement (at the national and the education system level) are held constant. Since we exploit within-country variability of a single country, we partially account for the driving force of legal deterrents – measures of legal enforcement, recommendations, deontological professional standards, etc. Secondly, we assume that much of the motivations and incentives to act illicitly, associated with the social structure and territorial characteristics, can be captured through the socioeconomic composition controls (at the school and the province level), as well as by the territorial fixed effects (at the regional and the macro-area level).

The variables used as controls measures in the equations were at classroom, school and province level. At the classroom level, we included the vector C_{csp}, which covers aggregated student background data: the socio-demographic factors considered were the fraction of female students and the fraction of first-generation and second-generation immigrant students in the classroom. We also controlled for the classroom size S_{csp}, expressed as the number of students enrolled. At the province level, we accounted for a range of social, economic and geographic characteristics, which was included in the vector P_p of province characteristics data. The variables considered in this vector were: GDP per capita; unemployment rate; whether provinces contained a metropolitan area; population size and density of province; share of adult population participating in education, and share of adult population with a low level of education. In this vector we also included province-aggregated data related to the school system, such as the share of early school leavers, an index of teacher precariousness (teachers with a temporary contract), and the share of teacher turnover. Finally, in order to control for systemic and cultural differences at the territorial level, we separately introduced a set of macro-area and region fixed-effects. ε_{csp} is the error term. Standard errors were clustered at the province level.

Another model was run for the differential effect on cheating depending on the high or low stakes nature of the test. Here we take advantage of the fact that the results of the tests

taken at the end of the 8th grade have formal consequences for students, since they contribute to the final mark when leaving lower secondary school. The equation is as follows:

$$\begin{aligned}\text{logit}\{\Pr(cheat_{cspg} = 1 \,|\mathbf{X}_{cspg})\} \\ = \alpha + \gamma_1 S_g + \gamma_2 O_{cspg} + \gamma_3 H_{cspg} + \gamma_4 F_{cspg} + \gamma_5 E_{cspg} + \gamma_6 T_{spg} \\ + \gamma_7 A_{pg} + \gamma_8 D_{pg} + \gamma_9 C_{cspg} + \gamma_{10} P_{pg} + \varepsilon_{cspg}\end{aligned} \quad (4)$$

where $cheat_{cspg}$ is whether a classroom *c* in school *s* in province *p* in grade *g* is suspected of cheating. In this equation we included the vector S_g, which covers dummies indicating whether the test was in the 5th year of primary education, or in the 1st or 3rd year of lower secondary education. These dummies should inform on the differential effects on cheating of taking the test in different grades, net from different factors at classroom, school and province level.

Nevertheless, since we used the grade fixed effect as a proxy of the high-stakes nature of the test (tests in the third grade of middle school have formal consequences for students), we have to be aware of a potential problem of omitted variables: the phenomenon of cheating may be, for instance, associated with the teacher profile, which in turn may vary significantly depending on the grade. Therefore, to partially account for teacher characteristics, and despite the limited availability of school-level data, we included grade-variant control factors, which could drive the results in terms of cheating: provincial-level factors related to teacher characteristics (teacher absenteeism, turnover rate, index of teacher precariousness), and school-level predictors of deviant behaviour (students excluded from the test taking, school social tracking, and school and classroom size). The remaining control variables were as in equation 3, except for data on student socioeconomic background, which was not available for students in the 3rd year of lower secondary.

After reporting the coefficients, the results were interpreted as *average marginal effects*, but also as *odds ratio*. The *average marginal effects* are calculated estimating the average of the classroom marginal effects, and expressed as a percentage, indicating how an increase in *x* is associated with an increase or decrease of the probability of *y* being equal to 1 (classroom suspected of cheating). While for dummy variables the marginal effect is expressed in comparison to the base, for continuous variables it is expressed for one-unit change in the explanatory factor. On the other hand, the *odds ratio* measures the probability of *y* being equal to 1 (classroom suspected of cheating), relative to the probability of *y* being equal to 0 (classroom not suspected of cheating). *p* is the probability that *y*=1 divided by the probability that *y*=0 (1-*p*).

The goodness of fit was measured with the percentage of values correctly predicted. First of all, taking the estimated coefficient $\hat{\beta}$, we calculate the predicted probability $\hat{p}$ that *y* would be equal to 1 (being cheater) for each classroom in the dataset:

$$\hat{p} = pr[y = 1|\mathbf{x}] = F(\mathbf{x}'\hat{\beta}) \quad (5)$$

In logit and probit models, the predicted probabilities are limited between 0 and 1, and indicate the likelihood that *y*=1. A predicted probability of being suspected of cheating is 0.2 means that a classroom is 20% likely to be cheater. We can say that this classroom is not likely to cheat (*y*=0), since predicted probability is lower than 0.5. If the value is higher than 0.5, we can predict that a classroom is likely to cheat (*y*=1). Once we have $\hat{p}$, to check the goodness of the model, we calculate the percentage of values correctly predicted. This is the proportion of true predictions to total predictions ($\hat{y}$ = y). We can say that our model is a good fit if we correctly give at least 70% of true or correct predictions. The models derived from equations 3 and 4 correctly predict between 88% (90th percentile cut-off) and 94% (95th percentile cut-off) of values, and therefore we can confirm that they are suitable to use for the empirical analysis.

MAIN RESULTS

In this section we begin with the general results, dealing with the territorial divide and the incentives to cheat – the high-stakes nature of tests and the lack of control during the test taking. Then we present the main findings, with the results related to the factors of social capital understood as network-based civic capital, and those in which cheating is associated with the social capital understood as culture-related civic capital.

General results

Territorial divide

Considering the territorial divide in Italy in terms of teacher cheating, our results confirm what other studies on school cheating have already highlighted: the geographical location of schools is one of the factors most correlated with teacher cheating. Cheating behaviours are much greater in the southern regions and the islands, particularly in the regions of Calabria and Sicily. In Table 3 we present the results, for primary and lower secondary school, of logistic regression models used to estimate the probability of being suspected of cheating, using the 90th percentile threshold calculated at national level.

The table clearly shows there is a non-random pattern in the geographical distribution of teacher cheating across macro-areas. According to the odds ratio, net of different multi-level factors, to be a cheater for a teacher located in the south and islands is between two and four times more likely than in the north-east of the country. As can be seen in Figure 1, this territorial pattern is even more noticeable at the regional level. Compared to the

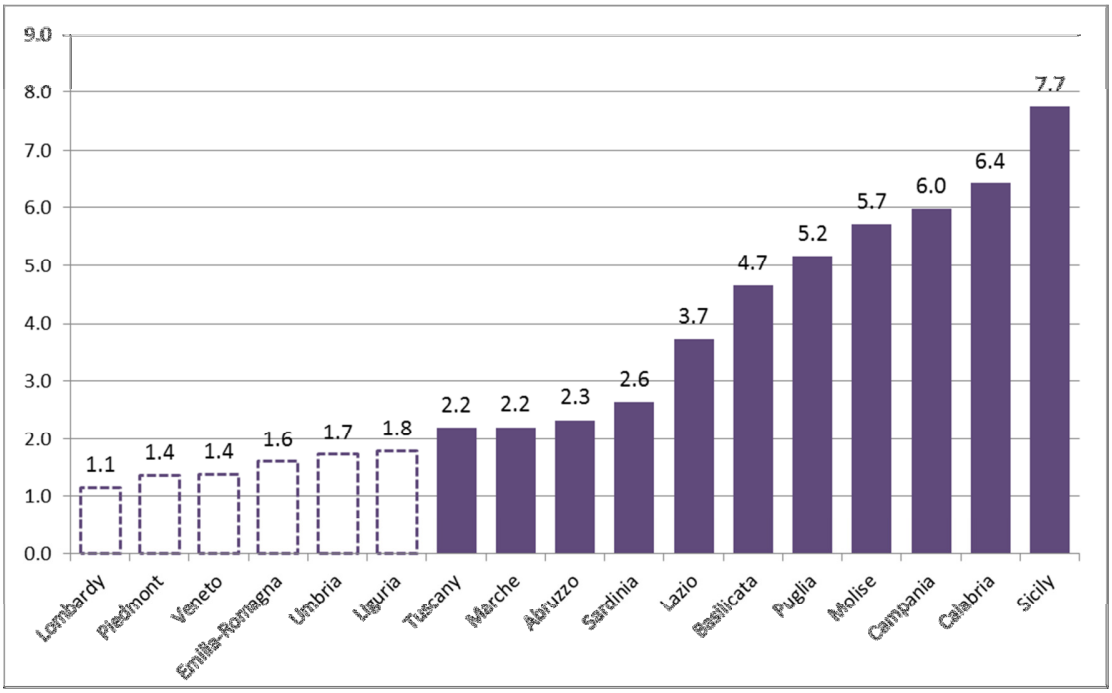
northern region of Friuli-Venezia Giulia, the region with the lowest levels of cheating, the probability of cheating of teachers located in southern regions, such as Molise, Campania, Calabria and Sicily, accounting for a wide range of classroom and contextual control factors, is between six and seven times higher. In central and southern regions, such as Lazio, Basilicata or Puglia, teacher cheating is between four and five times more likely, while there is no statistically significant difference with respect to the central region of Umbria and the other northern regions of Piedmont, Lombardy, Veneto, Emilia-Romagna, Umbria, and Liguria.

Table 3. Territorial divide: probability of teacher cheating by macro-areas (odds-ratio)

Suspected classroom of cheating (90^{th} cut-off)		Maths			Italian language		
		Primary Education	Lower Secondary	Upper Secondary	Primary Education	Lower Secondary	Upper Secondary
		Odds Ratio	Odds Ratio	Odds Ratio	Odds Ratio	Odds Ratio	Odds Ratio
Macro-area fixed effects (ref. North East)	North West	1.293*** (0.10)	1.287*** (0.11)	1.155*** (0.03)	1.222 (0.16)	0.843 (0.10)	1.374** (0.18)
	Centre	2.602*** (0.17)	1.927*** (0.20)	1.710*** (0.09)	2.044*** (0.28)	1.733*** (0.20)	2.148*** (0.29)
	South	4.820*** (1.51)	2.800*** (1.03)	3.707*** (0.82)	2.988*** (0.60)	2.252*** (0.41)	4.123*** (0.82)
	Islands	4.786*** (1.19)	3.598*** (1.23)	3.045*** (0.46)	3.352*** (0.67)	2.725*** (0.50)	4.088*** (0.81)
Observations		23772	22647	19743	23518	22619	19534
Pseudo R^2		0.2008	0.1432	0.1316	0.1663	0.1008	0.1303

Source: Invalsi dataset 2011-12. Sig. * p < 0.10, ** p < 0.05, *** p < 0.01. The dependent variable is the cheating indicator using the 90th percentile cut-off. The macro-area of reference is the North-East. Odds-ratio and average marginal effects have been calculated. In columns 4 to 6, the control variables include classroom, school and province factors. Classroom and school: classroom SES average, fraction of female students, the fraction of first-generation and second-generation immigrant students in the classroom, classroom size, and school size. Province: GDP per capita, unemployment rate, whether provinces contained a metropolitan area, population size and density of province, share of adult population participating in education, share of adult population with a low level of education, share of early school leavers, an index of precarious teachers, the share of teacher turnover, and the share of teacher absenteeism.

Figure 1. Territorial divide: probability of teacher cheating by region (odds-ratio)



The dependent variable is the cheating indicator using the 90th percentile cut-off. The region of reference is Friuli-Venezia Giulia. The Odds Ratio are the exponentiated regression coefficients. Columns with dashed lines mean that there is no statistically significant difference. Control variables include classroom, school and province factors. Classroom and school: classroom SES average, fraction of female students, the fraction of first-generation and second-generation immigrant students in the classroom, classroom size, and school size. Province: GDP per capita, unemployment rate, whether provinces contained a metropolitan area, population size and density of province, share of adult population participating in education, share of adult population with a low level of education, share of early school leavers, an index of precarious teachers, the share of teacher turnover, and the share of teacher absenteeism.

High-stakes testing for schools

As regards the high or low stakes nature of the test, we compared teacher cheating levels in the standardised tests of the 3rd year of middle school, whose results are high-stakes for students, with the cheating in the 5th year of primary and the 1st year of the lower secondary school. A first insight can be seen in Figure 2, which shows the percentage distribution of cheating, using both the 90th and the 95th percentile thresholds. We found a noticeable gap between teacher cheating levels, with the magnitude of teacher cheating in the 3rd year of lower secondary being twice that detected in the 5th year of primary and the 1st year of lower secondary. Moreover, teacher cheating in this 3rd year was three times greater than the cheating identified in primary education when considering a more restricted measure of cheating using the 95th percentile cut-off.

This is further confirmed when modelling the probability that a classroom would be cheater, depending on the grade when the tests were taken (Table 4). As hypothesised, teachers of the 3rd grade of middle school are significantly the most prone to cheat.

Although the tests of the 8th grade are not high-stakes for teachers, and taking into consideration the potential problem of omitted variables, the results allow us to confirm that teachers are notably more likely to cheat when a test has formal consequences for students. When considering the 90th percentile cut-off, odds-ratios indicate that the teachers in the 3rd grade are, in fact, twice as likely to cheat as those of primary and first year of secondary school. This likelihood increases when considering the 95th percentile threshold, a more restrictive measure of cheating. In this case, the likelihood of a teacher of the third year of lower secondary cheating is even higher: 3 times more likely with respect to primary education teachers, and 2.5 times compared to teachers of the 1st year of lower secondary school. Looking in terms of average marginal effects, Table 4 shows that teachers in the 5th year of primary education are between 5% and 6% less likely to cheat in comparison with teachers of the third year of lower secondary, while teachers in the first year of lower school are between 4.5% and 6.5% less likely.

Figure 2. Low vs. high stakes tests: percentage of teacher cheating across grades (maths)

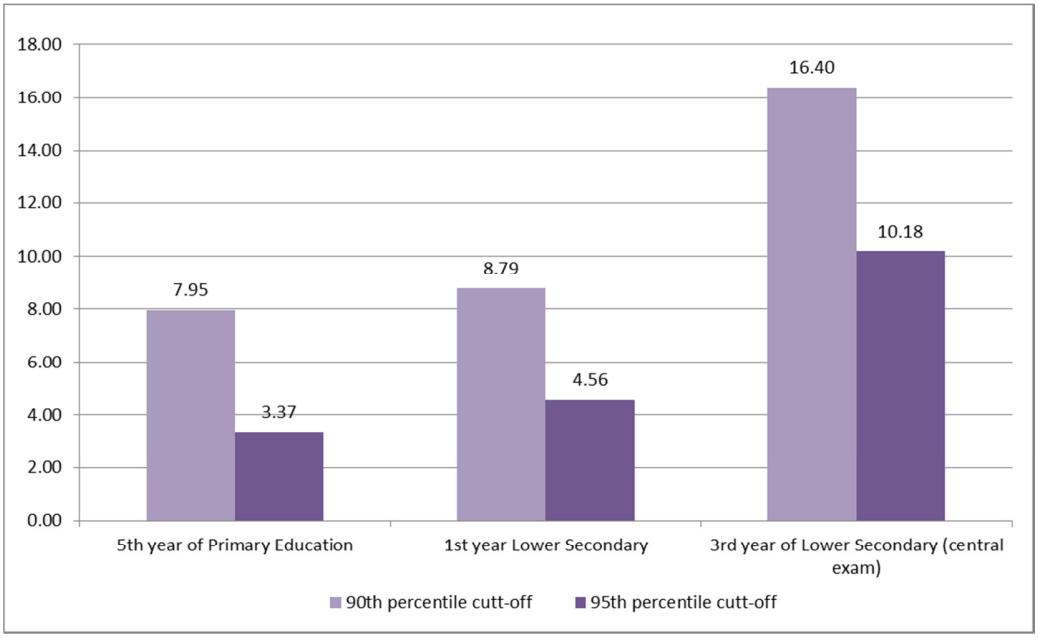


Table 4. Low vs. high stakes tests: probability of teacher cheating across grades

Suspected classroom of cheating		Coefficients		Odds ratio		Average marginal effects	
		90th percentile cut-off	95th percentile cut-off	90th percentile cut-off	95th percentile cut-off	90th percentile cut-off	95th percentile cut-off
Grade (ref. 3rd year of Lower secondary – central exam)	5th year of Primary education	-0.703*** (0.04)	-1.038*** (0.05)	0.495*** (0.02)	0.354*** (0.02)	-0.060*** (.003)	-0.053*** (.003)
	1st year Lower Secondary	-0.758*** (0.03)	-0.886*** (0.04)	0.468*** (0.01)	0.412*** (0.02)	-0.065*** (.002)	-0.045*** (.002)
Observations		75520	75520	75520	75520	75520	75520
Pseudo R^2		0.1806	0.1866	0.1806	0.1866		

Source: Invalsi dataset 2009-10. Robust standard errors clustered by school in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For average marginal effects we show the Delta-method standard errors. The dependent variable is the cheating indicator in the math tests using the 90th and the 95th percentile cut-offs. The Odds Ratio are the exponentiated regression coefficients. Control variables include classroom, school and province factors. Classroom and school: fraction of female students, the fraction of first-generation and second-generation immigrant students in the classroom, classroom size, and school size. Province: GDP per capita, unemployment rate, whether provinces contained a metropolitan area, population size and density of province, share of adult population participating in education, share of adult population with a low level of education, share of early school leavers, an index of precarious teachers, the share of teacher turnover, and the share of teacher absenteeism. It includes macro-area fixed effects. Region fixed effects were included in separate specifications (tables 8a and 8b, in appendix 2), but the sign and significance of the results was not affected.

External control and opportunity

This research confirms published results showing that external control of the classroom during the test is the most important deterrent to prevent illicit behaviours (Bertoni et al., 2013; Ferrer-Esteban, 2013; Paccagnella & Sestito, 2014). Having a controller during the test taking significantly reduces the likelihood of teacher cheating. Moreover, it confirms the spill-over effects of the presence of external controllers in other classrooms in the same school. As can be seen in table 5, if we consider the odds ratio, we can confirm that teachers in a non-monitored classroom are two to three times more likely to cheat than teachers in directly monitored classrooms. The average marginal effect obviously follows the same pattern: teachers with direct control during the test taking are between 6% (lower secondary) and 10% (primary) less likely to cheat, with respect to teachers without external control. Differences are more attenuated when considering indirect monitoring. The marginal effect indicates, in this case, that a teacher in a non-monitored classroom is between 1.4% and 2.7% more likely to fall into dishonest behaviours. Not allowing the opportunity to cheat is clearly one of the most powerful deterrents against cheating behaviour. The role of external control is further analysed in the third section of the results, where interaction effects with culture-based factors of civic capital were tested.

Table 5. Opportunity to cheat: external control and probability of teacher cheating (Italian language)

	Coefficients			Odds-ratio			Average marginal effects		
Suspected classroom of cheating (90th cut-off)	Primary	Lower Sec.	Upper Sec.	Primary	Lower Sec.	Upper Sec.	Primary	Lower Sec.	Upper Sec.
Controllers in classroom during the test taking (ref: no control)	-1.194*** (0.15)	-0.636*** (0.09)	-0.998*** (0.13)	0.303*** (0.05)	0.530*** (0.05)	0.369*** (0.05)	-0.100*** (0.013)	-0.058*** (0.008)	-0.089*** (0.011)
Controllers in school during the test taking (ref: no control)	-0.328*** (0.08)	-0.154** (0.07)	-0.275*** (0.07)	0.720*** (0.06)	0.858** (0.06)	0.760*** (0.05)	-0.027*** (0.006)	-0.014** (0.006)	-0.024*** (0.006)
Observations	23518	22619	19534	23518	22619	19534	23518	22619	19534
Pseudo R^2	0.1664	0.1010	0.1314	0.1664	0.1010	0.1314			

Source: Invalsi dataset 2011-12. Robust standard errors clustered by school in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The dependent variable is the cheating indicator using the 90th percentile cut-off. Odds-ratio and average marginal effects have been calculated. The control variables include classroom, school and province factors. Classroom and school: classroom SES average, fraction of female students, the fraction of first-generation and second-generation immigrant students in the classroom, classroom size, and school size. Province: GDP per capita, unemployment rate, whether provinces contained a metropolitan area, population size and density of province, share of adult population participating in education, share of adult population with a low level of education, share of early school leavers, an index of precarious teachers, the share of teacher turnover, and the share of teacher absenteeism. It includes macro-area fixed effects. Region fixed effects were included in separate specifications (tables 8a and 8b, in appendix 2), but the sign and significance of the results was not affected.

Social capital as network-based civic capital: bonding social capital and cheating

Studies have suggested that the higher levels of cheating in Southern Italy are due to both a lower endowment of bridging social capital, and a higher degree of bonding social capital (Bertoni et al., 2013). Other studies have demonstrated that cheating is positively associated with measures of *particularistic values* (Paccagnella & Sestito, 2014), which, in Putnam's view, can be understood as civic values related to network-based *bonding* social capital. For instance, Paccagnella and Sestito (2014) found strong associations between cheating and contexts where people use close local networks to a greater extent, to be informed, or to participate in associations with similar people. In this section we focus particularly on the framework of *network-based civic capital* to explain and understand teacher cheating. As indicated in the research questions, the objective is twofold: firstly, we want to explore to what extent cheating behaviours are more likely in classrooms where there are higher stocks of bonding social capital, expressed as a measure of classroom social homogeneity. Secondly, we analyse the rationale behind teacher cheating from the perspective of the social capital understood as community-based civic engagement: moral reasoning in close

and socially homogeneous groups may be targeted to support the most disadvantaged members, and therefore be socially justified, even if it does not respond to deontological professional standards and contravenes a more society-oriented civic engagement.

Conditions for cheating: classroom social homogeneity

As mentioned earlier, to the extent that a context is socially more homogeneous, stronger ties between teachers and students are more plausible. The first indicator is a measure of social dispersion within classrooms (so classroom social heterogeneity), which, inverted, becomes a proxy of strong ties established in a close and socially homogeneous context. Since we refer specifically to teacher cheating, we do not approach classroom social homogeneity to identify interactions among students as a determinant of cheating (Lucifora & Tonello, 2012), but in contexts where ties of mutual support between teacher and students are more likely to be found. As can be seen in tables 6a and 6b, the social dispersion indicator is a robust predictor of teacher cheating, net of classroom, school and territorial controls. Specifically, looking at the marginal effects, we observed that a one-unit increase in the standard deviation of the students' socioeconomic index within classrooms decreases the probability of a teacher being suspected of cheating. For a one-unit increase in social heterogeneity we expect the probability of cheating to be reduced by 1.6% in primary and 3% in lower and upper secondary. As regards the odds ratio, if we invert the scale, we can say that for a one-unit decrease in the measure of social dispersion, the probability for being suspected of cheating is 1.2 and 1.4 times as likely as not being suspected, in primary, and lower and upper secondary, respectively. Overall, teachers are more prone to support students, suggest answers or fill in the answer sheets when the classroom composition is socially more homogeneous, where network-based forms of bonding social capital are more likely to increase.

Rationale for teacher cheating: a compensation function for disadvantaged students

From the bonding approach of social capital, cheating can be understood as a strategy targeted to support socially and academically disadvantaged students. If we look at the fraction of grade-retained students in classrooms, we observe that teacher cheating is significantly associated with these academically disadvantaged students in the upper secondary for maths and Italian, and in the lower secondary only for maths. Indeed, a one-percent unit increase of grade-retained students in upper secondary school increases the probability of a teacher cheating by 12% in Italian and 8% in maths (Tables 6a and 6b). In the case of lower secondary teachers, the marginal effect increases to 8% for a one-percent unit increase in retained students. This is, however, an outcome that may be influenced by the differential presence of retained students depending on the education grades. The extent to which we move up through the grades, the percentage of grade-retained students

increases. It is almost non-existent in primary, representing only 3.4% of students. In the lower secondary this is about 7%, while in the upper secondary it dramatically rises to almost 20%.

In addition, to test whether cheating is more likely in classrooms with socially disadvantaged students, we estimated the probability that being suspected of cheating was dependant on the socioeconomic average status of classrooms. Here we compared classrooms with a high SES average with low and mid-SES classrooms. The results are significant and robust in all specifications: teachers in classrooms with low social composition are more prone to be suspected of cheating. Specifically, in primary and upper secondary, having a low SES composition doubles the probability of cheating in both Italian and maths (Tables 6a and 6b). In terms of average marginal effects, teachers in low SES classroom are 5% to 7% more likely to cheat than teachers in high SES classrooms. These results are attenuated, though still statistically significant, in the 1st year of the lower secondary school, and between mid SES and high SES classrooms.

Table 6a. Teacher cheating and *bonding* social capital related factors (Italian language)

	Coefficients			Odds-ratio			Average marginal effects		
Suspected classroom of cheating (90th cut-off)	Primary	Lower Sec.	Upper Sec.	Primary	Lower Sec.	Upper Sec.	Primary	Lower Sec.	Upper Sec.
Social student heterogeneity in classroom	-0.190* (0.10)	-0.333*** (0.11)	-0.351** (0.14)	0.827* (0.08)	0.717*** (0.08)	0.704** (0.10)	-0.016** (0.01)	-0.030*** (0.01)	-0.031** (0.01)
Mid-SES classroom (ref: Low-SES classrooms)	-0.381*** (0.06)	-0.107* (0.06)	-0.201*** (0.07)	0.683*** (0.04)	0.898* (0.05)	0.818*** (0.05)	-0.032*** (0.00)	-0.010* (0.01)	-0.018*** (0.01)
High-SES classroom (ref: Low-SES classrooms)	-0.656*** (0.08)	-0.134** (0.06)	-0.813*** (0.10)	0.519*** (0.04)	0.874** (0.06)	0.444*** (0.04)	-0.055*** (0.01)	-0.012** (0.01)	-0.072*** (0.01)
Fraction of grade-retained students in classroom	0.664 (0.48)	0.389 (0.30)	1.317*** (0.17)	1.943 (0.93)	1.475 (0.45)	3.733*** (0.62)	0.056 (0.04)	0.035 (0.03)	0.117*** (0.01)
Observations	23518	22619	19534	23518	22619	19534	23518	22619	19534
Pseudo R^2	0.1664	0.1010	0.1314	0.1664	0.1010	0.1314			

Table 6b. Teacher cheating and *bonding* social capital related factors (maths)

	Coefficients			Odds-ratio			Average marginal effects		
Suspected classroom of cheating (90^{th} cut-off)	Primary	Lower Sec.	Upper Sec.	Primary	Lower Sec.	Upper Sec.	Primary	Lower Sec.	Upper Sec.
Social student heterogeneity in classroom	-0.352*** (0.10)	-0.323*** (0.11)	-0.031 (0.17)	0.703*** (0.07)	0.724*** (0.08)	0.970 (0.16)	-0.028*** (0.01)	-0.028*** (0.01)	-0.003 (0.01)
Mid-SES classroom (ref: Low-SES classrooms)	-0.300*** (0.05)	-0.139*** (0.05)	-0.285*** (0.06)	0.741*** (0.04)	0.870*** (0.05)	0.752*** (0.04)	-0.024*** (0.00)	-0.012** (0.00)	-0.025*** (0.01)
High-SES classroom (ref: Low-SES classrooms)	-0.571*** (0.08)	-0.222*** (0.08)	-0.538*** (0.09)	0.565*** (0.04)	0.801*** (0.06)	0.584*** (0.05)	-0.046*** (0.01)	-0.019*** (0.01)	-0.048*** (0.01)
Fraction of grade-retained students in classroom	0.469 (0.52)	0.915** (0.40)	0.934*** (0.25)	1.598 (0.83)	2.497** (0.99)	2.545*** (0.64)	0.038 (0.04)	0.080** (0.03)	0.083*** (0.02)
Observations	23772	22647	19743	23772	22647	19743	23772	22647	19743
Pseudo R2	0.2015	0.1439	0.1327	0.2015	0.1439	0.1327			

Source: Invalsi dataset 2011-12. Robust standard errors clustered by school in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The dependent variable is the cheating indicator using the 90th percentile cut-off. The control variables include classroom, school and province factors. Classroom and school: classroom SES average, fraction of female students, the fraction of first-generation and second-generation immigrant students in the classroom, classroom size, and school size. Province: GDP per capita, unemployment rate, whether provinces contained a metropolitan area, population size and density of province, share of adult population participating in education, share of adult population with a low level of education, share of early school leavers, an index of precarious teachers, the share of teacher turnover, and the share of teacher absenteeism. It includes macro-area fixed effects. Region fixed effects were included in separate specifications (tables 8a and 8b, in appendix 2), but the sign and significance of the results was not affected.

Social capital as culture-based civic capital: cheating as deviant non-civic-minded behaviour

In this section we tested the hypothesis that teachers in contexts with lower levels of civic capital are more prone to cheat in standardised tests. As mentioned earlier, here we specifically explore to what extent cheating behaviours are associated with *culture-based deviant behaviours*, which are collectively shared and detrimental to the general public good. Beside deontological and ethical considerations, we consider them as non-civic-minded behaviours since they do not comply with legal requirements, or are against laws and regulations. Deviant behaviours are used to proxy the lack of civic social capital at two levels: schools and provinces.

School-related factors

Firstly, we report the results on the association between teacher cheating behaviours and deviant practices undertaken in a noticeable number of schools: social tracking of students between classrooms, and the exclusion of students from tests. These school system-related factors were found to have a strong predictive power, as they were significantly associated with the probability of teacher cheating. First, as shown in tables 7a and 7b, we observed that those schools more likely to have teachers who cheat also seem to undertake other dishonest strategies to raise the average performance. With primary education the only exception, schools that directly alter the test results also tend to exclude a higher proportion of students from test taking (expressed as the proportion of students who missed the test).

The second factor associated with teacher cheating is the practice of non-random allocation of students across classrooms, on the basis of their social background. In short, schools that undertake tracking practices, which result in social polarisation of classrooms, are also more likely to show higher levels of teacher cheating. As can be seen in tables 7a and 7b, a one-unit increase in the standard deviation of a classroom's socioeconomic index within schools increases the probability of the classroom being suspected of cheating by between 3% and 5% in Italian, and between 2% and 5% in maths. As regards the odds-ratio, for a one-unit increase in the measure of social tracking, the probability of being suspected of cheating is, on average across grades, 1.6 times as likely as not being suspected. It is worth noting that, in the case of upper secondary education, no significant association was found in maths. This is probably due to the fact that upper secondary schools are already highly tracked, having a more homogeneous social composition. On the contrary, the most solid association between cheating and social tracking, as a proxy of non-minded-civic behaviour, is in the first year of the lower secondary school: this is when students from the primary school are distributed according to the allocation criteria set by the school board.

Table 7a. Teacher cheating and schools' lack of civic capital (Italian language)

	Coefficients			Odds-ratio			Average marginal effects		
Suspected classroom of cheating (90th cut-off)	Primary	Lower Sec.	Upper Sec.	Primary	Lower Sec.	Upper Sec.	Primary	Lower Sec.	Upper Sec.
Students who missed the test	0.005 (0.01)	0.013*** (0.00)	0.018*** (0.01)	1.005 (0.01)	1.013*** (0.00)	1.018*** (0.01)	0.000 (0.00)	0.001*** (0.00)	0.002*** (0.00)
Social tracking between classrooms	0.329* (0.17)	0.541*** (0.18)	0.465** (0.20)	1.390* (0.24)	1.718*** (0.30)	1.591** (0.31)	0.028* (0.01)	0.049*** (0.16)	0.041** (0.02)
Teacher absenteeism (province)	0.217*** (0.03)	0.150*** (0.03)	0.161*** (0.03)	1.242*** (0.04)	1.162*** (0.04)	1.175*** (0.03)	0.018*** (0.00)	0.014*** (0.00)	0.014*** (0.00)
Observations	23518	22619	19534	23518	22619	19534	23518	22619	19534
Pseudo R^2	0.1664	0.1010	0.1314	0.1664	0.1010	0.1314			

Table 7b. Teacher cheating and schools' lack of civic capital (maths)

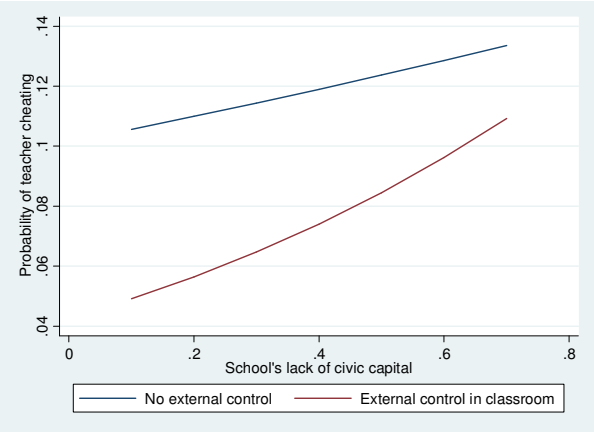
	Coefficients			Odds-ratio			Average marginal effects		
Suspected classroom of cheating (90th cut-off)	Primary	Lower Sec.	Upper Sec.	Primary	Lower Sec.	Upper Sec.	Primary	Lower Sec.	Upper Sec.
Students who missed the test	0.017*** (0.01)	0.028*** (0.00)	0.014*** (0.00)	1.017*** (0.01)	1.028*** (0.01)	1.014*** (0.00)	0.001*** (0.00)	0.002*** (0.00)	0.001*** (0.00)
Social tracking between classrooms	0.364* (0.20)	0.578*** (0.17)	0.207 (0.19)	1.439* (0.29)	1.782*** (0.31)	1.230 (0.24)	0.029* (0.02)	0.051*** (0.02)	0.018 (0.02)
Teacher absenteeism (province)	0.210*** (0.05)	0.170*** (0.04)	0.145*** (0.04)	1.234*** (0.06)	1.185*** (0.05)	1.156*** (0.04)	0.017*** (0.00)	0.015*** (0.00)	0.013*** (0.00)
Observations	23772	22647	19743	23772	22647	19743	23772	22647	19743
Pseudo R2	0.2015	0.1439	0.1327	0.2015	0.1439	0.1327			

Source: Invalsi dataset 2011-12. Robust standard errors clustered by school in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The dependent variable is the cheating indicator using the 90th percentile cut-off. The control variables include classroom, school and province factors. Classroom and school: classroom SES average, fraction of female students, the fraction of first-generation and second-generation immigrant students in the classroom, classroom size, and school size. Province: GDP per capita, unemployment rate, whether provinces contained a metropolitan area, population size and density of province, share of adult population participating in education, share of adult population with a low level of education, share of early school leavers, an index of precarious teachers, the share of teacher turnover, and the share of teacher absenteeism. It includes macro-area fixed effects. Region fixed effects were included in separate specifications (tables 8a and 8b, in appendix 2), but the sign and significance of the results was not affected.

Early in this paper we showed that the presence of external controllers during the test taking is a powerful deterrent for teacher cheating. Now we showed that practices of social tracking within schools, understood as a proxy of lack of civic capital, have a robust association with teacher cheating. At this point, we questioned to what extent deterrents of cheating are effective when considering the level of non-civic-minded engagement of schools. Consequently, we introduced interaction terms to test the change in the interaction effects of direct and indirect measures of external control on teacher cheating, depending upon the degree of civic capital in schools.

Figure 3. Interaction between schools' lack of civic capital and external control in classrooms (direct control)

a. Probability of teacher cheating



b. Difference in probability

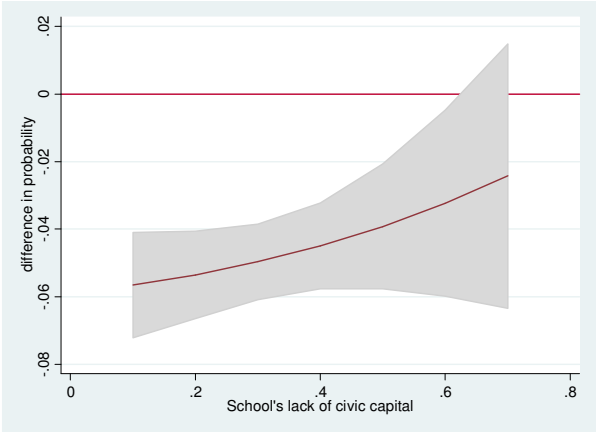
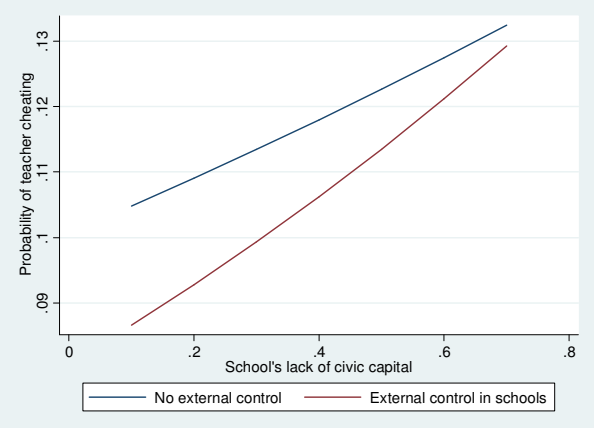
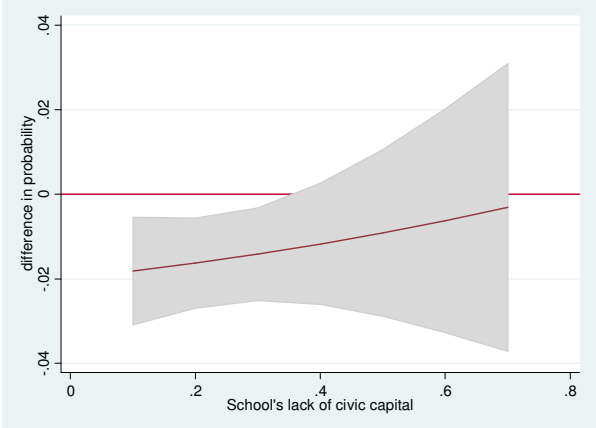


Figure 4. Interaction between schools' lack of civic capital and external control in schools (indirect control)

a. Probability of teacher cheating



b. Difference in probability



Figures 3a and 4a show that the strong association between the lack of civic capital and teacher cheating is also observed in classrooms with direct and indirect external control: although the predicted probability is always higher in unmonitored classrooms, even in classrooms with external control, the less civic capital in a school, the more cheating behaviours are found. If we compare the association trends between monitored and unmonitored classrooms, we observe that they tend to converge when schools show high levels of non-civic-minded behaviours. This means that, especially in the case of indirectly controlled classrooms, the predicted probability of cheating is the same as in unmonitored classrooms (13%).

This is best illustrated in figures 3b and 4b, which show the difference in probability of cheating, depending on the external control. In figure 3b, the difference in probability between classrooms with and without direct external control is not statistically significant when the lack of civic capital is very high. This means that direct control in classrooms during the test taking is an effective deterrent, except in those schools with very low levels of civic capital. Conversely, with regard to indirect control, figure 4 indicates that indirect control is an effective deterrent when the school has medium or high levels of civic-minded capital. Specifically, the difference between the levels of teacher cheating between unmonitored and indirectly monitored classrooms is no longer significant when the lack of civic-minded capital is in the 80th percentile (0.37).

Context-related factors

As indicated in table 1, the rate of teacher absenteeism may be approached as a school system-related factor, but also as a context-related factor. In either case, it is a powerful predictor of lack of civic capital, as the results indicate a solid association between teacher cheating and the extent to which teachers are absent from the school, significant and robust across the grades, and also in both tests (tables 7a and 7b, appendix 2). As regards the magnitude of the association, the tables show that, when the rate of teacher absenteeism at the province level increases by a one-percent unit, the probability of a teacher being suspected of cheating increases between 1.3% and 1.8%, depending on both the grade and the test. Even when controlling for regional dummies (instead of macro-area dummies), the association is positive and significant, with the sole exception of the maths test in the upper secondary school.

We further analysed the extent to which teachers in schools located within contexts with a lack of civic capital are more likely to cheat. As described earlier, we included, in separate specifications, three factors related to diverse illicit economic activities: the evasion of public television fees, the risk of tax evasion, and the estimated rate of submerged work. As can be seen in tables 9a and 9b (appendix 2), the results are in line with Paccagnella and Sestito (2014) when dealing with universalistic social values, which appear to be negatively correlated with cheating. In our case, since measures of social capital refer to the *lack* of

civic-minded capital in schools, we find that all measures of illicit economic activities are positively associated with teacher cheating.

Here we confirm that the cultural context in which schools are located is strongly correlated with cheating: teachers are more likely to cheat in contexts where more deviant behaviours – behaviours against laws and regulations – are found. This remains significant even after accounting for a wide range of province characteristics, including the socioeconomic composition, and also for cultural and economic differences across macro-areas. The only exception is the indicator of public TV fee evasion, which is not as strongly associated with cheating as are the other indicators of deviant behaviour. Specifically, the association is not significant in the lower secondary (maths), and in the upper secondary school (maths and Italian). This weaker association might be due to the heterogeneous profile of people who do not pay the public TV fee across provinces, which means that, in many cases, it is no longer in accordance with a lack of civic-minded capital.

CONCLUSIONS AND DISCUSSION

This paper contributes to the educational policy debate, providing a comprehensive framework to understand and explain the rationale and incentives behind teacher cheating in standardised tests. Aside from providing further evidence to confirm the general factors associated with cheating, which have been highlighted by other studies (relative to the geographical location of schools, the effect of high-stakes testing and external controllers during the test taking), we present empirical evidence to understand teacher cheating as both stock and lack of civic social capital. Obviously, here this capital is approached from two distinct perspectives: *network-based* and *culture-based*.

Firstly, the findings are consistent with the *network-based* approach, insofar as cheating behaviour was found to be linked to the *bonding* form of social capital. In socially homogeneous contexts, where strong ties are more likely, cheating levels were found to be higher and addressed to help, to a large extent, the most socially and academically disadvantaged students – students with a low socioeconomic status and grade-retained. Along with this theoretical perspective, teacher cheating may be socially admissible behaviour, even if it is detrimental to the society at large and contravenes the deontological standards of teachers, since it is presented as a form of community-based support. According to Bandura and et al. (1996), detrimental conducts may be considered personally and socially acceptable depending on the extent to which they are portrayed in the service of valued social or moral purposes. This is a key factor in the process of moral justification of such behaviour, and could be reinforced by an adverse socioeconomic context, exogenously determined, perceived as unfair.

The results are also in line with the *culture-based* approach of social capital, with teacher cheating understood as a deviant behaviour that could reflect collective non-civic-minded

beliefs and values. This has been highlighted by Paccagnella and Sestito (2014), who showed how cheating is negatively associated with proxies of trust towards education authorities and non-adherence to the rule of law. In our paper, since teacher cheating was found to be associated with other collectively shared illicit behaviours, we assume the culturalist approach of deviance and argue that deviance may be socially constructed and linked to rules of behaviour created in particular social sub-groups. The findings confirm how teachers in schools with lower levels of civic capital (non-civic-minded practices), or within contexts which have a lack of civic capital (teacher absenteeism or tax evasion), are significantly more likely to cheat. A relevant and original contribution of this paper refers to the significant and robust correlation between teacher cheating and school-based dishonest practices. Specifically, it was found that teachers are more likely to cheat in schools which undertake practices that go against legal requirements or recommendations, such as social tracking of students between classes, and exclusion of students from tests.

REFERENCES

Agasisti, T., & Falzetti, P. (2013). Between-classes sorting within schools and test scores. An empirical analysis of the Italian junior secondary schools. Working Paper, 20/2013. INVALSI.

Akers, R. L., & Jennings, W. G. (2009). Social learning theory. In J. Miller (Ed.), *21st Century criminology: a reference handbook*. Thousand Oaks: SAGE Publications, Inc.

Akers, R. L., Krohn, M. D., Lanza-Kaduce, L., & Radosevich, M. (1979). Social learning and deviant behavior: a specific test of a general theory. *American Sociological Review, 44*(4), 636-655.

Amrein, A. L., & Berliner, D. C. (2002). An analysis of some unintended and negative consequences of high-stakes testing. EPSL-0211-125-EPRU. Education Policy Research Unit, Arizona State University.

Angoff, W. H. (1974). The development of statistical indices for detecting cheaters. *Journal of the American Statistical Association, 69*(345), 44-49.

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education, 31*(2), 101-119.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71*(2).

Bélair, L. M. (2005). Les dérives de l'obligation de résultats ou l'art de surfer sans planche. In C. Lessard & P. Meirieu (Eds.), *L'obligation de résultats en éducation. Evolutions, perspectives et enjeux internationaux*. Bruxelles: De Boeck.

Belleza, F. S., & Belleza, S. F. (1989). Detection of cheating on multiple-choice tests by using error similarity analysis. *Teaching of Psychology, 16*(3), 151-155.

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

Campbell, D. T. (1976). Assessing the impact of planned social change. Occasional Paper Series, 8. The Public Affairs Center, Dartmouth College.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94, Supplement: Organizations and Institutions:Sociological and Economic Approaches to the Analysis of Social Structure*, S95-S120.

de Blasio, G., Scalise, D., & Sestito, P. (2014). Universalism vs. particularism: a round trip from sociology to economics. Questioni di Economia e Finanza, 212. Banca d'Italia.

Ferrer-Esteban, G. (2011). Beyond the traditional territorial divide in the Italian education system. Effects of system management factors on performance in lower secondary school. FGA Working Paper, 43. Fondazione G. Agnelli.

Ferrer-Esteban, G. (2013, 8 February). *Il 'cheating' nel sistema italiano di valutazione.* Paper presented at the workshop 'metodi di identificazione, analisi e trattamento del cheating', Roma.

Fondazione G. Agnelli. (2014). *La valutazione della scuola. A che cosa serve e perché è necessaria all'Italia*. Bari: Laterza.

Frary, R. B. (1993). Statistical detection of multiple-choice answer copying: review and commentary. *Applied Measurement in Education, 6*(2), 153-165.

Gittell, R., & Vidal, A. (1998). *Community organizing. Building social capital as a development strategy*. Thousand Oaks, CA: Sage Publications, Inc.

Guiso, L., Sapienza, P., & Zingales, L. (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives, 20*(2), 23-48.

Guiso, L., Sapienza, P., & Zingales, L. (2010). Civic capital as the missing link. NBER Working Paper, 15845. National Bureau of Economic Research.

Haladyna, T. M., Nolen, S. B., & Haas, N. S. (1991). Raising standardized test scores and the origins of test score pollution. *Educational Researcher, 20*(5), 2-7.

Horn, D. (2012). Catching Cheaters in Hungary estimating the ratio of suspicious classes on the National Assessment of Basic Competencies tests. Institute of Economics, Research Centre for Economic and Regional Studies (Hungarian Academy of Sciences), Department of Economics (Eötvös Lóránd University).

Jacob, B. A., & Levitt, S. D. (2003). Rotten apples: an investigation of the prevalence and predictors of teacher cheating. *The Quarterly Journal of Economics, 118*(3), 843-877.

Kohlberg, L. (1976). Moral stages and moralization: the cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behavior: theory, research and social issues*. Holt, NY: Rinehart and Winston.

Kohlberg, L. (1981). *The philosophy of moral development: moral stages and the idea of justice*. San Francisco: Harper and Row.

Lanza-Kaduce, L., & Klug, M. (1986). Learning to cheat: the interaction of moral-development and social learning theories. *Deviant behavior, 7*, 243-259.

Lucifora, C., & Tonello, M. (2012). Students' cheating as a social interaction: evidence from a randomized experiment in a national evaluation program. IZA Discussion Paper, 6967. Institute for the Study of Labor (IZA).

Madaus, G., West, M., Harmon, M., Lomax, R., & Viator, K. (1992). *The influence of testing on teaching math and science in grades 4-12*. Boston: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

Mons, N. (2009). *Theoretical and real effects of standardised assessment, background paper to the study: National Testing of Pupils in Europe: Objectives, Organisation and Use of Results*. Brussels: EACEA, Eurydice Network.

Morris, A. (2011). Student standardised testing: current practices in OECD countries and a literature review. OECD Education Working Papers, 65. OECD.

Nichols, S. L., & Berliner, D. C. (2002). The inevitable corruption of indicators and educators through high-stakes testing. EPSL-0503-101-EPRU. Education Policy Research Unit, Arizona State University.

Paccagnella, M., & Sestito, P. (2013, 8 february). *School cheating and social capital.* Paper presented at the Metodi di identificazione, analisi e trattamento del cheating, Roma.

Paccagnella, M., & Sestito, P. (2014). School cheating and social capital. Temi di Discussione (Working Papers), 952. Banca d'Italia.

Piaget, J. (1932). *Le jugement moral chez l'enfant*. Paris: F. Alcan.

Portes, A. (1998). Social capital: its origins and applications in modern sociology. *Annual Review of Sociology, 24*, 1-24.

Portes, A. (2000). The two meanings of social capital. *Sociological Forum, 15*(1), 1-12.

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy, 6*(1), 65-78.

Putnam, R. D. (2000). *Bowling alone: the collapse and revival of American community*. New York: Simon & Schuster.

Quintano, C., Castellano, R., & Longobardi, S. (2009). A fuzzy clustering approach to improve the accuracy of Italian student data. An experimental procedure to correct the impact of outliers on assessment test scores. *Statistica & Applicazioni, VII*(2), 149-171.

Solow, R. (1995). Trust: the social virtues and the creation of prosperity (book review). *The New Republic, 213*, 36-40.

Sotaridona, L. S. (2003). *Statistical Methods for the detection of answer copying on Achievement Tests.* University of Twente, Enschede.

Sotaridona, L. S., & van der Linden, W. J. (2006). Detecting answer copying when the regular response process follows a known response model. *Journal of Educational and Behavioral Statistics, 31*(3), 283-304.

Sutherland, E. H. (1947). *Principles of criminology*. Chicago: J.B. Lippincott Co.

Trigilia, C. (2011). Capitale sociale tra economia e sociologia: avanti con giudizio. In G. de Blasio & P. Sestito (Eds.), *Il capitale sociale: che cos'è e che cosa spiega*. Roma: Donzelli editore.

Wesolowsky, G. O. (1999). Detecting excessive similarity in answers on multiple choice exams. DeGroote School of Business Working Paper, 30. McMaster University, Michael G. DeGroote School of Business.

Woolcock, M., & Narayan, D. (2000). Social capital: implications for development theory, research, and policy. *The World Bank Research Observer, 15*(2), 225-249.

APPENDICES

Appendix 1. Construction of the cheating indicator based on suspicious answer strings

Measure 1: unlikely block of identical answers on consecutive questions

For the first measure, Jacob and Levitt (2003) predicted the likelihood that each student will give each answer on each question using past test scores, future test scores and background characteristics. Since we were not able to use longitudinal data, we only used background data – immigration background, SES status, and sex. We estimated a multinomial logit for each item on the exam (only one from different choices was correct) in order to predict how students responded to each question. We estimated this model using information from other students in the same grade and subject:

$$\Pr(Y_{isc} = j) = \frac{e^{\beta_j x_s}}{\sum_{j=1}^{J} e^{\beta_j x_s}} \qquad (6)$$

where Y_{isc} is the response of student *s* in class *c* on item *i*. The possible responses (J) range from 3 to 5, and x_s is a vector that includes the socio-demographic variables for student *s*. The predicted probability of a student of choosing a response is identified by the likelihood

to choose that response of other students (in the same grade and subject) with similar background characteristics. We cannot account for future and prior test scores, which unfortunately increase the likelihood of identifying unusually good teachers as cheaters. In contrast, we estimate the probability of selecting each possible response, rather than estimating the probability of choosing the correct response, conditional to relevant background variables. As stated by Jacob and Levitt (2003), by doing this we take advantage of any additional information that is provided by particular response patterns in a classroom. Then, using the estimates from this model, we calculated the predicted probability that each student would answer each item *i* in the way that she did, proving one measure per student *s* per item *i*:

$$p_{isc} = \frac{e^{\widehat{\beta}_k x_s}}{\sum_{j=1}^{J} e^{\widehat{\beta}_j x_s}} \quad \text{for } k = \text{response actually chosen by student s on item i.} \qquad (7)$$

Using the product over items within student, we calculated the probability that a student would answered a string of consecutive questions from item *m* to item *n* as she did:

$$p_{sc}^{mn} = \prod_{i=m}^{n} p_{isc} \qquad (8)$$

We took the product over all students in the classroom c, who had identical responses in the string. Then the product is:

$$\tilde{p}_{sc}^{mn} = \prod_{s \in \{z : s_{ic}^{mn} = \bar{s}_{sc}^{mn}\}} p_{sc}^{mn} \qquad (9)$$

where *z* is defined as a student, s_{sc}^{mn} as the string of responses for student *z* from item *m* to item *n*, and $\bar{s}_{sc}^{mn}$ as the string for student *s*. $\tilde{p}_{sc}^{mn}$ collapses to p_{sc}^{mn} for each student, and there will be *ns* distinct values within the class, to the extent that there are *ns* students in classroom *c*, and each student has a unique set of responses to these particular items. On the contrary, if students in class *c* have identical answers, only one value of $\tilde{p}_{sc}^{mn}$ will be found. This calculation has been repeated for all possible strings of length three to seven. Finally, the indicator of the least likely block of identical answers given on consecutive questions is created taking the minimum of the predicted block probability for each classroom:

$$\text{measure } 1_c = min_s(\tilde{p}_{sc}^{mn})$$

Measure 2: classroom average variance across all test items

The second measure, which aims to capture general patterns of similarity in student responses, is constructed in several steps. First, residuals for each of the possible choices that a student could have made for each item are calculated:

$$e_{jisc} = \begin{matrix} 0 - \frac{e^{\widehat{\beta}_j x_s}}{\sum_{j=1}^{J} e^{\widehat{\beta}_j x_s}} \text{if } j \neq k \\ 1 - \frac{e^{\widehat{\beta}_j x_s}}{\sum_{j=1}^{J} e^{\widehat{\beta}_j x_s}} \text{if } j = k \end{matrix} \quad (10)$$

where e_{jisc} is the residual for response *j* on item *i* by student *s* in classroom *c*.

Then, information for each student is combined in order to create a classroom level measure of the response to item *i*. First, we sum the residuals for each response across students within a classroom. This term should be near to zero if there is no within-class correlation:

$$e_{jic} = \sum_{s} e_{jisc} \quad (11)$$

Second, we sum across the possible responses for each item within classrooms. Then we square each of the component residual measures to accentuate outliers and divide by number of students in the class (ns_c) to normalize by class size:

$$v_{ic} = \frac{\sum_j e_{jic}^2}{ns_c} \quad (12)$$

The statistic v_{ic} captures something like the variance of student responses on item *i* within classroom *c*. In order to emphasize the classroom level tendencies in response patterns, we first sum across the residuals of each response across students and then sum the classroom level measures for each response, rather than summing across responses within student initially. The second measure of suspicious strings is the classroom average (across items) of this variance term across all test items:

$$measure\ 2_c = \bar{v}_c = \frac{\sum_i v_{ic}}{ni}$$

where *ni* is the number of items on the exam.

Measure 3: variance in the degree of correlation across questions within a classroom

$$measure\ 3_c = \sigma_{v_c}^2 = \frac{\sum_i (v_{ic} - \bar{v}_c)^2}{ni}$$

Measure 4: different student's response patterns

The fourth indicator focuses on the extent to which a student's response pattern was different from other student's with the same aggregate score that year. q_{isc} is equal to one if student *s* in classroom *c* answered item *i* correctly, and zero otherwise. A_s is equal to the aggregate score of student *s* on the exam. Then, it is determined what fraction of students at each aggregate score level answered each item correctly. If ns_A is equal to the number of students with an aggregate score of *A*, then this fraction $\bar{q}_i^A$ can be expressed as:

$$\bar{q}_i^A = \frac{\sum_{s \in \{z:A_z = A_s\}} q_{isc}}{ns_A} \tag{13}$$

Then, it is calculated a measure of how much the response pattern of student *s* differed from the response pattern of other students with the same aggregate score: it is subtracted a student's answer on item *i* from the mean response of all students with aggregate score *A*, squared these deviations and summed across all items on the exam:

$$Z_{sc} = \sum_i \left(q_{isc} - \bar{q}_i^A\right)^2 \tag{14}$$

The final indicator is calculated subtracting out the mean deviation for all students with the same aggregate score, $\bar{Z}^A$, and summing the students within each classroom:

$$measure\ 4_c = \sum_s (Z_{sc} - \bar{Z}^A)$$

As stated earlier, the overall measure of cheating is constructed, within a given subject and grade, by ranking classrooms on each of the four indicators, and taking the sum of squared ranks across the four measures. In the empirical work we use the 90th cut-offs (and the 95th in robustness checks[2]) to identify the classrooms suspected to cheat:

$$cheating_{cdg} = \left(rank\_measure1_{csg}\right)^2 + \left(rank\_measure2_{csg}\right)^2 + \left(rank\_measure3_{csg}\right)^2 + \left(rank\_measure4_{csg}\right)^2$$

where $cheating_{cdg}$ indicates whether a classroom *c* in subject *d* in grade *g* is suspected of cheating or not.

[2] Results available on request

Appendix 2. Baseline estimates

Table 8a. Teacher cheating and lack of civic capital: teacher absenteeism (Italian)

Suspected classroom of cheating	Primary Education - Teacher absenteeism	Lower Secondary - Teacher absenteeism	Upper Secondary - Teacher absenteeism	Primary Education - Teacher absenteeism	Lower Secondary - Teacher absenteeism	Upper Secondary - Teacher absenteeism
Social student heterogeneity in classroom	-0.190*	-0.333***	-0.351**	-0.173*	-0.311***	-0.339**
	(0.10)	(0.11)	(0.14)	(0.10)	(0.11)	(0.14)
Mid-SES classroom (ref: Low-SES classrooms)	-0.381***	-0.107*	-0.201***	-0.371***	-0.093	-0.165**
	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)
High-SES classroom (ref: Low-SES classrooms)	-0.656***	-0.134**	-0.813***	-0.643***	-0.130*	-0.766***
	(0.08)	(0.06)	(0.10)	(0.08)	(0.07)	(0.10)
Fraction of grade-retained students in classroom	0.664	0.389	1.317***	0.573	0.278	1.400***
	(0.48)	(0.30)	(0.17)	(0.49)	(0.31)	(0.16)
Students who missed the test	0.005	0.013***	0.018***	0.004	0.013***	0.019***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
Social tracking between classrooms	0.329*	0.541***	0.465**	0.298*	0.517***	0.520**
	(0.17)	(0.18)	(0.20)	(0.17)	(0.18)	(0.20)
Teacher absenteeism (province)	0.217***	0.150***	0.161***	0.184***	0.078***	0.130***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
Controllers in classroom during the test taking (ref: no control)	-1.194***	-0.636***	-0.998***	-1.100***	-0.550***	-0.884***
	(0.15)	(0.09)	(0.13)	(0.14)	(0.09)	(0.13)
Controllers in school during the test taking (ref: no control)	-0.328***	-0.154**	-0.275***	-0.226***	-0.065	-0.175***
	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.06)
Macro-area fixed effect	Yes	Yes	Yes	No	No	No
Region fixed effect	No	No	No	Yes	Yes	Yes
Observations	23518	22619	19534	23518	22619	19534
Pseudo R^2	0.1664	0.1010	0.1314	0.1741	0.1061	0.1365

Source: Invalsi dataset 2011-12. Robust standard errors clustered by province in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The dependent variable is the cheating indicator using the 90th percentile cut-off. The control variables include classroom, school and province factors. Classroom and school: classroom SES average, fraction of female students, the fraction of first-generation and second-generation immigrant students in the classroom, classroom size, and school size. Province: GDP per capita, unemployment rate, whether provinces contained a metropolitan area, population size and density of province, share of adult population participating in education, share of adult population with a low level of education, share of early school leavers, an index of precarious teachers, and the share of teacher turnover.

Table 8b. Teacher cheating and lack of civic capital: teacher absenteeism (maths)

Suspected classroom of cheating	Primary Education - Teacher absenteeism	Lower Secondary - Teacher absenteeism	Upper Secondary - Teacher absenteeism	Primary Education - Teacher absenteeism	Lower Secondary - Teacher absenteeism	Upper Secondary - Teacher absenteeism
Social student heterogeneity in classroom	-0.352***	-0.323***	-0.031	-0.336***	-0.302***	-0.033
	(0.10)	(0.11)	(0.17)	(0.11)	(0.11)	(0.17)
Mid-SES classroom (ref: Low-SES classrooms)	-0.300***	-0.139***	-0.285***	-0.296***	-0.124**	-0.247***
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)
High-SES classroom (ref: Low-SES classrooms)	-0.571***	-0.222***	-0.538***	-0.545***	-0.215***	-0.493***
	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)
Fraction of grade-retained students in classroom	0.469	0.915**	0.934***	0.334	0.780*	1.011***
	(0.52)	(0.40)	(0.25)	(0.51)	(0.41)	(0.24)
Students who missed the test	0.017***	0.028***	0.014***	0.015**	0.027***	0.016***
	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
Social tracking between classrooms	0.364*	0.578***	0.207	0.365*	0.520***	0.288
	(0.20)	(0.17)	(0.19)	(0.19)	(0.17)	(0.20)
Teacher absenteeism (province)	0.210***	0.170***	0.145***	0.157***	0.066*	0.048
	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
Controllers in classroom during the test taking (ref: no control)	-1.083***	-0.982***	-0.677***	-0.981***	-0.889***	-0.543***
	(0.11)	(0.14)	(0.11)	(0.11)	(0.14)	(0.10)
Controllers in school during the test taking (ref: no control)	-0.227**	-0.340***	-0.279***	-0.111	-0.241***	-0.168***
	(0.09)	(0.08)	(0.07)	(0.08)	(0.07)	(0.06)
Macro-area fixed effect	Yes	Yes	Yes	No	No	No
Region fixed effect	No	No	No	Yes	Yes	Yes
Observations	23772	22647	19743	23772	22647	19743
Pseudo R^2	0.2015	0.1439	0.1327	0.2119	0.1518	0.1403

Source: Invalsi dataset 2011-12. Robust standard errors clustered by province in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The dependent variable is the cheating indicator using the 90th percentile cut-off. The control variables include classroom, school and province factors. Classroom and school: classroom SES average, fraction of female students, the fraction of first-generation and second-generation immigrant students in the classroom, classroom size, and school size. Province: GDP per capita, unemployment rate, whether provinces contained a metropolitan area, population size and density of province, share of adult population participating in education, share of adult population with a low level of education, share of early school leavers, an index of precarious teachers, and the share of teacher turnover.

Table 9a. Teacher cheating and lack of civic capital: illicit economic activity (Italian language)

Suspected classroom of cheating	Primary Education - TV fee evasion	Lower Secondary - TV fee evasion	Upper Secondary - TV fee evasion	Primary Education - Risk of tax evasion	Lower Secondary - Risk of tax evasion	Upper Secondary - Risk of tax evasion	Primary Education - Shadow economy	Lower Secondary - Shadow economy	Upper Secondary - Shadow economy
Social student heterogeneity in classroom	-0.198**	-0.347***	-0.327**	-0.214**	-0.344***	-0.324**	-0.199**	-0.347***	-0.348**
	(0.10)	(0.11)	(0.14)	(0.10)	(0.11)	(0.14)	(0.10)	(0.11)	(0.14)
Mid-SES classroom (ref: Low-SES classrooms)	-0.364***	-0.095	-0.203***	-0.360***	-0.086	-0.203***	-0.358***	-0.090	-0.172***
	(0.05)	(0.06)	(0.07)	(0.05)	(0.06)	(0.07)	(0.05)	(0.06)	(0.06)
High-SES classroom (ref: Low-SES classrooms)	-0.655***	-0.130**	-0.795***	-0.641***	-0.122*	-0.796***	-0.625***	-0.120*	-0.758***
	(0.08)	(0.07)	(0.10)	(0.08)	(0.07)	(0.10)	(0.08)	(0.06)	(0.10)
Fraction of grade-retained students in classroom	0.827*	0.422	1.250***	0.822*	0.407	1.242***	0.746	0.363	1.283***
	(0.46)	(0.30)	(0.16)	(0.48)	(0.30)	(0.16)	(0.48)	(0.30)	(0.16)
Students who missed the test	0.005	0.013***	0.019***	0.003	0.012**	0.019***	0.006	0.013***	0.020***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
Social tracking between classrooms	0.397**	0.629***	0.470**	0.365**	0.580***	0.494**	0.339*	0.603***	0.546***
	(0.17)	(0.18)	(0.20)	(0.18)	(0.17)	(0.20)	(0.17)	(0.17)	(0.20)
Controllers in classroom during the test taking (ref: no control)	-1.208***	-0.610***	-1.009***	-1.179***	-0.581***	-0.987***	-1.161***	-0.554***	-0.946***
	(0.15)	(0.10)	(0.12)	(0.15)	(0.10)	(0.12)	(0.14)	(0.09)	(0.12)
Controllers in school during the test taking (ref: no control)	-0.338***	-0.148**	-0.283***	-0.311***	-0.122*	-0.264***	-0.275***	-0.090	-0.235***
	(0.08)	(0.07)	(0.07)	(0.08)	(0.07)	(0.06)	(0.08)	(0.06)	(0.06)
Deviant culture – illicit economic activity (province)	-0.017**	-0.014***	-0.007	-0.012***	-0.010***	-0.008***	0.073***	0.058***	0.048***
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Observations	23887	22992	19817	23887	22992	19817	23887	22992	19817
Pseudo R^2	0.1643	0.0980	0.1304	0.1647	0.0986	0.1311	0.1682	0.1018	0.1328

Source: Invalsi dataset 2011-12. Robust standard errors clustered by province in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The dependent variable is the cheating indicator using the 90th percentile cut-off. The control variables include classroom, school and province factors. Classroom and school: classroom SES average, fraction of female students, the fraction of first-generation and second-generation immigrant students in the classroom, classroom size, and school size. Province: GDP per capita, unemployment rate, whether provinces contained a metropolitan area, population size and density of province, share of adult population participating in education, share of adult population with a low level of education, share of early school leavers, an index of precarious teachers, and the share of teacher turnover. It includes macro-area fixed effects.

Table 9b. Teacher cheating and lack of civic capital: illicit economic activity (maths)

Suspected classroom of cheating	Primary Education - TV fee evasion	Lower Secondary - TV fee evasion	Upper Secondary - TV fee evasion	Primary Education - Risk of tax evasion	Lower Secondary - Risk of tax evasion	Upper Secondary - Risk of tax evasion	Primary Education - Shadow economy	Lower Secondary - Shadow economy	Upper Secondary - Shadow economy
Social student heterogeneity in classroom	-0.358***	-0.349***	-0.031	-0.378***	-0.345***	-0.028	-0.350***	-0.341***	-0.046
	(0.10)	(0.10)	(0.17)	(0.10)	(0.10)	(0.17)	(0.10)	(0.10)	(0.17)
Mid-SES classroom (ref: Low-SES classrooms)	-0.276***	-0.134**	-0.280***	-0.279***	-0.128**	-0.279***	-0.287***	-0.128**	-0.263***
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)
High-SES classroom (ref: Low-SES classrooms)	-0.560***	-0.210***	-0.521***	-0.544***	-0.205***	-0.520***	-0.549***	-0.202***	-0.509***
	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)
Fraction of grade-retained students in classroom	0.598	0.940**	0.872***	0.554	0.921**	0.865***	0.495	0.869**	0.922***
	(0.51)	(0.39)	(0.25)	(0.53)	(0.39)	(0.25)	(0.51)	(0.40)	(0.25)
Students who missed the test	0.018***	0.029***	0.016***	0.015***	0.028***	0.015***	0.021***	0.028***	0.017***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
Social tracking between classrooms	0.426**	0.619***	0.221	0.390*	0.590***	0.246	0.355*	0.625***	0.289
	(0.20)	(0.17)	(0.20)	(0.21)	(0.16)	(0.20)	(0.20)	(0.16)	(0.21)
Controllers in classroom during the test taking (ref: no control)	-1.075***	-0.973***	-0.683***	-1.032***	-0.948***	-0.655***	-1.022***	-0.935***	-0.645***
	(0.11)	(0.15)	(0.11)	(0.11)	(0.14)	(0.11)	(0.11)	(0.14)	(0.10)
Controllers in school during the test taking (ref: no control)	-0.234**	-0.346***	-0.279***	-0.193**	-0.323***	-0.254***	-0.163**	-0.306***	-0.251***
	(0.09)	(0.08)	(0.07)	(0.09)	(0.07)	(0.07)	(0.08)	(0.07)	(0.06)
Deviant culture – illicit economic activity (province)	-0.020***	-0.007	-0.008	-0.016***	-0.008**	-0.011***	0.068***	0.057***	0.047***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.02)	(0.01)	(0.01)
Observations	24142	23020	20031	24142	23020	20031	24142	23020	20031
Pseudo R^2	0.1999	0.1422	0.1325	0.2010	0.1430	0.1339	0.2040	0.1461	0.1369

Source: Invalsi dataset 2011-12. Robust standard errors clustered by province in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The dependent variable is the cheating indicator using the 90th percentile cut-off. The control variables include classroom, school and province factors. Classroom and school: classroom SES average, fraction of female students, the fraction of first-generation and second-generation immigrant students in the classroom, classroom size, and school size. Province: GDP per capita, unemployment rate, whether provinces contained a metropolitan area, population size and density of province, share of adult population participating in education, share of adult population with a low level of education, share of early school leavers, an index of precarious teachers, and the share of teacher turnover. It includes macro-area fixed effects.

Rationale and incentives for cheating in the standardised tests of the Italian assessment system

Gerard Ferrer-Esteban [a b]

[a] Fondazione Giovanni Agnelli

[b] Interdisciplinary Group on Educational Policies

(IGEP), University of Barcelona