

Cavalieri, Marina; Finocchiaro Castro, Massimo; Guccio, Calogero

**Working Paper**

## Does the Fish Rot from the Head? Organised Crime and Educational Outcomes in Southern Italy

*Suggested Citation:* Cavalieri, Marina; Finocchiaro Castro, Massimo; Guccio, Calogero (2021) : Does the Fish Rot from the Head? Organised Crime and Educational Outcomes in Southern Italy, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/228976>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Does the Fish Rot from the Head? Organised Crime and Educational Outcomes in Southern Italy

## Abstract

This paper explores the relationship between the presence of organised crime in local institutions and the educational outcomes in Southern Italy. To this purpose, we have built a contemporary index of *Mafia* institutional infiltration that proxies the (scale of) values that parents transmit to their children and that are likely to impact on their educational achievements. Furthermore, combining contemporary individual-level educational outcomes with historical data on mafia infiltration, we control for endogeneity concerns through an IV strategy. We find that a standard deviation increase in our indicator of *Mafia* institutional infiltration leads to a decrease of about 2 percentage points in the outcomes achieved by primary school students undertaking the INVALSI test for literacy in Southern Italy. Our results are robust to the use of different measures of organised crime, to the inclusion of different sets of controls, different subsamples and to relaxing the exclusion restriction in the IV strategy.

**JEL classification:** D73; I21; H72; K42

**Keywords:** Organised crime; *Mafia*-type organisations; Education outcomes; Investments in education; INVALSI; Italy

## 1. Introduction

The detrimental effects of organised crime on society and its economic system are well-known in the world. In fact, it badly affects political activity and policy makers (De Feo and De Luca, 2017; Dal Bò *et al.* 2006), the economic output (Pinotti, 2015), growth and investments (Dixit, 2004; Peri, 2004; Daniele and Marani, 2011). Among all the socio-economic features influenced by organised crime, education stands out because of its relevant contribution to economic growth in the long term. The first papers investigating the nexus between education and crime have been Buonanno and Leonida (2006, 2009) and Machin *et al.*, (2011) showing the causal link between education and crime in England and Wales. Then, other scholars have shown that crime can negatively affect the level of economic resources available to education (Ferraz *et al.*, 2012), the human capital stock (Coniglio *et al.*, 2010) and the honest attitude of students taking school exit exams (Borcan *et al.*, 2017).

We contribute this stream of literature by exploring the relationship between the presence of organised crime into local institutions and the educational outcomes of students, undertaking standardised tests (INVALSI) in the second and fifth grade of the Italian education system to assess their abilities in Italian language. For this purpose, we measure the presence of organised crime at provincial level by means of a novel index of organised crime institutional infiltration (IIM). The IIM has been computed by the ratio between the population of municipalities whose councils have been dismissed because of organised crime infiltration (D.L. 31/05/1991, n. 164) from 1991 to the 31st of July 2017 and the total population of the province.

There are several mechanisms through which the presence of organised crime in local institutions can affect the educational outcomes. First, political support for representatives suspected to have connection with mafia organization represent an indirect indicator of the level of adherence (or at least the low level of refusal) of the local communities to the *Mafia* social value (De Feo and De Luca, 2017). Building on those results, the underlying idea behind the use of the IIM index is to proxy the set of (social, cultural, human, etc.) values that permeate a certain community and that are transmitted from parents to children. In our opinion, these parental values are likely to play a role on educational outcomes but cannot be

observed and measured directly. Therefore, we infer them from the voting behaviours at local level.

Second, a large presence of individuals in the local community that adhere to *Mafia* values increases the likelihood that they will pass these values on to their children. Hence, we refer to the stream of literature focusing on cultural transmission from parents to their sons (Bisin and Verdier, 2010). Cultural traits, such as obedience to parents and teachers, respect for the law, and tolerance for different opinions, can persist from generation to generation if they are transmitted from parents to children. Bisin *et al.* (2004) and Tabellini (2008) show that individuals receive a package of values, update them (for instance, reflecting the local environment where they live), and transmit the updated package to their children. On this matter, Albanese *et al.* (2016) confirm the strong correlation between the values received from parents and those that the children, once grown up, have passed down to their descendants.

Finally, we also refer to the literature investigating the relationship between education and crime (for a detailed review, see Rud *et al.*, 2013). In particular, Lochner and Moretti (2004) show that investments in education can curb criminal behaviour, occasionally outperforming law enforcement because it makes the choice of crime more costly. However, the main result of Lochner and Moretti (2004) may not always be verified if the environment in which agents live is too affected by criminal organisations. In fact, Cruz (2019) shows that investment in education should be complemented with enough public safety in order to be effective in reducing criminality and violence. On this vein, looking at Colombia, Sviatschi (2017) shows that exposing kids to illegal labour market makes them more likely to be criminals as adults. Similarly, Jarillo *et al.* (2016) report the effect of drug-related violence in Mexico on academic outcomes. Finally, Brillì and Tonello (2015) show the presence of a causality nexus between high school attendance and adolescent crime and that it makes less likely that a crime reducing effect of education takes place. Moreover, the authors report that the crime increasing effect found in the South of Italy is consistent with a criminal capital accumulation process, operating through social interactions and crime networks.

Our results show that debauched local environment due to *Mafia* infiltration in local institutions affects education outcomes. More specifically, we find that a standard deviation increase in our indicator of *Mafia* institutional infiltration leads to a decrease of about 2 percentage points in the outcomes achieved by primary school students undertaking the INVALSI test for literacy in Southern Italy. Our results are robust to the use of different measures of organised crime, to the inclusion of different sets of controls, different subsamples and to relaxing the exclusion restriction in the IV strategy.

Thus, our findings provide robust evidence in favour of the link between the presence of organised crime into the local institutions and poor educational outcome. In details, having applied the procedure to control for test manipulations suggested by Angrist *et al.* (2017), we show that the higher the infiltration of organised crime in the local institutions, the lower the results of INVALSI test on Italian language obtained by the students of the province under analysis is.

Robustness checks, controlling for measurement errors with respect to different and wider measures of organised crime and for endogeneity concerns, through an IV strategy have been run. For instance, Bermpei, Kalyvas, and Leonida (2020) use interactions and instruments to capture the impact of public corruption on banking. In line with the main related literature, we employ, as instruments, measures related to the historical diffusion of *Mafia* in Southern Italian regions. Since the pioneering work of Bandiera (2003) on the historical origins of *Mafia*, contemporary measures of organized crime have been instrumented using geographic and historical variables to capture the institutional, socioeconomic, and environmental features of the places where organized crime originally emerged (e.g., Albanese & Marinelli, 2013; Barone & Narciso, 2015; Buonanno *et al.*, 2015; De Feo & De Luca, 2017). Specifically, our instruments are the diffusion of mafia organizations (Del Monte and Pennacchio, 2012) and the land productivity (Albanese and Marinelli, 2013) at the end of XIX century in the Southern Italian provinces.

As already mentioned, our IV results show to be also robust to relaxing the exclusion restriction, acknowledging that employed instruments based on historical presence of organised crime may only be plausibly exogenous. More specifically,

following Conley *et al.*, (2012), we conduct some sensitivity analysis to test the extent to which our estimates survive after allowing for plausible amounts of imperfect exogeneity of the instruments.

We provide evidence on possible mechanisms through which a large presence of individuals in the local community that adhere to *Mafia* values increases the likelihood that they will pass these values on to their children. Therefore, when a social environment is affected by high levels of organised crime, the typical values transmitted by parents referring to the role of education and the positive effects stemming from putting high effort when at school get seriously weakened. Our results show that poor performances in the southern regions of Italy are accompanied by the high presence of organised crime that, thus, lowers the expected returns to investment in education. Hence, families pay less attention to the educational outcomes of their sons who, in turn, put less effort in studying.

Furthermore, our results provide a rational and economic-based explanation to the issue of geographical divide on educational outcomes in Italy. This issue is still at the centre of a lively debate in the literature. Some scholars have suggested that socio-cultural factors, such as low “civic trust” (Putnam, 1993) or “trust and respect for others, and confidence in individual self-determination” (Tabellini, 2010) are at play. Others (Lynn and Mikk, 2007; Lynn, 2010) show that the poor educational outcome<sup>1</sup> achieved in southern Italian regions compared to those of the northern Italian regions are due to regional differences in intelligence quotient (IQ). According to them, the lower IQ in southern regions of Italy may be due to genetic merge with populations from the Near East and North Africa.

The rest of the paper is organised as follows. In the next Section, we report the literature review on the effects and measurement of organised crime. Section 3 describes the education system in Italy and the INVALSI evaluation service, whereas Section 4 illustrates the data and the methodology implemented. Sections 5 and 6 present and discuss the empirical results. Finally, Section 7 offers some concluding remarks.

---

<sup>1</sup> It is worth noting that the above-mentioned authors compute regional IQs from the 2006 PISA (Program for International Student Assessment) study of reading comprehension, mathematical ability, and science understanding administered to 15-year old’s in 52 countries (OECD, 2007), whereas we use the INVASI tests administered to primary school students in Italy.

## 2. Organised crime: effects and measurement issues

### 2.1 *The economic sound effects of organised crime*

Organised crime is commonly perceived as negatively affecting economic systems under different dimensions. On the one hand, violence and illegal activities reduce the human capital stock (Coniglio *et al.*, 2010). On the other hand, high levels of organised crime can create an unfavourable business climate, which works as disincentive for foreign and national investments (Dixit, 2004). Also, the quality and the accountability of the political class can be badly affected by the presence of organised crime (see Kugler *et al.*, 2005; Dal Bò and Di Tella, 2003; Dal Bò *et al.*, 2006). For instance, Pinotti (2015) looks at the relationship between organised crime and measures of economic and political performance showing that the presence of organised crime is associated with low levels of per capita economic output. Detotto and Otranto (2010) show the macroeconomic consequences of criminal activity in Italy in the long period.

Daniele and Marani (2011) report the negative effect of high levels of organised crime on foreign and national investments in Italy. In addition, Galletta (2017) shows the presence of spill over effects on the level of public investments directed to municipalities close to those whose city council has been dismissed because of the presence of *Mafia*-connected officials. Peri (2004) shows that the presence of organised crime is the most relevant socio-economic variable negatively correlated to economic development in Italy.

Considering all the possible aspects of an economy that can be affected by organised crime, our paper explores the role of debauched institutions and organised crime on primary school student outcomes, contributing to the stream of the literature studying the presence and the impact of organised crime on educational outcomes. For instance, Ferraz *et al.* (2012) study the effect of missing resources, due to corruption, on educational outcome of primary school students in Brazil and find a significant negative association between corruption and school performance of primary school students. Also, Borcan *et al.* (2017) report that a corruption fighting initiative such as introducing closed circuit television (CCTV)

monitoring of high school exit exam in Romania is effective in reducing corruption and increases the effectiveness of the punishment threats. Looking at the Italian case, Coniglio *et al.* (2010) show that the presence of organised crime lowers the accumulation of human capital by reducing the incentive to invest in formal education.

## *2.2 Measuring organised crime*

Notwithstanding the above-mentioned effects of organised crime on economic systems, the attempts at measuring its diffusion are quite scant. In fact, official statistics are often lacking and the number of complaints usually under-reports the real dimension of such phenomenon. Indeed, quantifying and understanding organised crime is a very complex task, because of its vague, hidden and multi-faceted nature that encompasses a number of dimensions, well beyond the sole identification of organised groups and the definition of the committed crimes (Transcrime, 2012). Moreover, the lack of an agreed international definition (Levi, 1998; Fijnaut & Paoli 2004; Van Duynes & Van Dijck 2007; Von Lampe, 2015) has made it difficult to develop accurate and more rigorous both quantitative and qualitative measures of the phenomenon so as to make reliable comparisons of crime trends and characteristics across countries. Therefore, to date, the different measuring approaches followed by international and European organisations, either based on official recorded data<sup>2</sup> or on international survey<sup>3</sup>, have each its own advantages and disadvantages and respond to specific purposes (see, Van Dijck 2007 for a detailed discussion).

Looking at the Italian case, data on single aspects of organised crime is routinely collected by the Italian National Institute of Statistics (ISTAT) in the

---

<sup>2</sup> See, for example, Statistics on Crime Trends and the Operations of Criminal Justice Systems collected by the United Nations Office on Drugs and Crime, the Eurostat Crime and Criminal Justice Statistics.

<sup>3</sup> Among the others, the World Economic Forum annual survey on obstacles to businesses; the Merchant International Group on countries' investment risks, the PricewaterhouseCoopers Global Economic Crime Survey, the Kroll Global Fraud Survey, the Ernst and Young's Global Fraud Survey.



*Territorial Information System on Italian Justice*<sup>4</sup>. This includes information on extortions, bomb attacks, arson and crimes of criminal association, crimes against property, thefts and robberies, etc. Other data on organised crime are yearly provided by the Ministry of Interior as part of its institutional activity.

A concise index that is often used to map the infiltration of *Mafia*-type organisations at provincial level in Italy is the one proposed by Transcrime (2013)<sup>5</sup>. This index (*Indice di Presenza Mafiosa*, IPM<sup>6</sup>) uses information on the number of *Mafia* murders or attempted *Mafia* murders, persons charged with *Mafia* association (Decree Act 416bis), municipalities and public administrations whose councils have been dismissed because of *Mafia* infiltration (Legislative Decree 31/05/1991, n. 164), assets seized and confiscated to organised crime, active criminal groups described by DIA and DNA<sup>7</sup> reports. The IPM is then normalised, with values ranging from 0 (minimum) to 100 (maximum)<sup>8</sup>.

Figure 1 depicts the distribution of IPM values across Italian provinces. With few exceptions (e.g. Rome and Turin), the figure shows that the southern provinces of Italy are the most badly affected by *Mafia*-type phenomena. Specifically, Naples is the province with the highest IPM value, followed by Reggio Calabria, Vibo Valentia and Palermo. This is in line with the strong presence of *Mafia*-related organisations in traditionally controlled southern regions such as Campania (*Camorra*), Apulia (*Sacra Corona Unita*), Calabria (*'Ndrangheta*) and Sicily (*Cosa Nostra*).

**- Figure 1 around here -**

---

<sup>4</sup> To this purpose, ISTAT has adopted the operational definition of organised crime as used by the Italian Ministry of the Internal Affairs, which includes mafia murders, attacks, arsons, serious robberies (e.g. bank. post offices, conveyors of bank and postal values, etc).

<sup>5</sup> Transcrime is the joint research Centre of the Catholic University of the Sacred Heart, the University of Bologna and the University of Perugia.

<sup>6</sup> The IPM is part of a more complex composite index of territorial risk of organised crime infiltration at provincial level (*Indice di Rischio Territoriale*, RT) developed by Transcrime (2013).

<sup>7</sup> DIA is the Italian Counter-*mafia* Investigative Agency, whereas DNA is the National Counter-*mafia* and Counter-terrorism Investigative Agency.

<sup>8</sup> More specifically, in this paper, the index is normalized to 1 in order to make the values of the 107 provinces more homogeneous, assigning 107 points to the first province for IPM, 106 to second and so on. The scores are then recalculated by attributing 1 to the first province and 0 to the last.

However, the IPM index suffers from some weaknesses, being the most relevant to us that it captures different and strongly heterogeneous aspects that are not always directly related with the presence of organised crime in government institutions. Hence, it does not appear to be the most appropriate indicator to answer our research question – namely, to investigate the relationship between organised crime in government institutions and educational outcome.

To overcome this problem, we develop an original index of *Mafia* institutional infiltration (IIM) computed as the ratio between the population of municipalities whose councils have been dismissed because of *Mafia* infiltration (Legislative Decree n. 164/91) from 1991 to the 31st of July 2017 and the total population of the province<sup>9</sup>. The IIM index will be discussed in the following sections. Nonetheless, Figure 2 provides a picture of the different levels of *Mafia* institutional infiltration across Italian provinces as measured by the IIM index. Compared to the previous Figure 1, here the phenomenon appears to be less homogeneously distributed in the South of the country, with few provinces (e.g., Potenza, Taranto and Enna) even showing very low IIM values.

**- Figure 2 around here -**

As mentioned in Section 1, the IIM index proxies the set of (social, cultural, human, etc.) values that permeate a certain community. Overall, the presence of criminal organisations, in fact, measured by means of different indexes, robustly affects the distribution of natural resources, the weakness of local institutions, the political competitions (Buonanno et al, 2015; Daniele and Geys, 2015; Acemoglu et al., 2019), and, in general, several economic factors environment as key determining factors (Albanese and Marinelli, 2013; Pinotti, 2015). On this vein, we employ the IIM index to assess the effect of the occurrence of *Mafia* institutional infiltration on educational outcomes.

---

<sup>9</sup>To avoid the risk of the index taking values above 1, the city councils that have been dismissed more than one time during the study period are counted just once. Furthermore, the index is normalised to the scale 0-1 according to a min-max criterion, where 1 = the highest value in the 107 provinces.

### 3. The evaluation of educational outcomes in Italian compulsory school

#### 3.1. The INVALSI test

The implementation of a national evaluation service of students' achievements in the Italian compulsory education system<sup>10</sup> has been long and troublesome. It can be traced back to the end of 1990 (Legislative Decree n. 258/1999), when the reform that granted school autonomy also led to the establishment of the National Institute for the Evaluation of the Educational System of Instruction and Training (hereafter, INVALSI).

In the period 2001-2003, three pilot surveys on learning outcomes in schools were completed but only among a limited sample and on a voluntary basis. In 2004 (Legislative Decree 286/2004), the institute was reorganised, formally assigning it the task of managing the systematic and recurrent evaluation of the learning outcomes in the Italian education system. During the same year, INVALSI carried out the first compulsory survey of primary schools. This was, however, strongly criticised, on the basis of the lack of a robust research protocol that was likely to generate not reliable empirical findings.

It was only in 2007 with the *Quaderno Bianco sulla Scuola*, a strategic document set by the government Prodi, that the need for a standardised national test was defined to counter “the lack of the culture of evaluation in Italy”, to avoid the risk of a highly fragmented education system (due to greater school autonomy), to improve schools' capacity to raise the quality of educational outputs and to meet learning outcome targets at national level. Afterwards, the test contents have been redefined based on experts' opinion and following an intense public debate.

Therefore, in 2007 INVALSI piloted standardised tests in a small sample of voluntary schools in the second and fifth grade, to assess students' abilities in Italian language and mathematics. The year after, these tests become compulsory for all schools at given grades (second, fifth, sixth, eighth and tenth grade)<sup>11</sup>. Students are

---

<sup>11</sup> More specifically, test taken in the sixth grade has been suppressed since the 2013/14 school year. For a detailed review of the Italian education system, see Appendix A.

thus requested to compile a hand-written test (the same for everyone), including both multiple choices and open-ended questions. These reflect Italian reading comprehension and grammar competencies as well as knowledge in mathematics, which they should have learned in their school career<sup>12</sup>. The questionnaire is also designed to collect several additional information about students themselves, their family, their parents' educational level and their socio-economic situation, which is then employed to construct a concise index of individual economic, social and cultural status (ESCS index).

The test is usually administered in the spring period by teachers of the schools (other than those of the class and of the subject tested). These are also asked to interpret and faithfully report students' original answers on pre-defined machine-readable score forms, which have then to be sent back to the INVALSI. Along with the risk of cheating during the test compilation, the transcription process by teachers leaves space to ex-post score manipulation phenomena. To control for these moral hazard problems, a representative random sample of schools is selected every year at national level where tests are monitored by external evaluators.

### *3.2. The INVALSI literature*

A recent body of empirical research has been devoted to the identification of the core determinants of the INVALSI test scores in Italy, distinguishing between students', family, school, classmates (peers) and context characteristics. Applying multilevel linear models, existing literature on individual and peer effects focuses on socio-economic status, gender and ethnic differences. Thus, evidence is found of gender (favouring males in mathematics achievements) as well as immigrant (i.e., 1<sup>st</sup> and 2<sup>nd</sup> generation)/native achievement gaps (Azzolini, 2011; Contini, 2013; Contini *et al.*, 2017). It has also emerged that the proportion of early enrollees in the classroom has a positive impact on average student's achievements (Ordine *et al.*, 2015). School characteristics have been primarily studied looking at the effect

---

<sup>12</sup> Since 2018 a new test to assess students' ability in English language has been introduced. Furthermore, all tests for the eight and tenth grades are now computer-based.

of school body composition, school size, managerial practices and characteristics of school principals (e.g. Masci *et al.*, 2018). Overall, class effects on student achievements have been found to be larger than school ones (Masci *et al.* 2016).

The existence of a geographical “learning divide” in the INVALSI student achievements has been a widely explored issue. As it can be seen by the left-hand side panels of Figures 3 and 4, the two distributions of the raw (observed) scores<sup>13</sup> in Italian language, respectively, at provincial level for the school years 2009-11 provide a similar and quite counterintuitive picture: second and fifth grade students in the South of Italy outperform those in the North. This geographical gradient seems highly unreliable in the light of the very much convergent regional patterns emerging from the Programme for International Student Assessment (OECD-PISA) (among others: Bratti *et al.*, 2007; Quintano *et al.*, 2009) and other international surveys<sup>14</sup> (Montanaro, 2008; IEA, 2011). Furthermore, the fact that southern students perform better than northern ones is in contrast with the marked backwardness of the South of Italy along many socio-economic indicators (e.g., unemployment rate, per-capita income, labour productivity, quality of governance and civic life indices, etc.) and the well-established literature on the direction of correlation between main predictors and student achievements (Battistin *et al.*, 2017).

- *Figure 3 around here* -

- *Figure 4 around here* -

---

<sup>13</sup> The scores have been standardised into a range [0;100], that represents the percentage of right answers to the questions of the test.

<sup>14</sup> Other international surveys largely employed in the Italian literature on educational outcomes include the Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS), which are both conducted by the International Association for the Evaluation of Educational Assessment (IEA). The former takes place every four years in 60 countries to measure mathematics and science achievements of fourth and eighth grade students. The latter is generally targeted to fourth graders to assess, every five years, students’ reading achievements.

The above problem has been largely discussed in the literature (Agasisti and Vittadini, 2012) and several explanations have been proposed for it, among which a higher motivation for southern students when compiling the test (INVALSI, 2008) and diverging grading standards among macro-regions (Argentin and Triventi, 2015). However, the most credible and agreed explanation for reconciling geographical patterns emerging from test surveys in Italy relies on the differences existing in the administration of tests. In the case of the INVALSI tests, the reliance on local single teachers (as compared to teams of teachers in TIMSS and PIRLS) is likely to give rise to opportunistic behaviours by both teachers and students. Score manipulation can take the form of either conventional cheating (by both students and teachers) or teachers' shirking in score transcription (Lucifora and Tonello, 2016)<sup>15</sup>. Indeed, manipulation of INVALSI test scores has proved to be considerably more widespread in the South and the Islands (Angrist *et al.*, 2017; Battistin *et al.*, 2017), being also correlated with measures of social capital (Paccagnella and Sestito, 2014).

To deal with these opportunistic behaviours, since 2009 INVALSI has adopted two different strategies. First, external inspectors are sent every year to a random and representative sample of classes to monitor the test administration and the marks reporting process. In this regard, using a natural experiment, Bertoni *et al.*, (2013) show that the introduction of external examiners has sizeable both direct (comparison between monitored and un-monitored classes within the same school) and indirect (comparison of un-monitored classes in schools with and without external examiners) negative effects on test performance, ranging between 4.3 and 6.6 percent and between 1.2 and 1.9 percent, respectively.

The second strategy adopted by INVALSI consists in improving the accuracy of data by recalculating the test scores based on the probability that cheating occurs in each classroom during the test. To this purpose, a subject-specific index (Cheating

---

<sup>15</sup> Lucifora and Tonello (2015) distinguish between many forms of opportunistic behaviours by teachers during and after the test administration. Thus, teachers may adopt a benevolent attitude by providing discretionary help to some students (generally, the lower performing ones) or lowering monitoring intensity and letting them interact and collaborate. Once the test has been compiled by the students, teachers may also deliberately altering or even changing student responses on answer sheets. The authors find significant social multiplier effects in student cheating, resulting from students' interactions in presence of teachers' tolerating behaviours.

Propensity Indicator, CPI) has been built according to which a class is considered at higher risk of cheating (i.e. “suspicious”) the more homogeneous is the answer behaviour for each item of the subject, the lower is the percentage of missing data, the higher is the average score, and the more close to zero is the within-class variability of the final score.<sup>16</sup>

The correction made by INVALSI to adjust for manipulation is, however, only able to affect marginally the geographical patterns emerging from raw scores (Battistin *et al.*, 2017). More specifically, Battistin *et al.* (2017) attempt to unravel true geographical patterns of INVALSI score data by constructing a manipulation index that replicates the CPI employed by INVALSI but replaces fuzzy clustering with hard clustering methods. This allows computing a binary variable to detect class outliers and to derive bounds for the average test scores net of manipulation. These bounds are found to be tight enough to reverse the evidence from raw score data for both Italian language and mathematics.

Therefore, a recent stream of literature has tried to develop statistical algorithms other and more sophisticated than the CPI used by INVALSI to detect cheating behaviours and correct observed test scores. Using a Maimonides-style instrumental variables (IV) approach (Angrist and Lavy, 1999), Angrist *et al.* (2017) estimate the effects of class size on INVALSI achievements, finding larger returns to class reductions in Southern Italy. However, these effects disappear after controlling for score manipulation (mostly in the form of teachers’ shirking), which is more likely to occur in smaller classes. To measure manipulation, the authors develop a specific indicator based on evidence of implausible score levels of students in a class, a doubtful within-class average, an unlikely small dispersion of test scores, an unusually low proportion of missing items and a high concentration in response patterns (as measured by the Herfindahl index). The indicator is similar to the CPI employed by INVALSI but generates a dummy (=1 for classes where score manipulation seems likely and 0 otherwise) instead of a continuous variable. The pictures emerging from the use of this dummy at provincial level for manipulation rates in Italian language and Mathematics are depicted in the central

---

<sup>16</sup> For further details on the construction of the CPI index and the ‘fuzzy clustering approach’ see Quintano *et al.* (2009).

hand side panels of Figures 3 and 4, respectively. They show a marked presence of manipulation especially in the southern areas of the country and, thus, corroborate the hypothesis of opportunistic behaviours playing a relevant role in explaining the geographical distribution of uncorrected scores. Indeed, by restricting the analysis only to non-suspicious classes according to the procedures followed by Angrist *et al.* (2017), the right-hand side panels of Figures 3 and 4 show that the regional gradient reverse for scores in Italian language but not for those in Mathematics. This result has influenced the choice of our dataset, which focuses mainly on Italian language achievements. Nonetheless we find comparable results also employing Math scores<sup>17</sup>.

## 4. Data

### 4.1 Data on Mafia infiltration

Data on the local government dissolutions due to (presumed) links to organised crime under the law no. 164/1991 is provided by the Italian Ministry of Internal Affairs and are used to develop the IIM index at provincial level, which, by construction (see Section 2.2), indicates the (normalised) proportion of the provincial population potentially exposed to *Mafia* phenomena. As said above, the IIM index is used to proxy the values transmitted by this population to their children. Indeed, we assume that the population exposed to mafia phenomena shares a common vision, concerning the importance of investing in education and the potential returns from it. This vision influences the effort that children put in

---

<sup>17</sup> See Table B.1 in the Appendix B, in which we report the estimates for model (1) focusing on math standardized scores. In Table B.1, although the geographical variables seem to confirm the need for a more adequate correction of cheating manipulation (Battistin *et al.*, 2017), the estimated impact of the index of organised crime is still significant and with the expected sign. Estimates are available from the authors upon request.



studying and, hence, their educational achievements. An alternative measure of organised crime employed in this paper is the IPM index of the presence of mafia-type organisations developed by Transcrime (2013), as reported in Section 2.2.

Table 1 provides the descriptive statistics of the two indices by geographical macro-area. Overall, some points are worth noting. First, looking at the entire Italian territory, both indexes show an extremely high dispersion of values. Considering the three macro-areas, as expected, moving from North to the South of the country, a clear geographical gradient appears, particularly for IPM. Notwithstanding the lower average values, the Northern part of Italy is not completely immune to the risk of criminal infiltration. However, the emerging picture across the Northern provinces is quite homogeneous (the standard variation is always well below the national average, especially for IIM). On the contrary, in the South of the country, the phenomenon of organised crime is more relevant but a higher variability in values exists across the single provinces.

*- Table 1 around here -*

#### *4.2 Data on educational outcomes*

To evaluate students' achievements, we use the same dataset employed by Angrist *et al.* (2017)<sup>18</sup>. Accordingly, data refers to INVALSI tests for Italian language in elementary schools (second and fifth grade students) in the 2009-11 academic school years<sup>19</sup>.

In our empirical analysis, we use data at class-level. Applying the same restrictions as in Angrist *et al.* (2017)<sup>20</sup>, the final sample comprises 140,010 classes. Our raw educational outcome variable is the average class percentage of corrected answers in Italian language (ANSWERS\_ITAL\_PCT). The variable is also

---

<sup>18</sup> Available online at the following link, <https://www.aeaweb.org/articles?id=10.1257/app.20160267&&from=f>

<sup>19</sup> The choice of neglecting test achievements in Mathematics is due to the previously discussed evidence concerning the effects of the manipulation correction.

<sup>20</sup> The applied restrictions concern class and school size, and the exclusion of classes in private schools. For further details, see Angrist *et al.* (2017).

standardised by grade and year of survey to have a mean of zero and a unit variance (ANSWERS\_ITAL\_STD), thus making OLS estimator unbiased. To account for score manipulation, we employ the dummy variable developed by Angrist *et al.* (2017), indicating classes where score manipulation seems likely (CHEAT\_ITAL= 1 for a “suspected” class and 0 otherwise).<sup>21</sup>

Table 2 reports descriptive statistics for our educational outcome variables and the index of score manipulation by geographical macro-area. The highest raw average scores are in the southern classes, where, however, dispersion in values is very high. As for score manipulation, Table 2 shows that, on average, around 5.5% of classes are flagged as “suspicious”. Manipulation rates, however, varies greatly throughout the country, reaching a minimum in the North (around 1.5% of classes) and a maximum in the South (around 11% of classes). Hence, the phenomenon seems to be more concentrated right there where the observed achievements in Italian language appear to be higher.

**- Table 2 around here -**

In Table 3, the analysis on educational outcome has been restrained to the sub-sample of non “suspicious” classes for score manipulation. Looking at the three macro-areas the previous picture reverses: best scores are now achieved by the northern classes while the worst scores are those gained by the southern ones.

**- Table 3 around here -**

#### 4.3 *Instruments*

In our IV analysis, we employ two instruments based on historical variables. Both are computed at the provincial level and available only for the Southern Italian regions. In this perspective, we follow Albanese and Marinelli (2013) that employ

---

<sup>21</sup> When using the cheating variable our initial sample is further restricted to 140,003 classes.

instruments potentially relevant to our framework. In fact, the previous literature suggests that the origin of organised crime in the South of Italy is largely uncorrelated with modern economic factors but rather to historical characteristics or local situations that have allowed the consolidation of this type of criminal organizations.

Specifically, we employ, as potential instruments, the diffusion of *Mafia* organizations at the end of XIX century (MAFIA\_HIST) at the provincial level, provided by Del Monte and Pennacchio (2012), and the land productivity (LANDPROD) measured by the total value of land tax per square kilometres at provincial level, taken from Albanese and Marinelli (2013). We discuss the validity of these instruments in Section 6.2.

## **5. Preliminary findings on the whole Italian sample**

### *5.1 Setting the stage*

In our opinion, the choice of adopting the data and the empirical approach provided by Angrist et al., (2017) has relevant advantages. First, due to its adoption in the literature it would be easier to draw some comparative results with the previous findings. Second, we can run our empirical exercise on the role of organised crime of class achievement reducing the potential risks of omitted variable bias and measurement errors.

However, contrary to Angrist *et al.* (2017), we are not interested in understanding the effects of class size on test achievements. Therefore, in this paper we employ a slightly different and simpler model specification, which, after controlling for several factors, assesses whether the class achievement in Italian language in a given province is associated to the presence of organised crime into the institutions of that province. To this purpose, through the school code, each class scores are linked to the province where the school is located.

Nevertheless, in our model, we employ a full set of controls at class, school and geographical level as in Angrist *et al.* (2017). More specifically, to control for

score cheating and manipulation, we include both the manipulation dummy (CHEAT\_ITAL) and a variable indicating the presence of randomly assigned external monitors (MONITORED=1 if the class is selected for monitoring and 0 otherwise).<sup>22</sup> The other control variables are dummies for the geographic macro-area location of the class (CENTRE, NORTH and SOUTH), the number of students enrolled in the grade at school (STUDENTS), the class size (CLSIZE), the proportion of female students in the class (FEMALE), the proportion of immigrant students (IMMIGRANTS\_BROAD), the proportions of students whose father is a high school dropout (DAD\_LOWEDU) or a high school graduate (DAD\_MIDEDU) or a college graduate (DAD\_HIGHEDU), the proportions of students whose mother is unemployed (MOM\_UNEMP) or a housewife (MOM\_HOUSEW) or employed (MOM\_EMPLOYED), and variables for missing values of these categories. Stratification controls include the year of the survey, the class grade and the region in which the class is located. Finally, we use standard errors clustered at both school and grade level for all estimates. Table 4 lists and describes all the covariates used in the models.

**- Table 4 around here -**

Before estimating our model, we conduct a preliminary exploratory analysis of our class data. Figure 5 graphs kernel density distributions of unstandardized (i.e., top, left and right-hand side panels) and standardised (i.e., bottom, left and right-hand side panels) class scores by geographical macro-area. These are obtained by pooling together score data on second and fifth grade classes for the school years 2009-2011. Looking at the two distributions on the left-hand side where class scores are not corrected for manipulation, those geographic patterns concerning the average class scores and score dispersion are clearly visible. Once data are corrected for manipulation (by excluding the “suspicious” classes) (i.e., distributions in the right-hand side) the kernel pictures become similar to those already described in

---

<sup>22</sup> Differently from the paper by Angrist *et al.* (2017), in our model we do not employ class size cut-offs based on Maimonides’ Rule.

Table 3. In details, the score distributions of the southern macro-area shift to the left, and the shape of their right tails starts resembling that of the distributions by the other macro-areas. Consequently, the mean values of the score distributions of the southern macro-area become worse than those by the other two macro-areas.

**- Figure 5 around here -**

However, the high heterogeneity of the southern regions in their manipulation behaviours (as resulting from the leopard-like image reported in Figure 1) poses some relevant questions. For, we question whether worse class achievements in the South of Italy are due to a geographic problem or to some ‘environmental’ factors (i.e., the infiltration of organised crime into the institutions) that are geographically distributed and likely to lower the expected returns to investment in education. Under the latter hypothesis, the geographic factor would act as a mere confounding in analysing the relationship between educational outcomes and organised crime.

To shed a light on this issue, in Figure 6, we replicate the previous kernel distributions of class scores but by the level (low, middle and high) of the IIM variable. It is worth noting that the corresponding kernel estimates in Figures 5 and 6 provide almost overlapping graphs, thus supporting our hypothesis on the role played by the institutional infiltration of organised crime in explaining the educational outcomes.

**- Figure 6 around here -**

## *5.2 Preliminary findings*

To test our hypothesis concerning the relationship between the presence of organised crime in government institutions and the INVALSI outcomes achieved by primary school students in Italy we follow several steps. First, we employ the full set of information on classes located in all Italian regions. Hence, we test the following linear model:

$$y_{i,g,k,j,t} = \alpha + \beta_1 AREA_j + \beta_2 CRIME_j + \beta_3 SCHOOL_{i,j,t} + \beta_4 CLASS_{i,g,k,j,t} + \beta_5 FIXED_{EF_{i,g,k,j,t}} + \epsilon_{i,g,k,j,t} \quad (1)$$

where  $y_{i,g,k,j,t}$  is the outcome score in class  $i$  in grade  $g$  at school  $k$  in province  $j$  in year  $t$ ,  $AREA$  is a categorical variable indicating the three Italian macro-areas (South is used as reference category),  $CRIME$  is our variable measuring the infiltration of organised crime into the institutions at provincial level (either IIM or IPM). The model also controls for a set of variables at  $SCHOOL$  and  $CLASS$  level (including the manipulation dummy), which have been previously discussed. Fixed effects ( $FIXED\_EF$ ) for grade, survey year and region are also included to capture specific unobservable factors that may affect the individual outcome. Standard errors clustered at both school and grade level are used for all estimates.

Looking at Eq. (1), conventional picture predicts that, after controlling for manipulation,  $\beta_1 > 0$ , meaning that the geographical factor matters in explaining test achievements in Italian language (i.e. the North and Centre outperform the South). The specific hypothesis that we want to test in this paper is that the organised crime matters, too (i.e.,  $\beta_2 < 0$ ).

Table 5 reports OLS estimates for different specifications of Eq. (1), using, as dependent variable, the standardised scores on Italian language ( $ANSWERS\_ITAL\_STD$ ). In column (1), we included macro-area controls only. After controlling for “suspicious” classes, the estimation results confirm the existence of “geographical” patterns, with higher performances in the North and in the Centre (the two coefficients are positive and highly statistically significant) compared to the South of the country. In column (2), we test the alternative specification where the IIM index at provincial level is included in the model, but the geographical variables are excluded. As expected, the sign of the IIM variable is negative and highly significant (at the 1% level). In column (3), the  $AREA$  variables along with the IIM index have been included in the model to check whether the effects emerging from the previous two specifications should be regarded or not as two sides of the same coin. As shown in Table 5, both the

geographical variables and the crime indicator continue to be highly statistically significant (at the 1% level) and with the correct signs, thus providing support that they are capturing different phenomena and telling different stories. In the last two columns of Table 5 the IIM index is replaced by the IPM indicator, alone (column (4)) and along with the geographical variables (column (5)).

The estimated coefficient accounting for the effect of organised crime on the educational outcome is very stable across specifications. Using the specifications in columns 2 and 3, one standard deviation increases in the IIM index (equivalent to 0.2081, see Table 4) is associated with a reduction in educational outcome of 0.0211 ( $0.2081 \times 0.1013$ ). Similarly, one standard deviation increases in the IPC index (equivalent to 0.2903, see Table 4), is associated with a reduction in educational outcome of 0.0278 ( $0.2903 \times 0.0956$ ) equivalent to about a twentieth of a standard deviation of class achievement. This effect seems somewhat small, although it is equivalent to almost half of the relative impact of geographical areas on the average class achievement. In fact, using the most conservative specification in column 6, one standard deviation increase in the variable NORTH is associated with an increase in class achievement of 0.0487, in comparison with the SOUTH<sup>23</sup>.

*- Table 5 around here -*

To fully exploit the variation of our dataset at provincial level, we collapse all the data at provincial level, adjusting the standard errors by using the number of students in each province as weights. The new dataset is thus a panel data of three waves conducted for the second and fifth grade in each province for a total of 618 observations. The results of this first assessment are reported in Table 6. Although some covariates have lost the statistical significance, by and large, the effects of institutional infiltration of organised crime on average class achievement at provincial level have been confirmed<sup>24</sup>.

---

<sup>23</sup> For the sake of clarity, we do not report the coefficient estimates of the other control variables in equation (1) throughout the paper. Overall, the estimates are coherent with the findings from previous studies. The full estimates are available from the authors upon request.

<sup>24</sup> Further robustness checks can be found in Appendix A.2

- Table 6 around here -

## 6. Organised crime and educational outcomes in Southern Italy

### 6.1 Baseline estimates

In this Section, we restrict the analysis to a subsample of classes located in southern regions that have been historically characterised by a high presence of *Mafia*-type organisations. By making this, we leave aside the North-South divide in educational achievements and we focus on explaining the existing variability in educational outcomes among southern classes. Doing so, we control for the effect of omitted characteristics that potentially can affect educational achievement. Furthermore, as shown by Figure 1 and 3, there are large differences in the presence of organised crime in this case too.

First, we compare the educational outcomes achieved by classes located in regions with an historical presence of organised crime with the performance of classes located in regions without an historical presence of organised crime<sup>25</sup>, respectively, by estimating the following specification:

$$y_{i,g,k,j,t} = \alpha + \beta_1 HM\_regions_j + \beta_2 CRIME_j + \beta_3 SCHOOL_{i,j,t} + \beta_4 CLASS_{i,g,k,j,t} + \beta_5 FIXED_{EF_{i,g,k,j,t}} + \epsilon_{i,g,k,j,t} \quad (2)$$

where  $y_{i,g,k,j,t}$  is the average outcome score in class  $i$  in grade  $g$  at school  $k$  in province  $j$  in year  $t$ ,  $HM\_regions$  is a dummy equal to 1 for Sicily, Calabria, and Campania, and equal to 0 for other regions,  $CRIME$  is our variable measuring the

---

<sup>25</sup> Two other regions in the South, Puglia and Basilicata, have also witnessed the presence of criminal organizations since the mid-1970s (Pinotti, 2015). Thus, we also run our analysis with the inclusion of those regions. The results largely overlap those presented here. These additional estimates are available from the authors upon request.



infiltration of organised crime into the institutions at provincial level (either IIM or IPM). The model also controls for a set of variables at SCHOOL and CLASS level (including the manipulation dummy), which have been previously discussed. Fixed effects (FIXED\_EF) for grade, survey year and region are also included. Table 7 shows that, once again, both the coefficient of IPM and IIM are still highly significant and with the expected sign. The results show that classes located in regions with historical presence of *Mafia* organizations do not gain higher education outcomes than those achieved by classes located in regions without historical presence of *Mafia* organizations. For instance, this result may be due to the high heterogeneity in terms of historical presence of *Mafia* organisations within each southern region.

**- Table 7 around here -**

Table 8 provides the results of the subsample of southern regions including classes located in Campania, Calabria and Sicily only (columns (1) and (2)) and classes also located in Apulia (columns (3) and (4)). Finally, columns (5) and (6) report the estimates for all classes located in the southern regions. For each subsample, the estimates, employing either IIM or IPM, provide further support to the role of organised crime in affecting educational performance at class level (*i.e.*, all coefficients are negative and highly significant at the 1% level). Once again, the estimated coefficient of the effect of institutional penetration of organised crime on the educational outcome is remarkably stable across specifications. One standard deviation increases in the IIM index is associated with a reduction in educational outcome of ranging between 0.0181 to 0.0195. Conversely, IPC index shows larger variation. In fact, one standard deviation increases in the IPC index is associated with a reduction in educational outcome, ranging between 0.0435 to 0.0994 in different subsamples. Thus, the latter index captures different and heterogeneous aspects of organised crime.

**- Table 8 around here -**

## 6.2 Endogeneity issues

We are aware that, in previous estimates, some endogeneity concerns may rise, and measurement errors as well as the presence of possible omitted factors may be another potential source of bias, implying that OLS estimates cannot be interpreted as causal. Thus, to overcome these problems, we employ an instrumental variable (IV) strategy to control for such issues.

Since the pioneering work of Bandiera (2003) on the historical origins of *Mafia*, contemporary measures of organized crime have been explained using geographic and historical variables to capture the institutional, socioeconomic, and environmental features of the places where organized crime originally emerged (e.g., Albanese & Marinelli, 2013; Barone & Narciso, 2015; Buonanno et al., 2015; De Feo & De Luca, 2017).

For instance, Gambetta, (1996) observes that the origin of the Sicilian *Mafia* in XIX century has been traced back to the demand for protection from southern landlords when legal enforcement was not effective. In this perspective, Bandiera (2003) shows that, due to the importance of agriculture, the value of land area played a fundamental role in the rise of organised crime. In the same line of reasoning, other papers argue that the Sicilian *Mafia* emerged in response to a demand for protection from the more profitable economic sectors in XIX century (e.g. Buonanno et al., 2015, document the existence of a systematic causal link from sulphur availability to the emerge of *Mafia* in Sicilian municipalities, whereas Dimico et al., 2012, discuss the origin of the Sicilian *Mafia* and verify that its territorial expansion is linked to the land availability for the cultivation of citrus fruits).

In line with the abovementioned literature, we employ, as instruments, measures related to the historical diffusion of *Mafia* in Southern Italian provinces. In this perspective, we follow Albanese and Marinelli (2013) that employ some instruments potentially relevant to our framework. In fact, the previous literature suggests that the origin of organised crime in the South of Italy is largely uncorrelated with modern economic factors but rather to historical characteristics or local situations that have allowed the consolidation of this type of criminal

organizations. In particular, we employ, as instruments, the diffusion of *Mafia* organizations at the end of XIX century (MAFIA\_HIST) at the provincial level provided by Del Monte and Pennacchio (2012) and the land productivity (LANDPROD) measured by the total value of land tax per square kilometres at provincial level (Albanese and Marinelli, 2013).

We therefore exploit the effect of the historical diffusion of *Mafia* in Southern Italian provinces as an exogenous source of variation in current presence of organised crime in local institutions.

As a preliminary evidence, Table 9 reports the pairwise correlations between organised crime index and historical instruments in our dataset. Our instruments are significantly correlated with organised crime index (in particular with IIM). As expected, correlation coefficients are always positive and not negligible. This evidence supports, in the first place, our empirical strategy of using historical values as instruments for current presence of organised crime in local institutions.

**- Table 9 around here -**

Thus, we re-estimate (2) by 2SLS approach. The results reported in Table 10 support the connection we have previously conjectured. Specifically, in Table 10, columns from 1 to 3 report the results of our IV strategy employing the IIM index, whereas columns from 4 to 6 show the same estimates using the IPC index. Overall, the 2SLS estimates are qualitatively consistent with the OLS results reported in the last two columns of Table 8, and the magnitude of the coefficients is quite stable across specifications. Nevertheless, the magnitude of our IV estimates is largely higher than the corresponding OLS ones, suggesting that the latter are downward biased. In fact, using the full specification (column 3), we find that a one standard deviation increase in IIM causes about 0.05 points decrease in the standardized class achievement. An equivalent increase in IPC (column 6) causes a decrease in the same outcome of about two-thirds of its standard deviation.

**- Table 10 around here -**

The standard diagnostic tests reported in the Table 10 seem to generally support our IV identification strategy. Specifically, in each estimate, the endogeneity tests reject the null hypothesis that both IIM and IPC index can be treated as exogenous at conventional significance levels. Furthermore, the F-statistics of the Kleibergen-Paap test for weak identification reported in Table 9 indicate that our instruments LANDPROD and MAFIA\_HIST are not weak.

In all estimates, the first-stage coefficients are always negative and highly statistically significant, consistently with the idea that past socioeconomic characteristics and historical presence of *Mafia* traits shape the current presence of criminal organizations and negatively influence the educational outcomes.

Similarly, when we include both the instruments LANDPROD and MAFIA\_HIST in columns (3) and (6), the first-stage coefficients are still negative and statistically significant. Finally, the Hansen's J statistic does not reject our overidentifying restrictions, thus supporting the orthogonality condition (i.e., the error term in the second-stage equation is orthogonal to our instruments).

Although the reported tests do provide evidence for the validity of our historical instruments, we recognize that the requirement of perfect exogeneity is a quite strong one that potentially is unlikely to hold closely. Furthermore, the magnitude of the coefficients is even four times larger in some IV specifications relative to OLS, as shown in Table 10, suggesting the possibility that exclusion restriction assumption has not been met flawlessly.

To address the last concern about the validity of the historical instruments, we investigate the robustness of our estimates in circumstances in which the instrumental variable does not perfectly meets the exclusion restriction. In particular, we follow the approach proposed by Conley et al. (2012) to construct a confidence interval of the coefficient of interest employing an instrument that is only plausibly exogenous. More precisely, we replace the assumption of the exclusion restriction of no direct effect of the instrument on the outcome, considering a bound that consists of the union of all confidence intervals in the assumed range of reasonable values. Thus, relaxing the assumption of perfect exogeneity and examining the bounds, we are able to estimate the true effect of the

organised crime on educational outcomes when we deviate from perfect exogeneity.

In Table 11, we report the lower and upper 95% confidence intervals of the coefficients of interest for non-zero values of the direct effect of the instrument on the outcome in different plausible ranges (that should be equal to zero to perfectly met exclusion restriction). Remarkably, regardless of the magnitude of the (non-zero) plausibly endogenous relationship between the instruments and the outcome, the confidence bound of the estimated relationship of interest is quite consistent. This confirms that the impact of organised crime is negative and significant, even if we consider the adopted instruments only plausibly exogenous.

- *Table 11 around here* -

## **7. Concluding remarks**

Building on the commonly accepted fact that organised crime harms society under different aspects, in this paper we have focused our attention on its effects on education because of its long-term relevant contribution to economic growth. Our work adds on to the existing literature in different ways. First, we investigate the relationship between the infiltration of organised crime into the Italian institutions (used as a proxy for the parental values transmitted to children, among which the lower expected returns to investment in education) and the INVALSI achievements of primary school students in Italian language. In particular, having controlled for test manipulations as suggested by Angrist *et al.* (2017), we find that the higher the infiltration of organised crime, the lower the average percentages of corrected answers in the Italian language obtained by the classes of the province under analysis is.

We show that values transmitted by the family have a role on students' educational achievements. Where these values reflect *Mafia* culture and its priorities, students are expected to performance worse at school. In other words, we provide evidence of the fact that “the fish rots from the head”. However, we

acknowledge that returns from investment in education may not be so relevant if the local level of public safety is not high enough. Thus, the incentives to study might decrease and achieve a low output, showing that the benefits of schooling can only be seized if they are complemented with enough public safety (Cruz, 2019).

Second, we develop a new measure of the presence of organised crime at provincial level by constructing a novel index of *Mafia* institutional infiltration (i.e. the IIM index), computed by the ratio between the population of municipalities whose councils have been dismissed because of organised crime infiltration from 1991 to the 31st of July 2017 and the total population of the province.

Finally, we provide a rational explanation for the existing geographical “learning divide” between the North and South of Italy, alternative to those suggesting as main determinant low levels of either socio-cultural values (e.g., Putnam, 1993: low “civic trust”; Tabellini, 2010: “trust and respect for others, and confidence in individual self-determination”) or IQ in southern regions of Italy (Lynn and Mikk, 2007; Lynn, 2010). Our results show that poor performances in the southern provinces of Italy are due to the high presence of organised crime that lowers the expected returns to investment in education. Consequently, families pay less attention to the educational outcomes of their sons who, in turn, put less effort in studying.

From a policy perspective, the results arising from this paper outline few directions for an educational policy agenda. Specifically, to achieve better students’ performance in those geographical areas where INVALSI scores are lower, policy makers should pursue a twofold approach by: 1) increasing the perceived returns to investment in education (e.g. by reducing the direct and indirect educational costs; by providing fiscal deductions for school expenditures, by giving subsidies to targeted families, etc.); 2) by contrasting the infiltration of organised crime into the institutions and fighting the *Mafia* culture, that is by “cutting the head of the fish”.

## References

Acemoglu, D., De Feo, G., & De Luca G. (2019). Weak states: Causes and consequences of the Sicilian mafia. *The Review of Economic Studies*, rdz009, <https://doi.org/10.1093/restud/rdz009>.

- Agasisti, T., & Vittadini, G. (2012). Regional economic disparities as determinants of students' achievement in Italy. *Research in Applied Economics*, 4(2), pp. 33-54.
- Albanese, G., & Marinelli, G. (2013). Organised crime and productivity: Evidence from firm-level data. *Rivista Italiana degli Economisti*, 18, pp. 367-394.
- Albanese, G., De Blasio, G., & Sestito, P. (2016). My parents taught me. Evidence on the family transmission of values. *Journal of Population Economics*. 29(2), pp. 571-592.
- Angrist, J. D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, 114(2), pp. 533-575.
- Angrist, J. D., Battistin, E., & Vuri, D. (2017). In a small moment: class size and moral hazard in the Mezzogiorno. *American Economic Journal: Applied Economics*. 9(4), pp. 216-249.
- Argentin, G., & Triventi, M. (2015). The North-South Divide in School Grading Standards: New Evidence from National Assessments of the Italian Student Population. *Italian Journal of Sociology of Education*, 7(2), pp. 157-185.
- Azzolini, D. (2011). A new form of educational inequality? What we know and what we still do not know about the immigrant-native gap in Italian schools. *Italian Journal of Sociology of Education*, 3(1), pp. 197-222.
- Bandiera, O. (2003). Land reform, the market for protection, and the origins of the Sicilian mafia: theory and evidence, *Journal of Law, Economics, and Organization*, vol. 19(1), pp. 218-44.
- Barbieri G., Rossetti, C., & Sestito, P. (2017). Teacher motivation and student learning, *Politica Economica/Journal of Economic Policy*, 33(1), pp. 59-72,
- Battistin, E., De Nadai, M., & Vuri, D. (2017). Counting rotten apples: Student achievement and score manipulation in Italian elementary schools. *Journal of Econometrics*, 200(2), pp. 344-362.
- Bermpei, T., Kalyvas, A. N., & Leonida, L. (2020). Local Public Corruption and Bank Lending Activity in the United States. *Journal of Business Ethics*, pp. 1-26.
- 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, pp. 65-77.
- Bisin, A., Topa, G., & Verdier, T. (2004). Religious intermarriage and socialisation in the US. *Journal of Political Economy* 112(3), pp. 615-664.

- Bisin, A., & Verdier, T. (2010). "The economics of cultural transmission and socialization". In: Behnabib J, Bisin A, Jackson M (eds.) *Handbook of Socio Economics* vol.1A. Elsevier, New York.
- Borcan, O., Lindahl, M. & Mitrut, A. (2017). Fighting corruption in education: What Works and who benefits?. *American Economic Journal: Economic Policy*, 9(1), pp. 180-209.
- Bratti, M., Checchi, D., & Filippin, A. (2007). Geographical differences in Italian students' mathematical competencies: evidence from PISA 2003. *Giornale degli Economisti e Annali di Economia*, 66(3), pp. 299-333.
- Brilli, Y., & Tonello, M. (2015), *Rethinking the crime reducing effect of education? Mechanisms and evidence from regional divides*. Temi di Discussione, n. 1008, Bank of Italy.
- Buonanno, P., & Leonida, L. (2006). Non-market effects of education on crime: Evidence from Italian regions. *Economics of Education Review*, 28(1), pp. 11-17.
- Buonanno, P., & Leonida, L. (2006). Education and crime: evidence from Italian regions. *Applied Economics Letters*, 13(11), pp. 709-713.
- Buonanno, P., Vanin, P., Prarolo, G. and Durante, R. (2015). Poor institutions, rich mines: resource curse and the origins of the Sicilian mafia, *The Economic Journal*, 125(586), pp. F175-202.
- Buonanno, P., Prarolo, G., and Vanin, P. (2016). Organised crime and electoral outcomes. evidence from Sicily at the turn of the xxi century. *European Journal of Political Economy*, 41, pp.61-74.
- Coniglio, N.D., Celi, G. & Scagliusi, C. (2010). Organised crime, migration and human capital formation: Evidence from the South of Italy, *Southern Europe Research in Economic Studies* (S.E.R.I.E.S), Working Paper No. 0028.
- Conley, T. G., Hansen, C. B. & Rossi, P. E. (2012). Plausibly Exogenous. *The Review of Economics and Statistics*, 94(1), 260–272.
- Contini, D. (2013). Immigrant background peer effects in Italian schools. *Social Science Research*, 42(4), pp. 1122-1142.
- Contini, D., Di Tommaso, M.L., & Mendolia, S. (2017). The gender gap in mathematics achievement: Evidence from Italian data. *Economics of Education Review*, 58, pp. 32-42.
- Cruz, L.I.G. (2019). Policing, schooling and human capital accumulation. *Journal of Economic Behavior and Organization*, 159, pp.572-597.



- Dal Bó, E., Dal Bó, P., & Di Tella, R. (2006). “Plata o plomo?”: Bribe and punishment in a theory of political influence. *American Political Science Review*, 100(1), pp. 41-53.
- Dal Bó, E., & Di Tella, R., (2003). Capture by threat. *Journal of Political Economy*, 111(5), pp. 1123-1154.
- Daniele, V., & Marani, U. (2011). Organised crime, the quality of local institutions and FDI in Italy: A panel data analysis. *The Economic Journal*, 125, pp. F233-255.
- Daniele, V., & Geys, B. (2013). Organised crime, institutions and political quality: Empirical evidence from Italian municipalities. *European Journal of Political Economy*, 27(1), pp. 132-142.
- De Feo, G., & De Luca, G. (2017). Mafia in the ballot box. *American Economic Journal: Economic Policy*, 9(3), pp. 134–167.
- Del Monte A. & Pennacchio L. (2012). Agricultural Productivity, Banditry and Criminal Organisations in Post-Unification Italy. *Rivista Italiana degli Economisti*, 17, pp. 347-378.
- Detotto, C., & Otranto, E. (2010). Does Crime Affect Economic Growth? *Kyklos*, 63(3), pp. 330-345.
- Dimico, A., Isopi, A. & Olsson, O. (2012). Origins of the Sicilian mafia: the market for lemons’, *Scandinavian Working Papers in Economics* No. 532.
- Dixit, A. K. (2004). *Lawlessness and economics: alternative modes of governance*. Princeton, NJ: Princeton University Press.
- Ferraz, C., Finan, F., & Moreira, D.B. (2012). Corrupting learning: Evidence from missing federal education funds in Brazil. *Journal of Public Economics*, 96(9), pp. 712-726.
- Fijnaut C., & Paoli, L. (2004). The History of the concept. In: C. Fijnaut and L. Paoli (eds.), *Organised crime in Europe; Concepts, patterns and control policies in the European Union and beyond*. Dordrecht: Springer.
- Galletta, S. (2017). Law enforcement, municipal budgets and spillover effects: Evidence from a quasi-experiment in Italy. *Journal of Urban Economics*, 101, pp. 90-105.
- Gambetta, D. (1993). *The Sicilian Mafia: The Business of Private Protection*, Cambridge, MA: Harvard University Press.
- Kugler, M., Verdier, T., & Zenou, Y., (2005). Organised crime, corruption and punishment. *Journal of Public Economics*, 89(9), pp. 1639-1663.

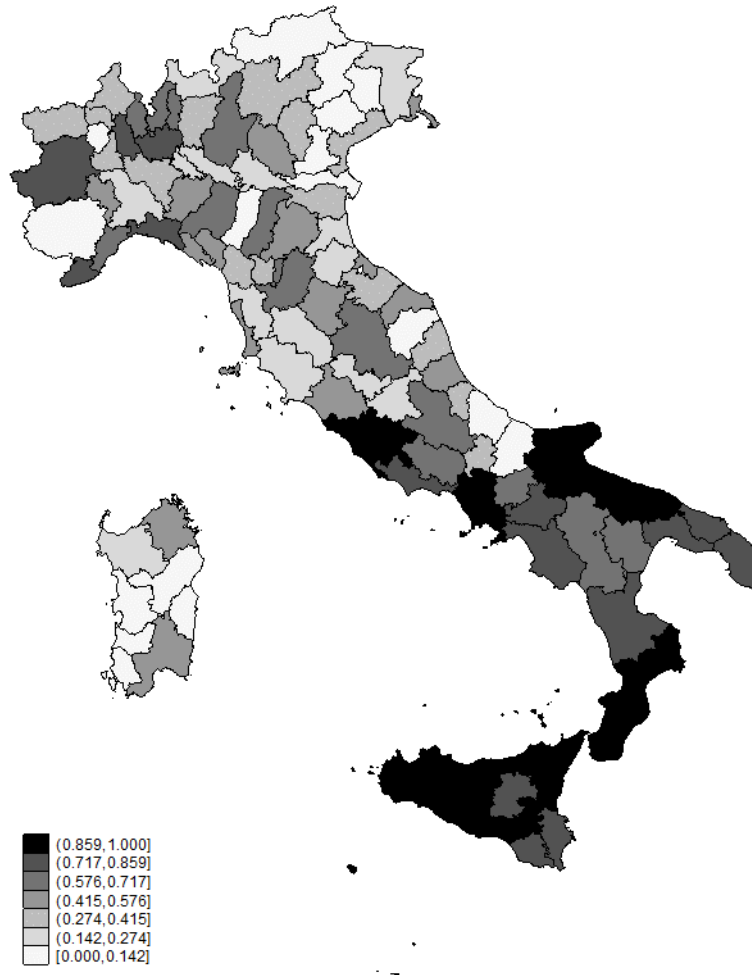
- Jarillo, B., Magaloni, B., Franco, E., & Robels, G., (2016). How the Mexican drug war affects kids and schools? Evidence on effects and mechanisms. *International Journal of Educational Development*, 51, pp. 135-146.
- International Association for the Evaluation of Educational Assessment (IEA) (2011): “IEA’s progress in international reading literacy study - Indagini IEA 2011 PIRLS e TIMSS,” Technical Report. Retrieved from: [http://www.invalsi.it/invalsi/ri/timss2011/documenti/Rapporto\\_PIRLS\\_TIMSS.pdf](http://www.invalsi.it/invalsi/ri/timss2011/documenti/Rapporto_PIRLS_TIMSS.pdf).
- Istituto per la Valutazione del Sistema di Istruzione e Formazione (INVALSI) (2008). *La prova nazionale al termine del primo ciclo, 2007/08*, INVALSI, Roma.
- Levi, M. (1998). Perspectives on “organised crime”: An overview’ , *The Howard Journal*. 37(4): pp. 335-345 .
- Lynn, R. (2010). In Italy, north–south differences in IQ predict differences in income, education, infant mortality, stature, and literacy. *Intelligence*, 38(1), pp. 93-100.
- Lynn, R., & Mikk, J. (2007). National differences in intelligence and educational attainment. *Intelligence*, 35(2), pp. 115-121.
- Lochner, L., & Moretti, E., 2004. The effect of education on crime: evidence from prison inmates, arrests and self-reports. *American Economic Review* 94, 155-189.
- Lucifora, C., & Tonello, M. (2015). Cheating and social interactions. Evidence from a randomized experiment in a national evaluation program. *Journal of Economic Behavior & Organization*, 115, pp. 45-66.
- Lucifora, C., & Tonello, M. (2016). *Monitoring and sanctioning cheating at school: What works? Evidence from a national evaluation program*. No. def051, Università Cattolica del Sacro Cuore, Dipartimenti e Istituti di Scienze Economiche (DISCE). Retrieved from: <http://dipartimenti.unicatt.it/economia-finanza-def051.pdf>.
- Machin, S., Marie, O., & Vujić, S. (2011). The crime reducing effect of education. *The Economic Journal*, 121(552), pp. 463-484
- Masci, C., De Witte, K., & Agasisti, T. (2018). The influence of school size, principal characteristics and school management practices on educational performance: An efficiency analysis of Italian students attending middle schools. *Socio-Economic Planning Sciences*, 61(March), pp. 52-69.
- Masci, C., Ieva, F., Agasisti, T., & Paganoni, A.M. (2016). Does class matter more than school? Evidence from a multilevel statistical analysis on Italian junior secondary school students. *Socio-Economic Planning Sciences*, 54(June), pp. 47-57.

- Montanaro, P., (2008). *I divari territoriali nella preparazione degli studenti italiani: evidenze dalle indagini nazionali e internazionali*. Occasional papers, n. 14, Rome: Bank of Italy.
- Ordine, P., Rose, G., & Sposato, D. (2015). Early enrollees and peer age effect: First evidence from INVALSI data. *Journal of Education and Training Studies*, 3(2), pp. 70-82.
- Organisation for Economic Co-operation and Development (OECD) (2007). *PISA 2006: Science Competencies for Tomorrows World*. Paris: OECD.
- Organisation for Economic Co-operation and Development (OECD) (2016). *Education at a Glance 2016: OECD Indicators*. Paris: OECD.
- Paccagnella, M., & Sestito, P. (2014). School cheating and social capital. *Education Economics*, 22(4), pp. 367–388.
- Peri, G. (2004). Socio-cultural variables and economic success: evidence from Italian provinces 1951-1991. *Topics in Macroeconomics*, 4(1), pp. 1-36.
- Pinotti, P. (2015). The economic costs of organised crime: Evidence from Southern Italy. *The Economic Journal*, 125(586), pp. F154-F178.
- Putnam, R.D. (1993). *Making democracy work*, Princeton, NJ, Princeton University Press.
- Quintano, C., Castellano, R., & Longobardi, S. (2009). The influence of social and economic factors on the competencies of Italian students. A multilevel analysis of data OECD-PISA2006. *Rivista di Economia e Statistica del Territorio*, 2, pp. 109-149.
- Rud, I., Van Klaveren, C., Groot, W., Maassen van den Brink, H., 2013. Education and Youth Crime: a Review of the Literature, TIER Working Paper 13/06.
- Sviatschi, M. M, 2018. Making a narco: Childhood exposure to illegal labor markets and criminal life paths, Manuscript, Department of Economics, Princeton University. [https://rpd.princeton.edu/pdf/files/sviatschi\\_making\\_a\\_narco\\_march2018.pdf](https://rpd.princeton.edu/pdf/files/sviatschi_making_a_narco_march2018.pdf)
- Tabellini, G. (2008). The scope of cooperation: values and incentives. *Quarterly Journal of Economics*, 123(3), pp. 905-950.
- Tabellini, G. (2010). Culture and institutions: economic development in the regions of Europe. *Journal of the European Economic Association*, 8(4), pp. 677-716.
- Transcrime (2012). *Imprese vittime di criminalità in Italia*. Trento: Università degli Studi di Trento. Available at: [www.transcrime.cs.unitn.it/tc/1075.php](http://www.transcrime.cs.unitn.it/tc/1075.php).
- Transcrime. (2013). *Progetto PON Sicurezza 2007-2013: Gli investimenti delle mafie*. Rapporto Linea 1. Milano: Ministero dell'Interno.

- Van Dijk, J. (2007). Mafia markers: assessing organised crime and its impact upon societies. *Trends in Organised Crime*, 10(4): 39-56.
- Van Duyne, P.C., & van Dijck, M. (2007). Assessing organised crime: the sad state of an impossible art. In: F. Bovenkerk, and M. Levi (eds.). *The organised crime community: essays in honour of Alan A. Block*, New York: Springer, pp.101-124.
- Von Lampe, K. (2015). "Definitions of organised crime". Retrieved from: [https://www.ceic.gouv.qc.ca/fileadmin/Fichiers\\_client/centre\\_documentaire/C EIC-R-3538.pdf](https://www.ceic.gouv.qc.ca/fileadmin/Fichiers_client/centre_documentaire/C EIC-R-3538.pdf) (last access: 29 January 2018).

## TABLES AND FIGURES

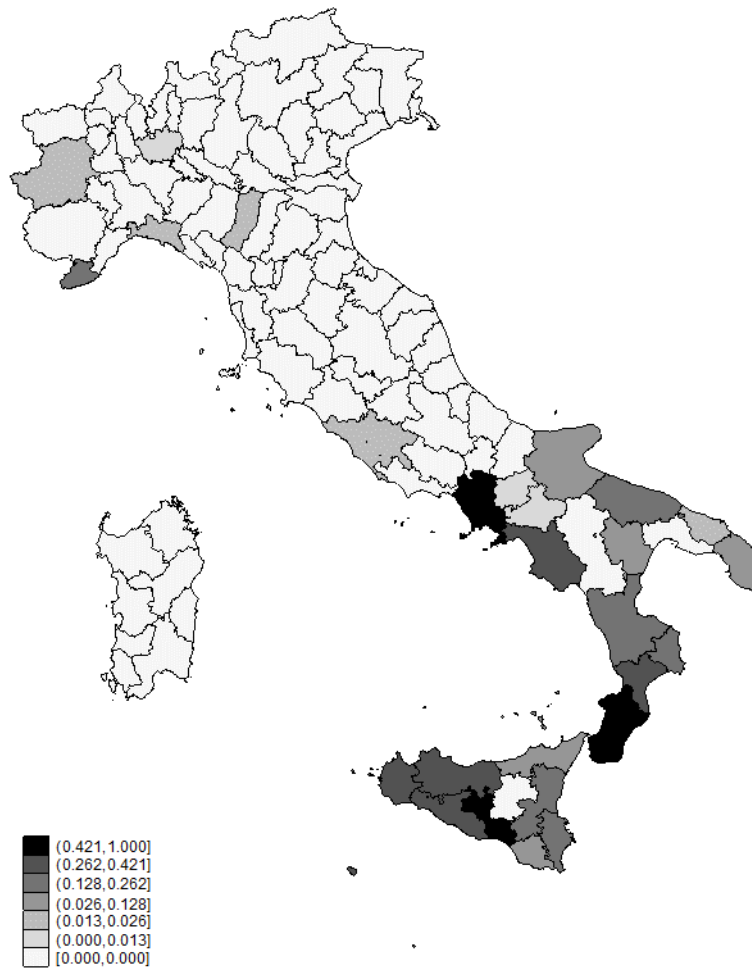
**Figure 1** - IPM index at provincial level



*Source:* our elaboration on data provided by Transcrime (2013).

*Note:* Map displays the index IPM (*Indice di Presenza Mafiosa*) of the infiltration of *Mafia*-type organisations at provincial level in Italy computed by Transcrime (2013). The index is normalised to the scale 0-1 according to a min-max criterion, where 1 = the highest value in the 107 provinces.

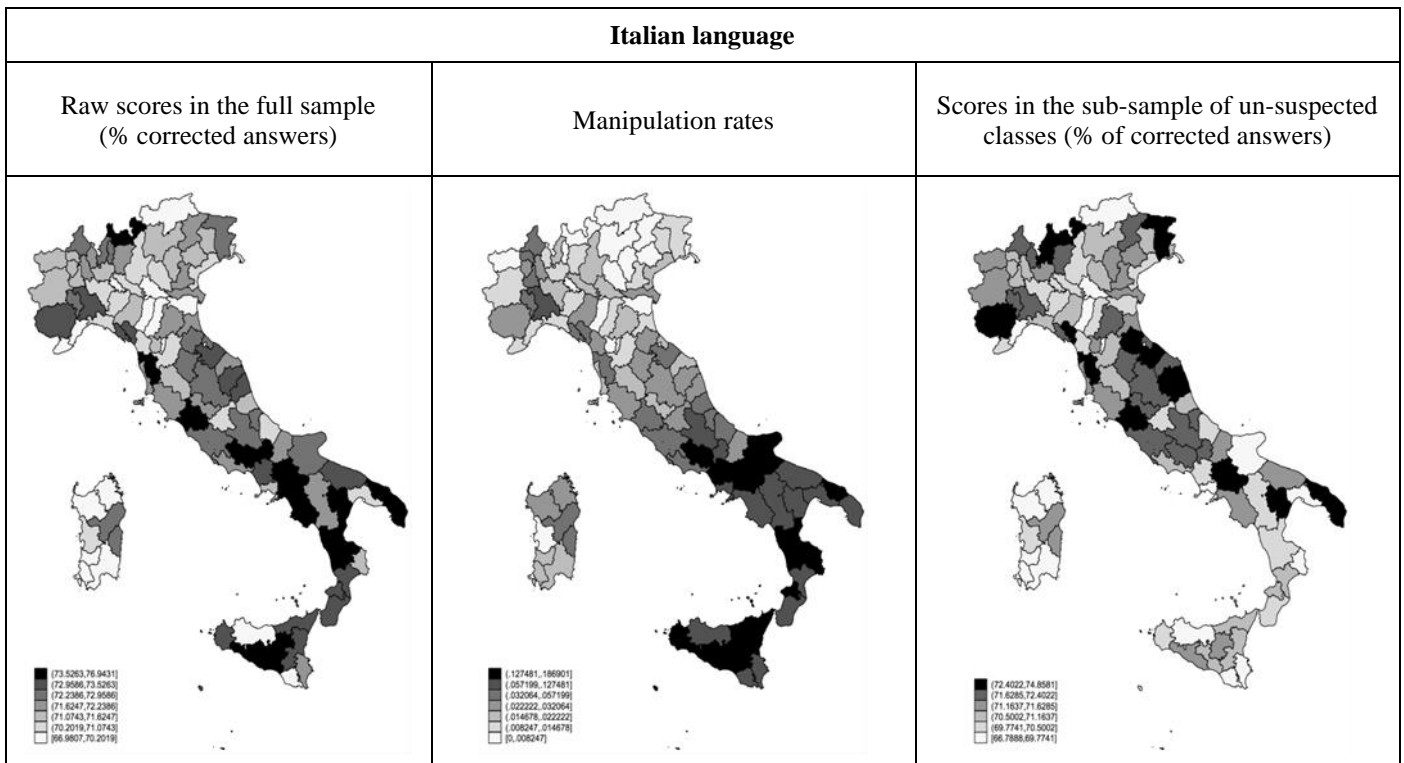
**Figure 2 - IIM index at provincial level**



*Source:* our elaboration on data provided by Italian Ministry of Internal Affairs.

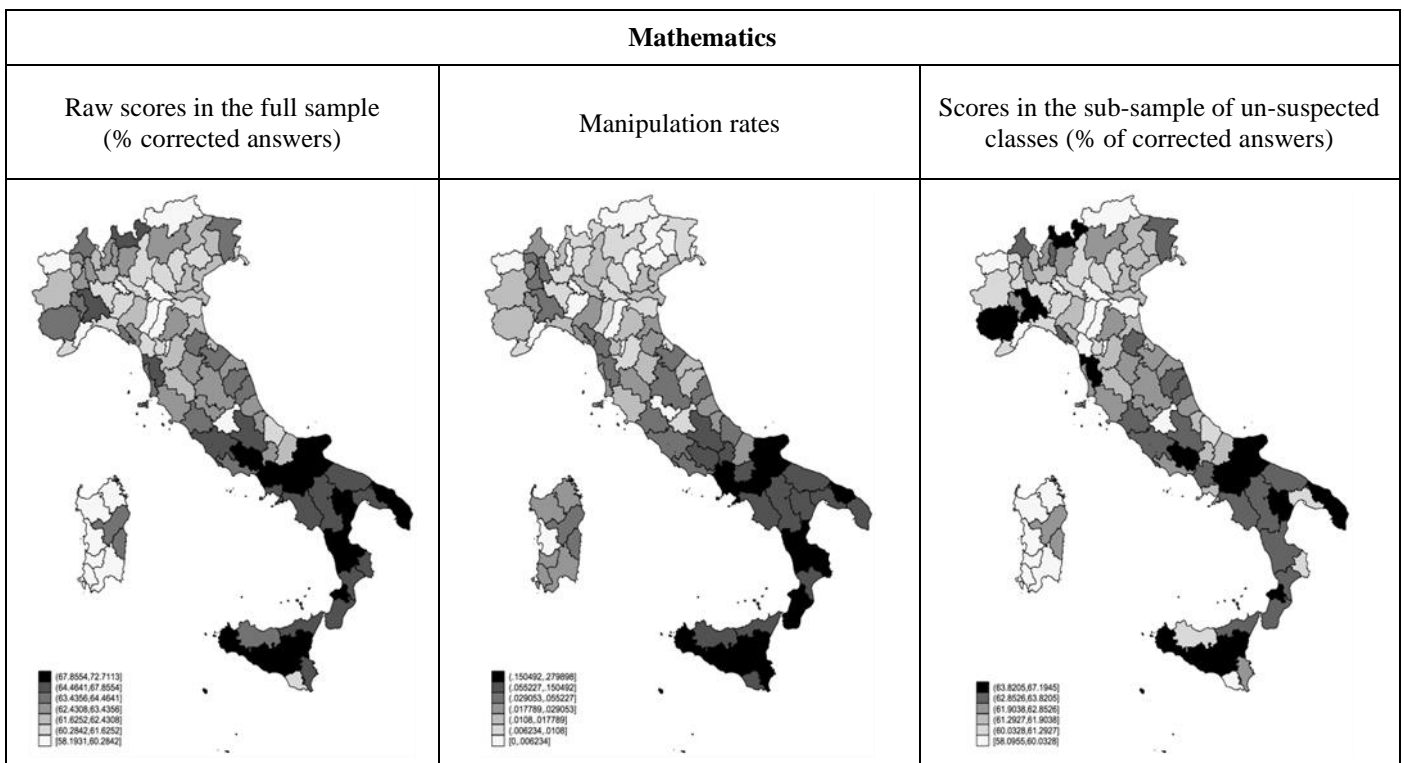
*Note:* Map displays the index of *Mafia* institutional infiltration (IIM) computed as the ratio between the population of municipalities whose councils have been dismissed because of mafia infiltration (Legislative Decree n. 164/91) from 1991 to the 31st of July 2017 and the total population of the province. The index is normalised to the scale 0-1 according to a min-max criterion, where 1 = the highest value in the 107 provinces.

**Figure 3** – Class scores in Italian language and manipulation rates (at provincial level)



Source: our elaboration on data provided by Angrist *et al.* (2017) by pooling second and fifth grade students for the school years 2009-2011.

**Figure 4** - Class scores in Mathematics and manipulation rates (at provincial level)



Source: our elaboration on data provided by Angrist *et al.* (2017) by pooling second and fifth grade students for the school years 2009-2011.

**Table 1 - Descriptive statistics of crime indices by geographical macro-area**

Variable	Obs.	Mean	Std. Dev.	Min	Max
<b>Italy</b>					
IIM	140,010	0.1071	0.2081	0.0000	1.0000
IPM	140,010	0.6273	0.2903	0.0094	1.0000
<b>North</b>					
IIM	61,938	0.0049	0.0204	0.0000	0.2427
IPM	61,938	0.4708	0.2589	0.0094	0.8585
<b>Centre</b>					
IIM	25,560	0.0058	0.0081	0.0000	0.0172
IPM	25,560	0.6171	0.2462	0.1132	0.8868
<b>South</b>					
IIM	52,512	0.2769	0.2623	0.0000	1.0000
IPM	52,512	0.8167	0.2266	0.0189	1.0000

Source: our elaboration on data provided by the Italian Ministry of Internal Affairs and Transcrime (2013).

**Table 2 - Descriptive statistics of educational outcomes and score manipulation index by geographical macro-area**

Variable	Obs.	Mean	Std. Dev.	Min	Max
<b>Italy</b>					
<b>North</b>					
ANSWERS_ITAL_PCT	61,938	71.4879	8.4254	24.8869	100.0000
ANSWERS_ITAL_STD	61,938	-0.0226	0.4074	-2.0187	1.7214
CHEAT_ITAL	61,936	0.0148	0.1206	0.0000	1.0000
<b>Centre</b>					
ANSWERS_ITAL_PCT	25,560	72.5024	9.3896	26.9231	100.0000
ANSWERS_ITAL_STD	25,560	0.0386	0.4726	-2.2027	1.8039
CHEAT_ITAL	25,557	0.0416	0.1996	0.0000	1.0000
<b>South</b>					
ANSWERS_ITAL_PCT	52,512	72.5653	12.1865	20.9790	100.0000
ANSWERS_ITAL_STD	52,512	0.0346	0.6491	-3.0167	1.7865
CHEAT_ITAL	52,510	0.1099	0.3128	0.0000	1.0000

Source: our elaboration on data provided by Angrist *et al.* (2017) by pooling second and fifth grade students for the school years 2009-2011.



**Table 3** - Educational outcome in the sub-sample of not “suspicious” classes by geographical macro-area

Variable	Obs.	Mean	Std. Dev.	Min	Max
<b>Italy</b>					
ANSWERS_ITAL_PCT	132,256	70.9396	9.2336	20.9790	99.2823
ANSWERS_ITAL_STD	132,256	-0.0545	0.4593	-3.0167	1.7447
<b>North</b>					
ANSWERS_ITAL_PCT	24,495	71.6777	8.6624	26.9231	96.7213
ANSWERS_ITAL_STD	24,495	-0.0079	0.4231	-2.2027	1.5864
<b>Centre</b>					
ANSWERS_ITAL_PCT	61,022	71.1982	8.1373	24.8869	97.8395
ANSWERS_ITAL_STD	61,022	-0.0389	0.3871	-2.0187	1.5077
<b>South</b>					
ANSWERS_ITAL_PCT	46,739	70.2151	10.7044	20.9790	99.2823
ANSWERS_ITAL_STD	46,739	-0.0992	0.5514	-3.0167	1.7447

*Source:* our elaboration on data provided by Angrist *et al.* (2017) by pooling second and fifth grade students for the school years 2009-2011.

**Table 4 - Summary statistics**

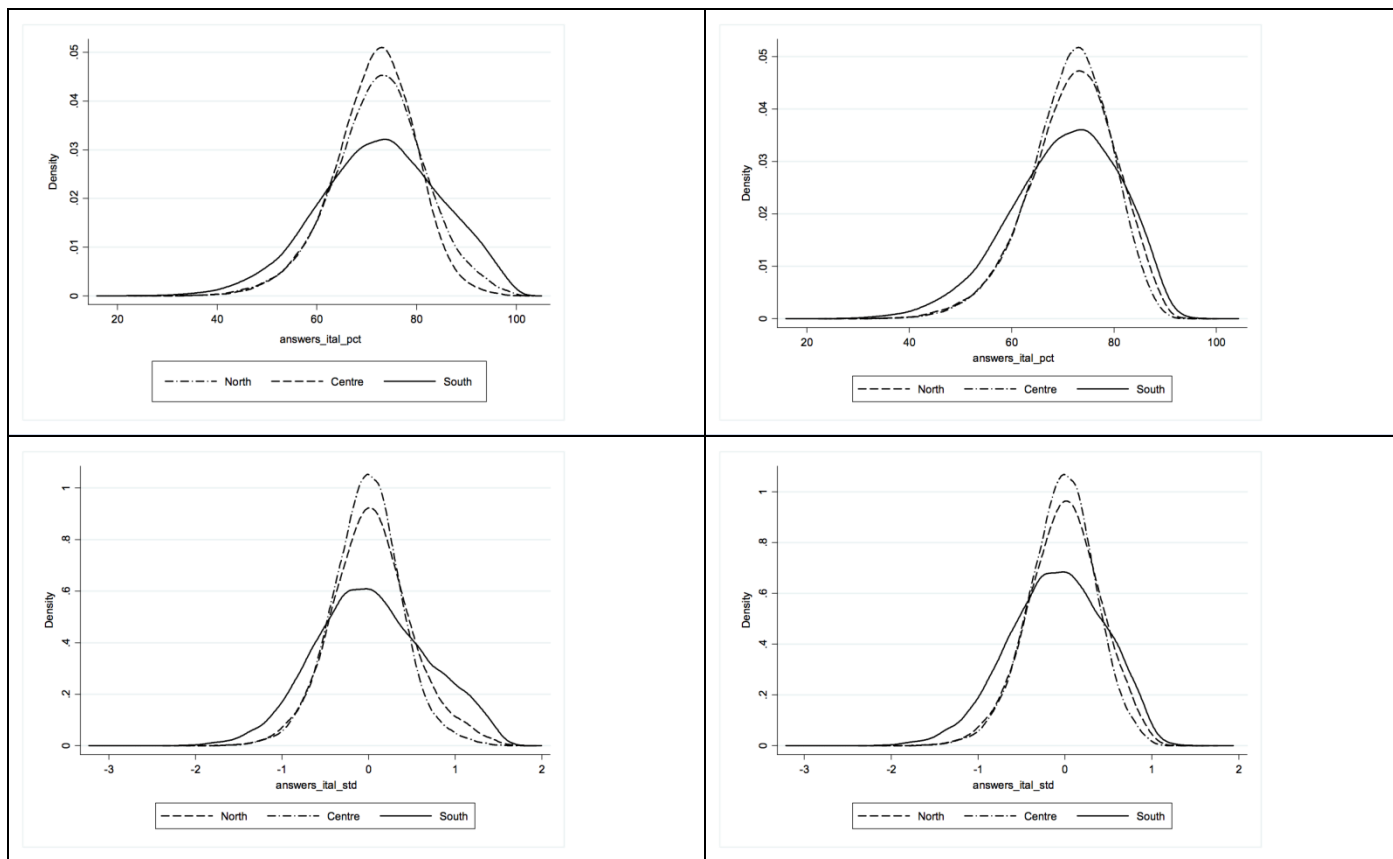
Variables	Meaning	Mean	St. Dev.
<i>Outcomes</i>			
ANSWERS_ITAL_PCT	Class percentage of corrected answers in Italian language	72.0772	10.1722
ANSWERS_ITAL_STD	Standardised class percentage of corrected answers in Italian language	0.0100	0.5226
<i>Controls for geographical area</i>			
NORTH	Dummy for the class in the North of the country	0.4424	0.4967
CENTRE	Dummy for the class in the Centre of the country	0.1826	0.3863
SOUTH	Dummy for the class in the South of the country	0.3751	0.4841
<i>Controls for organised crime at provincial level</i>			
IIM	Index of <i>mafia</i> institutional infiltration	0.1071	0.2081
IPM <sup>a</sup>	Index of infiltration of <i>Mafia</i> -type organisations	0.6273	0.2903
RT <sup>a</sup>	Composite index of territorial risk of organised crime infiltration	0.4097	0.2555
IR <sup>a</sup>	Index of wealth deprivation	0.5132	0.2694
IS <sup>a</sup>	Index of the diffusion of shadow economy	0.3895	0.2084
IT <sup>a</sup>	Index of lack of market openness and technological adoption	0.4629	0.2400
HM_regions	Dummy for regions with historical presence of <i>Mafia</i>	0.0934	0.2909
<i>Instruments</i>			
LANDPROD <sup>b</sup>	Land productivity at the end of XIX century	752.8110	917.2359
MAFIA_HIST <sup>b</sup>	Diffusion of <i>Mafia</i> organizations at the end of XIX century	1.4521	1.5955
<i>Other controls at class level</i>			
STUDENTS	Grade enrolment at school	53.2908	30.5651
CLSIZE	Class size	19.8848	3.5765
CHEAT_ITAL	Dummy for class with compromised scores (“suspicious” class)	0.0553	0.2286
MONITORED	Class externally monitored	0.0688	0.2531
FEMALE	Proportion of female students	0.4820	0.1205
M_FEMALE	Proportion of students with missing gender	0.0208	0.1085
IMMIGRANTS_BROAD	Proportion of immigrant students	0.0981	0.1198
M_ORIGIN	Proportion of students with missing origin	0.0296	0.1544
DAD_LOWEDU	Proportion of students whose father is High School dropout	0.4273	0.2561
DAD_MIDEDU	Proportion of students whose father is High School graduate	0.2546	0.1678
DAD_HIGHEDU	Proportion of students whose father is College graduate	0.0994	0.1168
M_DAD_EDU	Proportion of students with missing father education	0.2187	0.3361
MOM_UNEMP	Proportion of students whose mother is unemployed	0.0366	0.0552
MOM_HOUSEW	Proportion of students whose mother is a housewife	0.3206	0.2348
MOM_EMPLOYED	Proportion of students whose mother is employed	0.4503	0.2658
M_MOM_OCC	Proportion of students with missing mother occupation	0.1925	0.3239
M_MOM_EDU	Proportion of students with missing mother education	0.2051	0.3399

<sup>a</sup> Our computation on data provided by Transcrime (2013)

<sup>b</sup> Our computation on data provided by Del Monte and Pennacchio (2012) and Albanese and Marinelli (2013).

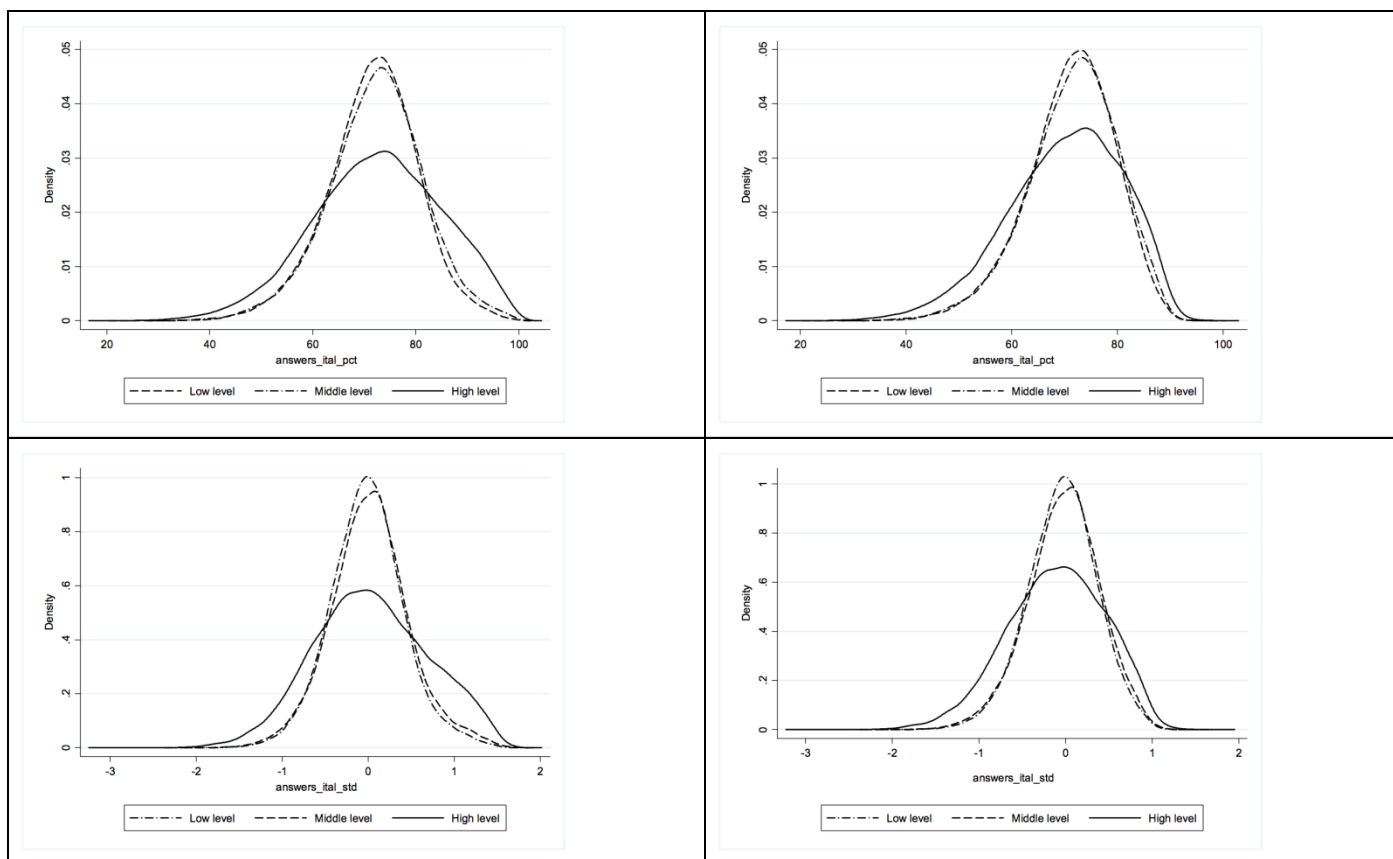
*Note:* Stratification controls also for survey year, grade and region.

**Figure 5** - Kernel density estimates of unstandardized (top) and standardised (bottom) class scores in Italian languages by geographical macro-area (left: without correction for score manipulation; right: with correction for score manipulation)



Source: our elaboration on data provided by Angrist *et al.* (2017) by pooling second and fifth grade students for the school years 2009-2011.

**Figure 6** - Kernel density estimates of unstandardized (top) and standardised (bottom) class scores Italian languages by level of IIM (left: without correction for score manipulation; right: with correction for score manipulation)



Source: our elaboration on data provided by Angrist *et al.* (2017) by pooling second and fifth grade students for the school years 2009-2011.

**Table 5 - OLS results – full sample (dependent variable: standardised Italian language score)**

Variables	(1)	(2)	(3)	(4)	(5)
NORTH	0.1063***		0.1071***		0.1260***
	(0.0157)		(0.0156)		(0.0160)
CENTRE	0.0732***		0.0737***		0.0671***
	(0.0117)		(0.0117)		(0.0119)
IIM		-0.1013***	-0.1013***		
		(0.0187)	(0.0187)		
IPM				-0.0956***	-0.0956***
				(0.0090)	(0.0090)
Full set of other controls	yes	yes	yes	yes	yes
Control for grade	yes	yes	yes	yes	yes
Control for survey year	yes	yes	yes	yes	yes
Control for region	yes	yes	yes	yes	yes
Observations	140,003	140,003	140,003	140,003	140,003
R-squared	0.3313	0.3317	0.3317	0.3322	0.3322

Source: our elaboration on data provided by Ministry of Internal Affairs, Transcrime (2013) and Angrist *et al.* (2017) by pooling second and fifth grade students for the school years 2009-2011.

Note: The table reports OLS estimates. The unit of observation is an individual. The models in the Table also controls for a full set of variables at SCHOOL and CLASS level. Fixed effects for GRADE, SURVEY\_YEAR and REGION are also included. Robust standard errors, clustered on school and grade, are shown in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 6 - OLS results – sample collapsed at provincial level (dependent variable: standardised Italian language score)**

Variables	(1)	(2)	(3)	(4)	(5)
NORTH	0.0872***		0.0884***		0.0892***
	(0.0162)		(0.0164)		(0.0166)
CENTRE	0.0639***		0.0735***		0.0693***
	(0.0190)		(0.0192)		(0.0200)
IIM		-0.0929***	-0.0905***		
		(0.0279)	(0.0286)		
IPM				-0.0200	-0.0233
				(0.0188)	(0.0190)
Full set of other controls	yes	yes	yes	yes	yes
Control for grade	yes	yes	yes	yes	yes
Control for survey year	yes	yes	yes	yes	yes
Observations	618	618	618	618	618
R-squared	0.5332	0.5183	0.5418	0.5097	0.5344

Source: our elaboration on data provided by Ministry of Internal Affairs, Transcrime (2013) and Angrist *et al.* (2017) by pooling second and fifth grade students for the school years 2009-2011.

Note: The table reports OLS estimates. The unit of observation is at provincial level. The models in the Table also controls for a full set of variables at SCHOOL and CLASS level. Fixed effects for GRADE and SURVEY\_YEAR are also included. Robust standard errors are shown in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 7** - OLS results – control for regions with historical presence of organised crime - full sample (dependent variable: standardised Italian language score)

Variables	(1)	(2)
HM_regions	-0.0254* (0.0140)	0.0051 (0.0145)
IIM	-0.1013*** (0.0187)	
IPM		-0.0956*** (0.0090)
Full set of other controls	yes	yes
Control for grade	yes	yes
Control for survey year	yes	yes
Control for region	yes	yes
Observations	140,003	140,003
R-squared	0.3317	0.3322

Source: our elaboration on data provided by Ministry of Interior, Transcrime (2013) and Angrist *et al.* (2017) by pooling second and fifth grade students for the school years 2009-2011.

Note: The table reports OLS estimates. The unit of observation is an individual. The models in the Table also controls for a full set of variables at SCHOOL and CLASS level. Fixed effects for grade, survey year and region are also included. Robust standard errors, clustered on school and grade, are shown in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 8** - OLS results – subsample of classes in traditionally *Mafia*-controlled regions (Campania, Calabria and Sicily in columns (1) and (2); also Apulia in columns (3) and (4)); all southern regions in columns (5) and (6) (dependent variable: standardised Italian language score)

Variables	(1)	(2)	(3)	(4)	(5)	(6)
IIM	-0.0936*** (0.0192)		-0.0869*** (0.0188)		-0.0912*** (0.0187)	
IPM		-0.3423*** (0.0537)		-0.3350*** (0.0487)		-0.1499*** (0.0302)
Full set of other controls	yes	yes	yes	yes	yes	yes
Control for grade	yes	yes	yes	yes	yes	yes
Control for survey year	yes	yes	yes	yes	yes	yes
Control for region	yes	yes	yes	yes	yes	yes
Observations	33,286	33,286	44,130	44,130	52,510	52,510
R-squared	0.4026	0.4032	0.4001	0.4009	0.3886	0.3886

Source: our elaboration on data provided by Ministry of Interior, Transcrime (2013) and Angrist *et al.* (2017) by pooling second and fifth grade students for the school years 2009-2011.

Note: The table reports OLS estimates. The unit of observation is an individual. The models in the Table also controls for a full set of variables at SCHOOL and CLASS level. Fixed effects for grade, survey year and region are also included. Robust standard errors, clustered on school and grade, are shown in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 9** – Pairwise correlation between organised crime index and historical instruments

		(1)	(2)	(3)	(4)
(1)	IIM	1.0000			
(2)	IPM	0.6215*	1.0000		
(3)	LANDPROD	0.6695*	0.3987*	1.0000	
(4)	MAFIA_HIST	0.7886*	0.4630*	0.6884*	1.0000

Source: our elaboration on data provided by Ministry of Internal Affairs, Transcrime (2013), Del Monte and Pennacchio (2012) and Albanese and Marinelli (2013).

Note: \* p<0.01

**Table 10** - IV estimates – all southern regions (Instrumented: IIM and IPM - dependent variable: standardised Italian language score)

Variables	(1)	(2)	(3)	(4)	(5)	(6)
IIM	-0.3255*** (0.0399)	-0.2382*** (0.0283)	-0.2481*** (0.0282)			
IPM				-0.7206*** (0.0880)	-0.6573*** (0.0788)	-0.6824*** (0.0746)
<i>Instruments</i>						
LANDPROD <sup>a</sup>	-0.0002*** (0.0000)		-0.0000*** (0.0000)	-0.0001*** (0.0000)		-0.0000*** (0.0000)
MAFIA_HIST <sup>a</sup>		-0.1094*** (0.0014)	-0.0982*** (0.0022)		-0.0396*** (0.0007)	-0.0271*** (0.0011)
Endogeneity test <sup>b</sup>	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
F-statistics <sup>c</sup>	16.42	13.99	10.78	14.04	15.88	11.27
Overidentification test <sup>d</sup>			0.1749			0.3832
Full set of other controls	yes	yes	yes	yes	yes	yes
Control for grade	yes	yes	yes	yes	yes	yes
Control for survey year	yes	yes	yes	yes	yes	yes
Control for region	yes	yes	yes	yes	yes	yes
Observations	52,510	52,510	52,510	52,510	52,510	52,510

<sup>a</sup> First-stage coefficients of instruments.

<sup>b</sup> Reports the p-value of the endogeneity test.

<sup>c</sup> F-statistic of the Kleibergen-Paap rk Wald test for weak identification.

<sup>d</sup> The overidentification test reports the p-value of the Hansen's J statistic of overidentifying restrictions.

Source: our elaboration on data provided by Ministry of Interior, Transcrime (2013) and Angrist et al. (2017) by pooling second and fifth grade students for the school years 2009-2011.

Note: The Table reports IV estimates. The models in the Table also controls for a full set of variables at SCHOOL and CLASS level. Fixed effects for GRADE, SURVEY\_YEAR and REGION are also included. Robust standard errors, clustered on school and grade, are shown in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 11** – Confidence bounds relaxing the exogeneity assumption of the instrument: IV estimates – all southern regions (Instrumented: IIM and IPM - dependent variable: standardised Italian language score)

Range of plausible values		IIM		IPM	
LANDPROD		Lower	Upper	Lower	Upper
[-0.0001, +0.0001]		-0.4102	-0.2409	-0.9073	-0.5338
[-0.0002, +0.0002]		-0.4166	-0.2345	-0.9217	-0.5195
[-0.0005, +0.0005]		-0.4360	-0.2152	-0.9646	-0.4767
MAFIA_HIST		Lower	Upper	Lower	Upper
[-0.1, +0.1]		-0.3853	-0.0915	-0.9385	-0.3775
[-0.2, +0.2]		-0.4312	-0.0459	-1.0654	-0.2519
[-0.5, +0.5]		-0.4771	-0.0002	-1.2778	-0.1010
LANDPROD	MAFIA_HIST	Lower	Upper	Lower	Upper
[-0.0001, +0.0001]	[-0.1, +0.1]	-0.3160	-0.1836	-0.8619	-0.5031
[-0.0002, +0.0002]	[-0.2, +0.2]	-0.3558	-0.1483	-0.9624	-0.4032
[-0.0005, +0.0005]	[-0.5, +0.5]	-0.4117	-0.0852	-1.1072	-0.2590

Source: our elaboration on data provided by Ministry of Interior, Transcrime (2013) and Angrist et al. (2017) by pooling second and fifth grade students for the school years 2009-2011.

Notes: 95% confidence lower and upper bounds are estimated according to the approach proposed by Conley et al. (2012). All estimates include the entire set of control as in Table 11.

## APPENDIX A

### A.1 A snapshot of the education system

In Italy education is a Constitutional right/duty (art. 33 and 34), publicly provided free of charge and compulsory for ten years between the ages of 6 and 16. More specifically, compulsory education covers eight years of the first cycle of education (i.e. five years of primary school and three years of lower secondary school), and the first two years of the second cycle. After completion of the first cycle of education, students can undertake the final two years of compulsory education (from 14 to 16 years of age) at an upper secondary school (*lyceum*, technical or vocational institutes), which lasts either five or four or three years. Alternatively, 15-year-olds can also spend the last year of compulsory education by attending an apprenticeship, organised by the Regions upon specific arrangements with the Ministry of Labour and the Ministry of Education, University and Research (MIUR).

Compulsory education can be undertaken at either public schools (free of charge), private but publicly subsidised schools (*scuole paritarie*), or, subject to certain conditions, at other private schools or even through home education. Parents are entitled to choose any school they want (even within the public sector) and, except for the rare case of oversubscription, they usually get it. Notwithstanding, families generally apply to the nearest school to their home, in the province of residence. At the end of upper secondary education, students sit a State examination. Only if they pass, they are allowed to access to tertiary education (university, high level arts and music education and higher technical institutes).

Following the 2001 Constitutional reform (Constitutional Act n.3/2001), legislative competencies on education issues and responsibilities for the organisation and administration of the education system are shared between the central and the regional level. At the central level, the MIUR is responsible for the organisation/programming/management of the overall school system (e.g. minimum national curriculum and quality standards) and its evaluation, the adherence to the European requirements, the determination and assignment of financial resources, and the recruitment of school staff. The MIUR operates at a decentralised level through the Regional School Offices (*Uffici Scolastici Regionali*, USR) and their provincial branches (*Uffici Scolastici Provinciali*, USP). The USR monitor compliance with the national education provisions and standards, as well as with the minimum performance requirements. At local level, the USP have no autonomy; they support schools with their administrative tasks and the planning of their own educational offers. Furthermore, once the USR have determined the total number of classes in the region, the USP decide their allocation between schools. Municipalities are mainly in charge of providing school buildings, equipment and facilities, as well as transport service; their assigned specific functions and responsibilities change according to the different levels of education.



Since 2000, schools at primary and secondary level have been granted legal status and autonomy in the fields of administration, education, teaching and research, though within the general framework laid down by the Ministry of Education to ensure uniformity throughout the Italian territory. Schools can be organised either as single level institutes or as “comprehensive” institutes that include multiple grades (usually, primary and lower secondary schools). According to the type of organisational structure and to the number of classes, Italian schools differ greatly in size.

All schools are run by a school manager (*Dirigente scolastico*), who is the legal representative of the institution and is responsible for its overall organisation, the management of human, technological and financial resources and the quality of the service provided. Teachers are mainly civil servants, recruited by competitive examinations and with virtually no risk of dismissal. They are paid a fixed salary, which remains one of the lowest among the OECD countries (OECD, 2016). Recently, in an attempt to incentivise teacher performance, the so-called “*Buona scuola*” reform (Law 107/2015) has introduced a merit-based component in the teachers’ salary, in the form of a one-off bonus. This is assigned yearly by an internal school committee to the best performing teachers, based on national guidelines for teachers’ evaluation. However, its limited entity is likely to have only a negligible impact on teachers’ motivation, which is found to be a relevant factor affecting pupils’ achievements in Italy (Barbieri *et al.*, 2017).

In an attempt to boost competition in the public education sector (both among schools and among teachers), the recent law 107/2015 has empowered the school principals to choose teachers to hire from territorial registers formed by networks of schools, taking into account teachers’ applications. Once appointed, teachers are reconfirmed every three years, in accordance with the Educational Syllabus (*Piano dell’Offerta Formativa*) of the school.

## A.2 Robustness checks on the whole Italian sample

In this Section we provide some robustness checks on the whole Italian sample. First, the consistency of our results across different *Mafia* measures reduces concerns for measurement errors. However, after having controlled for several determinants of educational outcomes and for regional fixed effects, the choice of a simple linear regression with an organised crime index as dependent variable may be subjected to picking up spurious correlations.

For, we provide several robustness assessments. First, to check for measurement errors of our results with respect to different and wider measures of organised crime and to control for environmental factors at provincial level that potentially affect the expected returns of the investment in human capital and thus the educational outcomes, we apply the composite index of territorial risk of organised crime infiltration at provincial level (*Indice di Rischio Territoriale*, RT) developed by Transcrime (2013) and its four pillars (one of which is the IPM index)<sup>26</sup>.

Although we control for regional fixed effects, adding these variables enables us to further control for environmental factors at provincial level that potentially affect the expected returns of the investment in human capital and thus the educational outcomes.

Looking at Table A.2.1, the coefficients of the RT index are never statistically significant, neither when the index is employed alone (column 1), nor when is used jointly with the macro-area variables (column 2), nor when also the IIM index is added to the specification (column 3).

Furthermore, both the macro-area variables and the IIM index continue to be highly significant (at the 1% level) and with the expected signs. Indeed, compared to our IIM index, the RT index provides a quite different picture of the organised crime phenomenon: it captures the potential risk of organised crime infiltration rather than the effective infiltration of *Mafia*-type organisations into local institutions as it appears from voting behaviours. However, only the latter phenomenon seems to be closely related to INVALSI class achievements (i.e., our dependent variable). In the last two columns of Table A.2.1 the effects of the four pillars are considered, without (column 4) and along with the macro-area variables (column (5)). Though the coefficients of the four pillars are always significant, not all of them have the expected sign. However, the

---

<sup>26</sup> The RT index is computed by aggregating the IPM and other three (normalised) indexes (pillars), using the generalised average method. The first pillar of the RT index is a measure of wealth deprivation (*Indice di Ricchezza*, IR) built considering per capita GDP, number of expensive cars (over 2000 cc.) purchased and average declared income used to compute Regional and Municipal surtax. The second pillar (*Indice di Economia Sommersa*, IS) captures the diffusion of shadow economy and is computed using the following simple indexes: levels of tax evasion, tax gap, and the number of notifications of suspicious transactions recorded by Bank of Italy. The last pillar (*Indice di Struttura del Mercato nel Territorio*, IT) is a measure of lack of market openness and technological adoption at provincial level. It has been obtained considering the expenditure in R&D, the endowment of infrastructure at provincial level and the ratio between the sum of imports and exports and the added value at provincial level. All the indexes are normalized to the scale 0-1 according to a min-max criterion, where 1 is equal to the highest value in the 107 provinces. The descriptive statistics of these variables are reported in Table 4. The distributions of the RT and its pillars' values are provided in the Appendix B.

IPM index continues to be significant, with a corrected sign and a beta coefficient slightly lower than that estimated in Table 5.

**Table A.2.1** - Robustness check for OLS estimates using the RT index (columns (1) to (3)) and its pillars (columns (4) and (5)) - full sample (dependent variable: standardised Italian language score)

Variables	(1)	(2)	(3)	(4)	(5)
NORTH		0.1072*** (0.0157)	0.1074*** (0.0157)		0.0693*** (0.0164)
CENTRE		0.0657*** (0.0128)	0.0708*** (0.0128)		0.0112 (0.0139)
RT	-0.0265 (0.0188)	-0.0265 (0.0188)	-0.0100 (0.0190)		
IIM			-0.1002*** (0.0189)		
IPM				-0.0290*** (0.0111)	-0.0290*** (0.0111)
IR				-0.2270*** (0.0231)	-0.2270*** (0.0231)
IS				0.2109*** (0.0187)	0.2109*** (0.0187)
IT				0.0785*** (0.0137)	0.0785*** (0.0137)
Full set of other controls	yes	yes	yes	yes	yes
Control for grade	yes	yes	yes	yes	yes
Control for survey year	yes	yes	yes	yes	yes
Control for region	yes	yes	yes	yes	yes
Observations	140,003	140,003	140,003	140,003	140,003
R-squared	0.3313	0.3313	0.3317	0.3346	0.3346

Source: our elaboration on data provided by Ministry of Interior, Transcrime (2013) and Angrist *et al.* (2017) by pooling second and fifth grade students for the school years 2009-2011.

Note: The table reports OLS estimates. The unit of observation is an individual. The models in the Table also controls for a full set of variables at SCHOOL and CLASS level. Fixed effects for GRADE, SURVEY\_YEAR and REGION are also included. Robust standard errors, clustered on school and grade, are shown in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Second, we re-run our regressions by excluding provinces where IIM is different from zero. In this subsample there are those provinces with at least one municipality whose council has been dismissed because of *Mafia* infiltration (Legislative Decree n. 164/91) from 1991 to the 31st of July 2017. Table A.2.2 reports the results of this test. It shows that both the coefficient of IPM and IIM are still highly significant with the expected sign.

In all robustness checks reported in this section, the estimated coefficient of the effect of institutional penetration of organised crime on the educational outcome is remarkable stable across specifications. One standard deviation increases in the IIM index is associated with a reduction in educational outcome of ranging between 0.0209 to 0.0224. Conversely, IPC index shows larger variation. In fact, one standard deviation increases in the IPC index is associated with a reduction in educational outcome ranging between 0.0084 to 0.1258.

**Table A.2.2** - Subsample of classes in provinces with IIM index different from zero (dependent variable: standardised Italian language score)

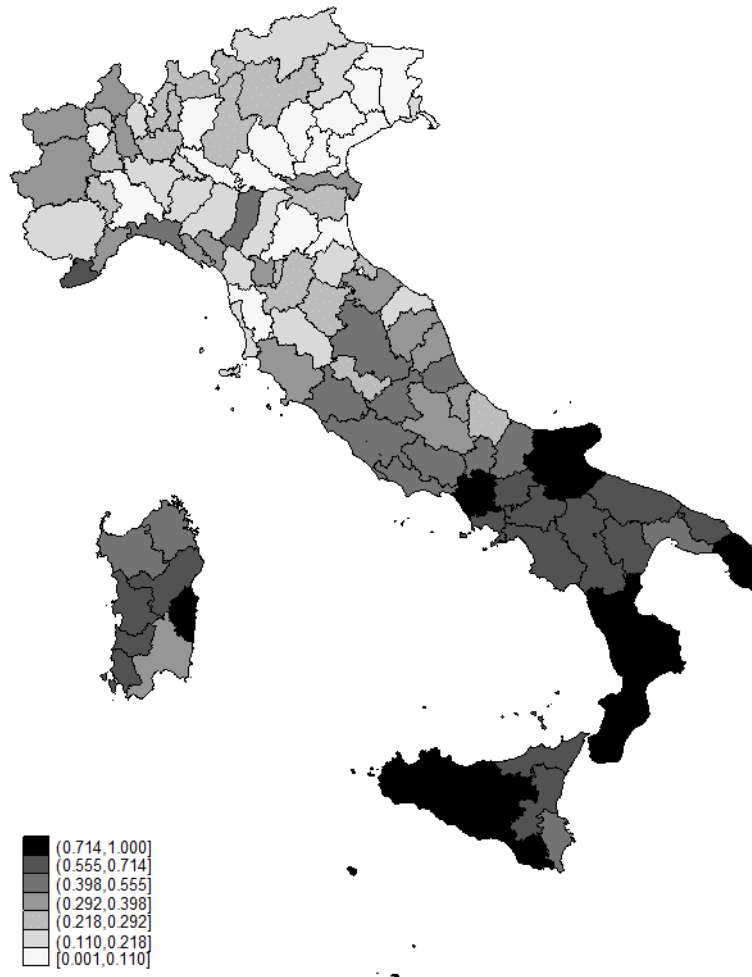
Variables	(1)	(2)	(3)
RT	0.1071* (0.0621)		
IIM		-0.1077*** (0.0189)	
IPM			-0.4333*** (0.0490)
Full set of other controls	yes	yes	yes
Control for grade	yes	yes	yes
Control for survey year	yes	yes	yes
Control for region	yes	yes	yes
Observations	68,039	68,039	68,039
R-squared	0.3753	0.3760	0.3771

*Source.* our elaboration on data provided by Ministry of Interior, Transcrime (2013) and Angrist et al. (2017) by pooling second and fifth grade students for the school years 2009-2011.

*Note:* The table reports OLS estimates. The unit of observation is an individual. The models in the Table also controls for a full set of variables at SCHOOL and CLASS level. Fixed effects for grade, survey year and region are also included. Robust standard errors, clustered on school and grade, are shown in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

## APPENDIX B

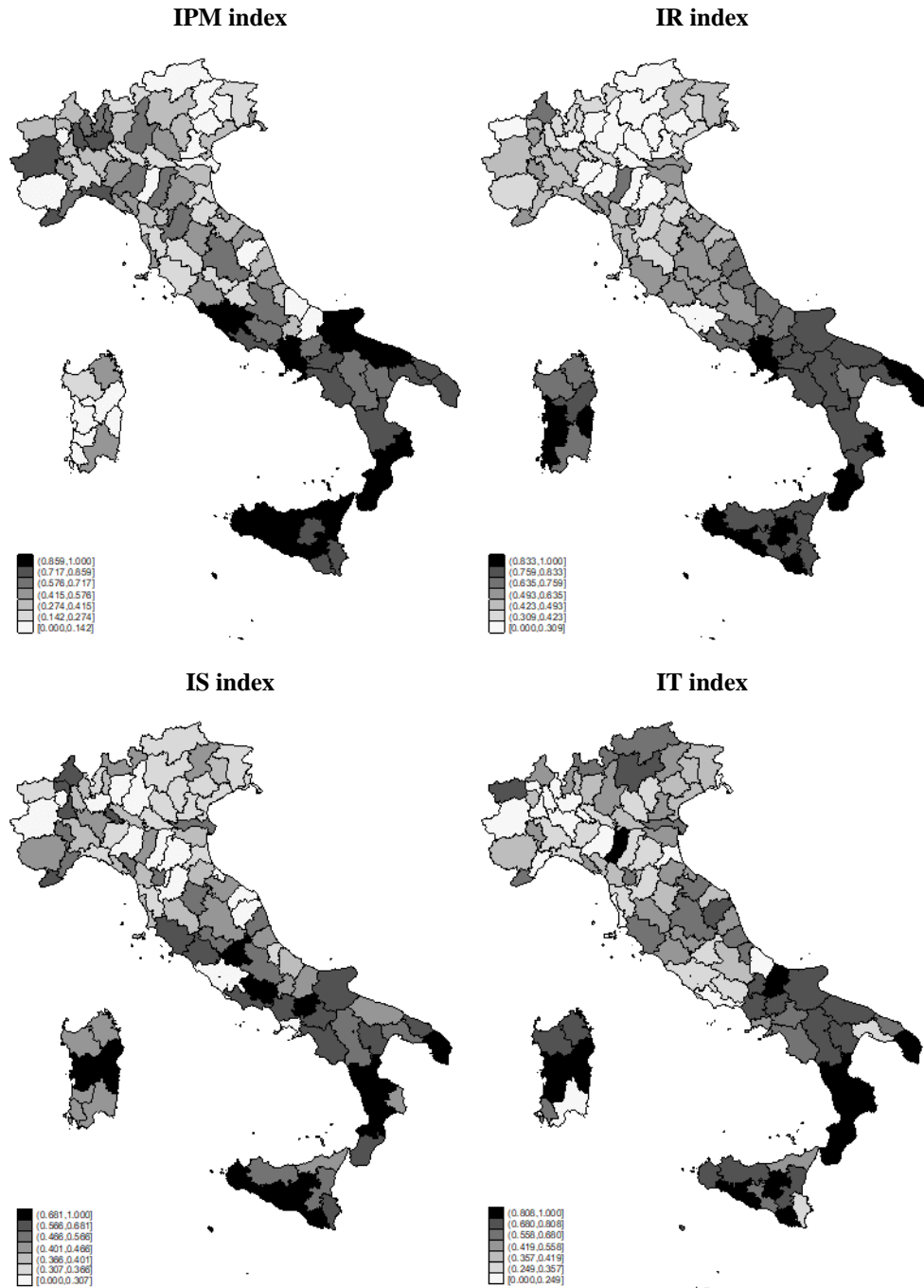
**Figure B.1** – RT synthetic index at provincial level



*Source:* our elaboration on data provided by Transcrime (2013).

*Note:* Map displays the composite index of territorial risk of organised crime infiltration at the provincial level (*Indice di Rischio Territoriale*, RT) developed by Transcrime (2013).

**Figure B.2 – Pillars of the RT synthetic index at provincial level.**



Source: our elaboration on data provided by Transcrime (2013).

Note: Maps display the four pillars of the RT composite index at provincial level. The first pillar is the IPM index of *Mafia*-type organized crime (*Indice di Presenza Mafiosa*) already reported in Figure 1. The second pillar is a measure of wealth deprivation (IR, *Indice di Ricchezza*) built considering per capita GDP, number of expensive cars (over 2000 cc.) purchased and average declared income used to compute Regional and Municipal surtax. The third pillar (IS, *Indice di Economia Sommersa*) captures the diffusion of shadow economy and is computed using the following simple indexes: levels of tax evasion, tax gap, and the number of notifications of suspicious transactions recorded by Bank of Italy. The last pillar (IT, *Indice di Struttura del Mercato nel Territorio*,) is a measure of lack of market openness and technological adoption at provincial level. It has been obtained considering the expenditure in R&D, the endowment of infrastructure at provincial level and the ratio between the sum of imports and exports and the added value at provincial level.

**Table B.1** - OLS results – dependent variable: standardised mathematics scores

Variables	(1)	(2)	(3)	(4)	(5)
<i>full sample</i>					
NORTH	0.0698*** (0.0214)		0.0701*** (0.0156)		0.0902*** (0.0217)
CENTRE	0.0123 (0.0117)		0.0125 (0.0161)		0.0060 (0.0163)
IIM		-0.0475** (0.0231)	-0.0475** (0.0231)		
IPM				-0.0991*** (0.0116)	-0.0991*** (0.0116)
Full set of other controls	yes	yes	yes	yes	yes
Control for grade	yes	yes	yes	yes	yes
Control for survey year	yes	yes	yes	yes	yes
Control for region	yes	yes	yes	yes	yes
Observations	139,996	139,996	139,996	139,996	139,996
R-squared	0.3469	0.3470	0.3470	0.3475	0.3475
<i>sample collapsed at provincial level</i>					
NORTH	0.0614*** (0.0190)		0.0642*** (0.0190)		0.0630*** (0.0192)
CENTRE	0.0513** (0.0224)		0.0629*** (0.0224)		0.0552** (0.0229)
IIM		-0.0949** (0.0370)	-0.1016*** (0.0382)		
IPM				-0.0104 (0.0227)	-0.0166 (0.0230)
Full set of other controls	yes	yes	yes	yes	yes
Control for grade	yes	yes	yes	yes	yes
Control for survey year	yes	yes	yes	yes	yes
Observations	618	618	618	618	618
R-squared	0.7423	0.7415	0.7461	0.7381	0.7425

*Source:* our elaboration on data provided by Ministry of Internal Affairs, Transcrime (2013) and Angrist *et al.* (2017) by pooling second and fifth grade students for the school years 2009-2011.

*Note:* Robust standard errors in parentheses; \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$