Shirking and social capital: evidence from skilled vs unskilled school workers*

Adriana Di Liberto
University of Cagliari
IZA and CRENoS

Marco Sideri
University of Cagliari
and CRENoS

September 8, 2017

Abstract

We study the relation between shirking behavior among workers in Italian public schools and the level of social capital across different regions. In particular, our analysis focuses on the role of regional (NUTS3 level) social capital on teachers' absenteeism rates. Exploiting a rich dataset from MIUR, we control for a number or school and local characteristics. In addition, to address likely endogeneity problems in OLS results, we instrument social capital with the presence of different foreign dominations that ruled Italian regions between the 12th and 19th century before the creation of the unified Italian State. Our results suggest that past historical institutions play a role on the current social capital level and show that the latter has a significant effect on the absenteeism rate of teachers in Italian schools.

Keywords: social capital, shirking, absenteeism, schools.
J.E.L. Classification: I20, I21, I28, J45.

PRELIMINARY DRAFT. DO NOT CITE WITHOUT PERMISSION.

*Contacts: adriana.diliberto@gmail.com, marco.sideri@crenos.unica.it.
The research leading to these results has received funding from the Autonomous Region of Sardinia's Regional Law 7/2007 (Capitale sociale e divari economici regionali, year 2012). We thank Fabiano Schivardi, Francesco Pigliaru, Giovanni Sulis and the seminar participants at the Conference "Capitale sociale e divari economici regionali", June 2016, University of Cagliari, for useful comments and suggestions.

1 Introduction

One of the elements that plays a key role in the students' outcomes at school is the teachers' behavior, including their absenteeism rates. In fact, many studies find evidence that teachers' absenteeism is usually correlated with a lower students' performance (Miller et al., 2008; Duflo et al., 2012; Herrmann and Rockoff, 2012) and try to analyze its determinants.

Absenteeism in schools may depend on various factors: the workplace environment and the managerial practices adopted by the school principal (Di Liberto et al., 2015; Bradley et al., 2007; Gaziel, 2004), school and students characteristics (Herrmann and Rockoff, 2012), local characteristics and non-monetary incentives (Chaudhury et al., 2006). However, we still have no evidence on the role played by the local social capital on teachers' absenteeism rates.

In particular, a vast literature use the sick leave absenteeism rates as a measure of the worker's shirking (Ichino and Maggi, 2000; Scoppa and Vuri, 2014; Bradley et al., 2007). Public sector workers, as school ones, are typically covered by the national insurance system when sick (as in the Italian case), but, their effective state of health cannot be efficiently and costless monitored. This implies an incentive to take more days off than what it is necessary, preserving the whole wage and possibly causing direct pecuniary costs (the cost of a substitute) and other non-pecuniary costs (organizational) to the schools.

This paper investigates whether the regional level of social capital affects the degree of teachers' absenteeism rates in Italian schools. To this aim, we use a very rich dataset provided by the Italian Ministry of Education (MIUR) that includes data on the level of workers' absences for each school in Italy during the school year 2010/11. It is well known that compared to northern Italian regions, the Italian *Mezzogiorno* is characterized by a persistent gap with lower levels of per-capita income, higher unemployment rate and lower human capital. Different social capital indicators also show that the latter is not an exception, while recent data imply that, unlike in the rpivate sector, the public sector workers' absenteeism rates are higher in the Southern regions (Scoppa and Vuri, 2014).

Overall, this suggests that local specific informal factors may be playing an important role, and that the Italian regional sample represents a good candidate to examine different functioning and effectiveness of local educational institutions in a developed economy. In this respect, our study can be classified within the literature that exploits the effects played by informal institutions on economic outcomes: indeed, the informal habits may have had a crucial role in shaping the functioning of similar formal institutions (schools) because of local differences in social capital (Putnam, 1994; Guiso et al., 2008; Tabellini, 2010).

The MIUR dataset contains absences distinguished in three categories: sick leaves, maternity leaves and other kind of leaves. Following the literature, we use sickness absenteeism rates as a measure of potential shirking behavior among school workers. Data are at school level, and this implies that we are not able to control for the workers' individual characteristics that, as shown by (Ichino and Maggi, 2000; Chaudhury et al., 2006), represent an important component in this type of analysis.

However, unlike other studies in this field, our data also include information on the absenteeism rates of two different school workers' categories, namely, teachers and ATA (administrative, auxiliary and technical) staff. That is, for each school we have data on the behavior of two groups of workers within the same working environment, and this enables us to exploit the observed differences in order to control for common characteristics that the two categories experience in the same working environment. This is important, since poor performers (both teachers and ATA workers) may be possibly sort themselves in schools where work environment is more cheat-friendly. In fact, in schools with a stronger monitoring, workers are more likely to pay both formal (enforced by the school principal) and informal (enforced also by their peers) sanctions. This may generate locational sorting (Ichino and Maggi, 2000; Bradley et al., 2007; Cornelissen et al., 2017) and affect our OLS estimates producing inconsistent results. In our case, we are able to mitigate this problem exploiting the absence rates of the two workers' categories: in our empirical analysis this enables us to control for school components, such as management, including ATA (administrative, auxiliary and technical staff) absenteeism rate within our controls. In sum, our empirical strategy includes the use of a rich set of covariates, plus a specific control with the ATA school staff absenteeism rates for (observed and unobserved) workplace characteristics and organizational arrangements.

Moreover, since we cannot exclude that unobserved heterogeneity or reverse causality are still a source of bias for our estimates, we also implement an IV strategy and use historical data on Italian regions past dominations as instruments. More precisely, following Di Liberto and Sideri (2015) our identification strategy exploits the Italian past history in order to build different sets of instruments.[1] Indeed, unlike most European countries, Italian history has been characterized by high levels of political fragmentation. Since the Middle Ages the Italian peninsula has been also subjected to different waves of colonisations and the numerous dominators that governed over centuries had very different cultural and political features and implemented highly heterogeneous formal and informal institutions

[1]On this see Acemoglu et al. (2001), Rodrick et al. (2004), Pande and Udry (2005), Guiso et al. (2008), Tabellini (2008), Bosker and Garretsen (2009) among the others.

in the administrated territories. The two extreme cases are usually identified by the State of the Church, that was an example of corrupt institutions and administrative inability, and Austria that is usually portrayed as a good administrator that did not implement exploiting or extracting policies.[2] Thus, we expect past dominations to be correlated with local social capital, but uncorrelated with the teachers' absenteeism and in terms of the empirical strategy.

We use two different sets of instruments. Our first instrument set uses a series of dummy variables that identify, for each province, the administration that occurred during the period of the Spanish domination in Italy, 1560-1659. This choice is based on two main reasons. First, during this period the Italian peninsula was ruled by different formal governments and each dominance has lasted for a sufficiently long period. Indeed, each province experienced the same formal government for the whole period. Second, Spain has been often portrayed by historians as having negatively affected the dominated areas also through its legacy of inefficient bureaucracy. Our second approach follows a different path with respect to previous studies which are typically based on specific historical events. Instead, here we build a matrix indicating, for each province, the kind and the duration (in years) of domination that ruled during the period between the 12th and 18th century. To this aim we collect data for all different regimes that governed each Italian province over seven centuries before the creation of the unified Italian State.[3]

Finally, the use of data on two different workers categories implies that we also need to take into account of the possible differences in skill levels between teachers and ATA workers. Evidence of different behavior among workers with different skills has been found by Mas and Moretti (2009) and Cornelissen et al. (2017). In particular, Mas and Moretti (2009) find evidence of significant differences in pro-social effects[4] depending on workers' skills (performance): low skilled workers are more responsive to changes in the average productivity of coworkers than high skilled workers. The same evidence can be found in Cornelissen et al. (2017), where workers with low performance (with respect to coworkers' level or to a social norm) have been found to experience feelings of guilt or shame due to "peer pressure" and these effects are larger for low skilled occupations. Conversely, Chaudhury et al. (2006) find evidence that high absence rates are associated with better educated teachers. In our case, those elements would result in a different role of social

[2]For more on that see Di liberto and Sideri, 2015.

[3]All details in Di liberto and Sideri, 2015.

[4]Mas and Moretti (2009) define pro-social effects as cases where workers experience disutility if their peers observe them in free-riding behaviors, with formal or informal sanctions.

capital for skilled (teachers) and low-skilled (ATA) workers in Italian schools. Our results suggest that social capital has a negative correlation with the absenteeism rates of teachers, while less so (non significant coefficient) for ATA workers. In particular, a unitary increase in social capital level would result in a decrease between 0.22 and 0.36 of teachers' sickness absence rates.

The paper is organized as follows. The next section briefly describes the most important features of the institutional setting of Italian schools workers. The third section is devoted to data description and analysis. Section 4 contains the model specification, while section 5 explains our main findings. Finally, we state our conclusions.

2 Data

Our main dataset is the MIUR (2012) "La scuola in chiaro" dataset that includes a number of schools and teachers characteristics for 10,197 Italian schools. That is, we have data on approximately the 90% of the Italian public schools (primary, lower and upper secondary school levels, including students aged 6 to 19 years old), except for those located in the regions/provinces of Aosta, Trento and Bolzano.[5] These data are also merged with additional variables on Italian history and area controls.

Table 1 reports the main descriptives, and shows that the average value of the teachers absenteeism is 8.19, i.e. each teacher, on average, takes around 8 days of sickness absences per year, with a standard deviation of 3.57. In fact, teachers absenteeism measures the average number of days of absence per teacher, within a school year, for each school of the Italian public educational system. The range goes from zero to 25. The second indicator measures the absenteeism rate of ATA workers. In this case the average number of days of absence is much larger, 15.06 sickness days per year with a standard deviation of 8.46.

Differences in the statistical distributions of absenteeism days between ATA and teachers are represented in two different maps. In order to make our two categories comparable, we have rescaled with a distributions ranging between 0 and 100 both measures of the teachers' and the ATA staff. The first map, Figure 1, represents the territorial distribution of absenteeism rate for teachers. It shows clearly that darker color regions (more days off) are mostly located in the central and southern areas, in particular Sicily, Sardinia, Calabria, Basilicata, Campania and Lazio. The lowest value for teachers' absenteeism (4.78

[5]Due to their status as autonomous provinces, these are excluded from the sample. In addition, for 1,191 schools absenteeism rates were not available, while 37 schools have been detected as outliers. So, our data correspond to the 89.3% of Italian public institutes.

days) corresponds to Cuneo (Nothern region), while the province showing the highest value is Reggio Calabria (13.09 days on average, located in the South).

Figure 2, representing the territorial distribution of absenteeism rate for ATA workers, tells a different story. In this case, we do not have a clear territorial characterization, and the map indicates that the phenomenon is quite random across Italian provinces. For ATA workers, the province with the lowest absenteeism is, again, Cuneo (10.05 days), while Agrigento (located in the south) is the area showing more days of absence among ATA workers (20.86).

In Figure 3, we have represented the differential between teachers' absenteeism rate and ATA absenteeism rate. In this case, we have a sort of summary of the previous two rates. The most important element is, again, the strong dualism between the Centre-North and southern areas; thus, not only northern areas show a lower absenteeism rate for teachers, but in the northern Italy we also have evidence of a smaller absenteeism differential compared to the other provinces: in those areas the two working categories behave in a more similar way. Darkest areas, in this case, are indicating that teachers have an average absenteeism rate higher than that of ATA, corresponding almost exactly to the *Mezzogiorno*.

Figure 4 shows the relationship between our indicator of absenteeism for teachers and our social capital variable.[6] The plot enables to distinguish between Southern areas (red dots) and other provinces (black triangles). First of all, there is a clear negative correlation between teachers' absenteeism and social capital, as expected. Second, we can clearly observe two different groups: almost all the Southern provinces are characterized by low levels of social capital (with few exceptions) and high absenteeism rates (as previously stated). In Figure 5, we repeat the same analysis but considering the absence days for ATA workers. In this case, the negative correlation is much weaker.

In Tables 3, 4 and 5 we can see a more detailed analysis of our main variables. In Table 3, we see the great differences for teachers' absences across Italian macro-regions: in the Central and Southern regions teachers show respectively 8.27 and 9.20 days of sickness absences; in the Northern areas, this average is, instead, lower and equal to 6.81. As said, absences for ATA workers are much higher in all areas but North-South differences are less remarkable.

In Table 4, we distinguish our data by school type. In this case, we see that teachers show the same average in all school levels but in primary schools: in the latter, in fact, teachers are absent 9.42 days on average, about 2 days more than their colleagues in other

[6]The social capital indicator belongs to Cartocci (2007).

school levels. The same conclusion may be reached also for ATA workers.[7] This fact may be explained by the presence of younger students: in primary schools pupils are aged between 6 and 11 years old, thus, they (and their teachers) are more exposed to diseases, typical of this age. In addition, in higher school levels, early leavers and school tracking may play an important role: students in upper secondary school are the least disadvantaged and, are usually the best students, also with a high socio-economic background. For this, they may be plausibly characterized by strong parental monitoring that also keeps teachers under a stricter control, limiting their potential shirking attitude. Finally, in Table 5, we can see a further detail for upper secondary schools. Compared to other school types, in lyceums ATA workers has a higher absenteeism rate on average (even though this difference is very small), while teachers take fewer days of absences: this fact may be due to the traditional, better environment in lyceums, often characterized by better performing students and better school principals[8] that seem to affect teachers but not ATA.

Our main regressor is social capital. We use a synthetic social capital index at NUTS3 level, provided by Cartocci (2007). It contains data on 1) blood donations, 2) sport participation, 3) dissemination of newspaper and 4) voter turnout. In particular, as blood donations data are used to assess a measure of "generalized morality" and sport participation is assumed to influence social capital since it supports the building of groups of mutual interest and promotes pro-social behavior, they are both important in our contest. In fact, both elements should play a significant role in explaining different (opportunistic) behaviors among working groups. Again, in Table 1 we can see some statistics about our social capital indicator: Italian regions are, as in the case of absenteeism rates, highly heterogeneously endowed. Again, Vibo Valentia and most southern provinces show the lowest values, while North-Centre provinces have the highest ones (in particular Bologna shows the best performance).

Besides, our analysis includes a number of school's and teacher's controls. In particular, we include: the number of teachers in the school, the share of teachers with a short term contract, the share of female teachers, the share of teachers aged over 55 and the teachers' turnover rate. The first one could be considered as an indirect measure of the school size: bigger schools may, in principle, be better organized and with stronger monitoring power. The second one is particularly important because it could capture the incentives for teachers with a short term contract to work harder than those with a permanent contract

[7] *Istituti comprensivi* where different levels of education (usually primary and lower secondary levels) coexist and are managed by the same school principal, show instead a middle value of other school levels.

[8] See on this Di Liberto et al. (2013)

and according to Bradley et al. (2007), we expect a negative effect of this on the teachers' absenteeism rate. The share of female teachers can give us important information about the composition of teachers' groups. On average, female teachers are about the 80%: in this case we should expect a positive correlation with absenteeism rates[9]. Teachers aged over 55 years measures the share of "old" teachers working in the school. As we know from other studies[10], compared to other countries, the age of Italian schools teachers is fairly high: on average, 33% of teachers are in this range (this share is 37% for southern provinces), with a maximum value of 87%. We expect that sick absences are positively correlated with age. These hypothesis are consistent with correlations reported in Table 2.

Finally, teachers' turnover may be correlated with absenteeism. Since Italian school principals have little control on teachers' transfers, turnover is almost exclusively voluntary, and our correlation matrix suggests a positive correlation.[11]

Also, we insert some regional geographic characteristics as additional controls. First, we average temperatures have been identified in this literature as a significant determinant of absenteeism.[12]. We also use two dummy variables capturing whether the school is located in a rural area. As found by Duflo et al. (2012), rural areas are often characterized by higher social monitoring and, in principle, this could discourage shirking behavior.

As said, Italian regions also show significant differences in GDP and unemployment rates and we control for both in our analysis. As stressed by Scoppa and Vuri (2014) there is an inverse relationship between regional unemployment and absenteeism at individual level and, following Ichino and Riphahn (2005), employees' sickness absences are positively related to the degree of job security. Scoppa and Vuri (2014) provide evidence on the impact of unemployment on workers' absenteeism at individual level. Controlling for a number of individual and firm characteristics, they find that the individual absenteeism rate is negatively and strongly related to the provincial unemployment rate. In particular, in high unemployment southern areas, shirking is dramatically lower than in northern areas, notwithstanding South Italy is characterized by lower levels of social capital and more widespread opportunistic behavior. This effect is larger in magnitude in small firms, in which the protection from dismissals for employees is lower. In addition, as a further evidence of the role played by the unemployment as deterrent for shirking, they show that public employees are not affected by local unemployment.

[9] See Miller et al. (2008); Herrmann and Rockoff (2012)

[10] See for instance Di Liberto et al. (2013)

[11] See on this Barbieri et al. (2013): they find evidence that teachers' mobility is driven mainly by geographical distance from the place of birth and by difficulties in teaching in the school.

[12] See Miller et al. (2008)

In addition, we need to take into account a possible overestimation of the absenteeism phenomenon: both Ichino and Maggi (2000) and Scoppa and Vuri (2014) argue about the possible overestimation of their results due to real illness episodes within absenteeism rate measurement. To control for this, Scoppa and Vuri (2014) insert in their model two variables related to life expectancy and mortality rate at regional level, while we use the mortality rate at NUTS3 level.

Then, as said, we also consider some cultural controls, namely, the extortion rate (to capture the effect of corruption and the presence of organized criminal groups) and alcohol consumption rate. The first one is the rate of extortions over 1,000 inhabitants. Data show that Treviso is the province with the lowest crime rate, while Catania as the poorest performer. Also, the alcohol consumption rate is a measure of "bad habits" among society and we include it in our controls. We also control for a composite indicator of educational infrastructures: the percentage of (public) primary schools provided with meals and equipped with a school-bus, the percentage of special classrooms in (public) secondary schools and the number of teachers for every 100 students. This aspect could be relevant in reducing absenteeism rates, as underlined by Chaudhury et al. (2006).

Finally, we take into account for some students' characteristics: the retention rate[13] and the presence of early leavers. Retention rate is measured as the percentage of rejected pupils in the school year 2010/2011, while the early leavers indicators is the average percentage of early leavers during the first three years of upper secondary school.

3 OLS specification and results

We study the effects of social capital on the school staff absences using a simple regression setting of the form:

$$Y_{ij} = \alpha + \beta SK_j + \gamma X_i^{'} + \delta Z_j^{'} + \theta S_j^{'} + \upsilon_{ij} \tag{1}$$

where Y_{ij} is either the ATA or teachers' absenteeism measure in school i and region j, SK_j is our indicator of social capital in region (NUTS3) j, X_{ij} is a set of school-teacher controls (number of teachers in the school, share of teachers with a short term contract, share of female teachers, share of teachers aged over 55 and teachers' turnover rate), Z_j a set of reional additional controls (temperature, mortality rate, rural and mountain areas

[13]Data for students' retention rate are available only for 3,615 observations, corresponding to 2,461 on 3,376 Upper secondary schools and 1,154 on 5,582 *Istituti Comprensivi.*

dummies, unemployment, added value) and S_j is a set of additional area charatcteristic variables (extortion rate, alcohol consumption rate and educational infrastructures index).

We set the scene in Table 6 where we show the OLS results when we alternatively use as dependent variable our measure of the absenteeism for teachers (columns 1 to 6), and our measure of the absenteeism for the ATA staff (columns 7 to 12). Robust standard errors, clustered at regional level, are reported in parenthesis.[14]

As suggested by the previous literature and by our descriptives, we expect a negative sign on β, implying lower absenteeism rates in areas with higher social capital levels.[15]

As we see, results are not entirely consistent with our expectations. Indeed, social capital has a negative and significant coefficient in all specifications with teachers absences as dependent variable. Model 1 shows the results for our most parsimonious specification, while, model 2 introduces a set of school type dummies, namely, *Istituto comprensivo*, primary schools (pupils aged 6 to 11 years old), and Lower Secondary schools (students aged 11 to 14 years old).[16] In model 3, we include our additional school-specific controls. In this case, we find a positive and significant correlation for the coefficients of the share of female teachers, and the share of "elder" teachers. These results have been also found by Bradley et al. (2007) and Herrmann and Rockoff (2012). The turnover rate shows a positive correlation as expected.

Model 4 adds our geographical controls: climate and two dummies for rural and mountain areas. As found in other studies, the former variable shows a positive and significant correlation with absenteeism.(Shi and Skuterud, 2014). It is often the case that, teachers live in urban areas, and this implies that it is more difficult to reach schools that are located in both rural or mountainous areas and this may possibly influence the teacher's absenteeism rates. In our regressions, the rural area dummy coefficient is not significant while the mountain area coefficient is positive and significant. Further, our mortality rate indicator shows, as expected, a positive and significant correlation with absenteeism. In this specification, we also include two economic measures: the regional unemployment rates and the added value. The unemployment rate coefficient, positive and significant, seems to confirm results suggested in other studies.[17] Finally in model 6, we insert our measure of students' characteristics. Note that this cause our sample to reduce significantly. As said

[14]Note that, since our dependent variable is identified at school level but the independent variable of interest only varies across regions, we always cluster the standard errors at the regions NUTS3 level.See Angrist and Pischke (2008)

[15]See for instances Chaudhury et al. (2006)

[16]The omitted dummy variable related to is the Upper Secondary schools one (14 to 19 years old).

[17]See Scoppa and Vuri (2014)

above, students' data measure the retention rates and early leaver rates. These phenomena are very rare at both primary and lower secondary school levels, while Italy shows one of the highest dropout rates for upper secondary school among OECD countries. As expected, the retention rate is positively correlated with the absenteeism rate, while the early leavers coefficient is not significant. Moreover, even using this subsample and specification, our main indicator of social capital is negative and significant.

The picture is different when we focus on the ATA staff. As also seen in our descriptive analysis, the negative relationship between social capital and the days of absence of the ATA staff seems weaker than that observed for teachers. Indeed, in models 7 to 12, the social capital shows a non-significant coefficient in all specifications.

This difference between the two workers categories suggests that workers' skill levels are playing a role here: high-skilled workers (teachers) seem to be influenced by the area social capital, while low-skills (ATA) are not. However, this analysis goes beyond the purpose of this study and, from now on we will focus on teachers.

4 Endogeneity issues

OLS results shown in Table 6 may be plagued by endogeneity problems. Following Ichino and Maggi (2000), our main concerns are the usual ones in this kind of analysis and arise from:

- Sorting problems: it is plausible that shirking teachers (or ATA) try to work in schools where they it is less costly to apply opportunistic behaviors.
- Omitted variables and measurement error.
- Reverse causality: social capital influences absenteeism rates but, in principle, we cannot exclude that the correlation could go in the opposite direction.
- Peer effects: individuals' behavior may influence group's behavior and viceversa, generating "reflection problems"; as we know from Ichino and Maggi (2000) and Bradley et al. (2007), group interaction effects are a source of endogeneity, thus, estimates of average regional social capital may be biased, due to the so-called "reflection problem", identified by Manski (1993).

Nevertheless, our data enable us to partly control for this problem. In fact, we can exploit the absenteeism measure for ATA workers and define the following alternative model specification:

$$Y_{kij} = \alpha + \beta SK_j + \mu Y_{aij} + \gamma X_i^{'} + \delta Z_j^{'} + \theta S_j^{'} + \upsilon_{ij} \quad (2)$$

where Y_{kij} is the teachers' absenteeism rate in school i in region j, and Y_{aij} is the ATA workers' absenteeism rate in school i in region j. This strategy allows us to consider the different behavior of the two groups of workers within the same school (working environment) and thus control for possible unobservable school characteristics (such as the quality of managerial practices applied by the school principals) possibly affecting the behavior of both workers' categories. In other words, this model should be able to mitigate sorting problems omitted variables bias.[18]

Results are in Table 7, where columns 1 to 6 replicate the same model specifications (set of additional controls) used in Table 6. Overall, all results fully confirm that the social capital coefficient is always negative and significant. Second, ATA absenteeism indicator always have a positive and significant coefficient: schools with high absenteeism rates for ATA show high absenteeism rates also for teachers. In sum, even controlling for ATA absenteeism rate, our results are very stable and confirm a strong negative correlation between social capital levels and teachers' shirking. Model 4 is the that with the minimum value of our main coefficient: it indicates that one unit more of social capital index corresponds to 0.17 days less of sickness absences per school. In sum, our OLS results suggest that if Vibo Valentia (lowest level) could increase its social capital level to that observed in Bologna (highest level), teachers absenteeism in its schools would decrease (on average for all its teachers) by a value between 0.60 (implied by the value of the coefficient in model 1) and 0.17 days (model 4).

5 Robustness analysis

5.1 IV strategy

As said above, including ATA absenteeism among our controls helped us to mitigate both sorting and omitted variables problems. However, endogeneity concerns may be still present and this motivates us to implement an IV strategy, and exploit the historical dataset described in Di Liberto and Sideri (2015). Given the cross-sectional nature of our analysis, this IV strategy should also address potential reflection problems.[19]. In our study we therefore use a two-stage least square approach. In particular:

[18] We need to assume that mobility opportunities across schools is the same for both teachers and ATA.
[19] On this see Mouw (2006)

$$\text{First stage:} \qquad SK_j = \zeta + \lambda HIST_i + \mu y_{aij} + \gamma X_i^{'} + \delta Z_j^{'} + +\theta S_j^{'} \epsilon_j \tag{3}$$

$$\text{Second stage:} \qquad y_{kij} = \alpha + \beta SK_j + \mu y_{aij} + \gamma X_i^{'} + \delta Z_j^{'} + +\theta S_j^{'} \upsilon_{ij} \tag{4}$$

where HIST refers to the set of historical variables. That is, as also stressed by Angrist and Piescke (2010) history matters since it enables researchers to find good instruments and to get through one of the main difficulties they have to face when endogeneity represents a real threat to estimate results. Thus, researchers need to identify plausible critical historical facts that may be palusibly used as instruments: in our case, we need to assume that past dominations have influenced current social capital levels, and may possibly influence current teachers' absenteeism rates but only through social capital. The importance of past administrations in the social capital endowments has been stressed by among the many others by Putnam (1994) and Tabellini (2010). According to these studies, there is a strong parallelism between civic degree in the medieval period and the current social capital endowments in Italian regions.

In choosing our instruments set we follow Di Liberto and Sideri (2015) that collected data for all different regimes that governed each Italian NUTS3 regions over seven centuries before the creation of the unified Italian State. In particular, this study considers "...the period between 1100 and 1800 where the historical lower bound is determined by the high political instability of the Peninsula from the Holy Roman Empire downfall until the Norman rise (about 1100) and also by the absence of reliable historical documents. The upper bound has been chosen, again, because since 1800 the Napoleonic era had established a situation of dramatic changes and instability in the Italian politics with a series of wars that persisted until the Italian Unity in 1861".[20] During these 700 years it is possible to identify the following dominations: the Normans, the Swabians, the Anjou, the Spanish (Aragonese until 1502), the Bourbons, the Papal State, the Savoy, the Austrians and the Republic of Venice, and construct a matrix that assigns to each province the number of years during which each regime has persisted in a specific territory. We do not provide further details here, leaving the interested reader to Di Liberto and Sideri (2015).

In Table 8, we find our IV results. We show first stage results (odd number columns) but we mainly discuss our second stage results that represent the focus of this study. IV results show that social capital coefficient, in all specifications, is significant and very stable, confirming previous results. This holds also the results found for the remaining regressors. Since we only have one endogenous variable (social capital) we conduct inference that is

[20] Di Liberto and Sideri (2015) p.21.

robust to weak instruments using Moreiras (2003) conditional likelihood ratio (CLR) test statistics. This allows us to create confidence intervals robust to weak instruments that we include among results, together with Limited Information Maximum Likelihood (LIML) estimates. Reassuringly, LIML estimates are always bounded and with the same sign of our social capital coefficient. Moreover, for each model we test for underidentification and for weak instruments. We firstly test for underidentification using the Kleibergen-Paap test [21]. The p-values always reject the null, suggesting that our model is not underidentified. However, this element is not sufficient for strong identification and we always report first-stage F statistics based on Cragg and Donald (1993). As a rule of thumb, then, we check if the first-stage F-statistic is larger than ten. Except for models 4 and 5, we always note that F-statistic is larger than ten. Finally, we also report the p-value of the Hansen-j test of over-identifying restrictions to check whether our instruments are together coherent[22] and it shows that, overall, our set of instruments is coherent: only the first two specifications show a p-value lower than 0.05 and greater than 0.01.

5.2 Different types of schools

As a final control, we also split our sample by the different school types. We have already seen in the descriptive analysis that the different school types show different characteristics: in particular, primary schools have a higher average absenteeism rate compared to other schools. Thus, we want to check whether different school types produce different results and, at the same time, and if our previous results on the overall negative role of social capital for absenteeism rate are robust to these different subsamples.

We replicate the IV specification described in the previous section and Table 9 reports our results: models from 1 to 3 focus on "Istituti comprensivi", 4 to 6 on primary schools, 7 to 9 on lower secondary and 10 to 12 on upper secondary schools. For upper secondary schools, unlike the other level of schooling, we insert among regressors a set of dummy identifying the different types: Lyceums, Vocational schools and technical schools.

For primary and lower secondary school subsamples highlight we observe some problems in the social capital LIML coefficient when we control for the full specification.[23] However, overall also this analysis confirms the previous results: the social capital coefficient is neg-

[21] See Kleibergen and Paap (2006): this test allows to determine whether the minimal canonical correlation between the endogenous variables and the instruments is statistically different from zero and is an alternative to the Anderson canonical test; we need to use it, in this specific case, because our IV regressions are made implementing clustered standard errors at NUTS3 level.

[22] On this, see Parente and Silva (2012)

[23] The coefficient interval shows a positive rather than negative sign in the upper bound when we control for local factors.

ative and significant in all specifications and we do not observe any significant changes also in the other coefficients.

6 Conclusions

In this paper, we investigate the relationship between social capital and shirking (measured as the average number of sickness days per teacher per school) among Italian teachers in public schools. To this aim, we have used a unique dataset provided by MIUR, containing a rich set of information of all Italian public schools for the year 2010-11. Our main results suggest that local social capital levels play an important role: indeed, in most specifications social capital coefficients are negative and significant, implying a negative association with the teachers' absenteeism rate, with a unitary increase in our social capital index resulting in a decrease of teachers' absenteeism rate between 0.36 and 0.22. Results are robust to the inclusion of different sets of covariates, and of a measure of the ATA staff absenteeism rates in our models. The latter is an important variable since it takes into account for likely problems arising from a locational sorting phenomenon.

Finally, to solve for possible additional endogeneity problems, we also apply an instrumental variables approach using as instruments the historical data on regional past domination as in Di Liberto and Sideri (2015). Again, these results fully corroborate the OLS ones.

In sum, our study suggests that differences in social capital levels significantly affect teachers' absenteeism rates. Future work should further investigate why different school workers groups with different skills (teachers and ATA) seem to be differently affected by the area social capital level.

Bibliography

J. D. Angrist and J.-S. Pischke. *Mostly Harmless Econometrics*. Princeton University Press, 2008.

G. Barbieri, C. Rossetti, and P. Sestito. Teacher mobility and student learning, 2013.

S. Bradley, C. Green, and G. Leeves. Worker absence and shirking: Evidence from matched teacher-school data. *Labour Economics*, 14(3):319–334, jun 2007. doi: 10.1016/j.labeco.2006.05.002.

R. Cartocci. *Mappe del tesoro. Atlante del capitale sociale in Italia*. Il Mulino, 2007. ISBN 9788815118608.

N. Chaudhury, J. Hammer, M. Kremer, K. Muralidharan, and F. H. Rogers. Missing in action: Teacher and health worker absence in developing countries. *Journal of Economic Perspective*, pages 1–91, 2006.

T. Cornelissen, C. Dustmann, and U. Schnberg. Peer effects in the workplace. *American Economic Review*, 107(2):425–456, feb 2017. doi: 10.1257/aer.20141300.

A. Di Liberto and M. Sideri. Past dominations, current institutions and the Italian regional economic performance. *European Journal of Political Economy*, 38(C):12–41, 2015. doi: 10.1016/j.ejpoleco.2014.1.

A. Di Liberto, F. Schivardi, M. Sideri, and G. Sulis. Le competenze manageriali dei dirigenti scolastici italiani. *FGA Working Paper*, 2013.

A. Di Liberto, F. Schivardi, and G. Sulis. Managerial practices and student performance. *Economic Policy*, 30(84):683–728, 2015.

E. Duflo, R. Hanna, and S. P. Ryan. Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4):1241–1278, jun 2012. doi: 10.1257/aer.102.4.1241.

H. H. Gaziel. Predictors of absenteeism among primary school teachers. *Social Psychology of Education*, 7(4):421–434, 2004. doi: 10.1007/s11218-004-5232-z.

L. Guiso, P. Sapienza, and L. Zingales. Social Capital as Good Culture. CEPR Discussion Papers 6657, C.E.P.R. Discussion Papers, Jan. 2008. URL `https://ideas.repec.org/p/cpr/ceprdp/6657.html`.

M. A. Herrmann and J. E. Rockoff. Worker absence and productivity: Evidence from teaching. *Journal of Labor Economics*, 30(4):749–782, oct 2012. doi: 10.1086/666537.

A. Ichino and G. Maggi. Work environment and individual background: Explaining regional shirking differentials in a large italian firm. *The Quarterly Journal of Economics*, 115(3):1057–1090, 2000. URL `https://ideas.repec.org/a/oup/qjecon/v115y2000i3p1057-1090..html`.

A. Ichino and R. T. Riphahn. The effect of employment protection on worker effort: Absenteeism during and after probation. *Journal of the European Economic Association*, 3(1):120–143, 03 2005. URL `https://ideas.repec.org/a/tpr/jeurec/v3y2005i1p120-143.html`.

F. Kleibergen and R. Paap. Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1):97–126, jul 2006. doi: 10.1016/j.jeconom.2005.02.011.

C. F. Manski. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531, jul 1993. doi: 10.2307/2298123.

A. Mas and E. Moretti. Peers at work. *American Economic Review*, 99(1):112–145, feb 2009. doi: 10.1257/aer.99.1.112.

R. T. Miller, R. J. Murnane, and J. B. Willett. Do teacher absences impact student achievement? longitudinal evidence from one urban school district. *Educational Evaluation and Policy Analysis*, 30(2):181–200, jun 2008. doi: 10.3102/0162373708318019.

M. J. Moreira. A conditional likelihood ratio test for structural models. *Econometrica*, 71 (4):1027–1048, jul 2003. doi: 10.1111/1468-0262.00438.

T. Mouw. Estimating the causal effect of social capital: A review of recent research. *Annual Review of Sociology*, 32(1):79–102, aug 2006. doi: 10.1146/annurev.soc.32.061604.123150.

P. M. Parente and J. S. Silva. A cautionary note on tests of overidentifying restrictions. *Economics Letters*, 115(2):314–317, may 2012. doi: 10.1016/j.econlet.2011.12.047.

R. D. Putnam. *Making Democracy Work*. University Press Group Ltd, 1994. ISBN 0691037388.

V. Scoppa and D. Vuri. Absenteeism, unemployment and employment protection legislation: evidence from italy. *IZA Journal of Labor Economics*, 3(1):3, 2014. doi: 10.1186/2193-8997-3-3.

J. Shi and M. Skuterud. Gone fishing! reported sickness absenteeism and the weather. *Economic Inquiry*, 53(1):388–405, jun 2014. doi: 10.1111/ecin.12109.

J. Stock and M. Yogo. Testing for weak instruments in linear IV regression. Technical report, NATIONAL BUREAU OF ECONOMIC RESEARCH, nov 2002.

G. Tabellini. Culture and institutions: Economic development in the regions of europe. *Journal of the European Economic Association*, 8(4):677–716, jun 2010. doi: 10.1111/j.1542-4774.2010.tb00537.x.

A The institutional setting

A.1 Teachers' and ATA school workers

As said, in Italy school personnel is organized in two categories: the administrative and technical staff (ATA) and the teachers. The administrative, technical and auxiliary staff (ATA) includes all workers in charge of general school administration and logistic tasks. Precisely, they are classified in four areas:

- Area A: in this area, we find the Schooling collaborator (CS), the lowest level among ATA; this position requires the achievement of three schooling years of secondary education;
- Area AS: in this area, the working profile is identified with the Schooling collaborator specialized in agricultural firms (CR); this figure is present only in agriculture professional schools and requires a professional diploma (secondary education) in the field of agricultural studies;
- Area B: in this area we find different profiles; the Administrative assistant (AA), present in all schools, requires a full diploma (5 years of secondary education); the Technical Assistant (AT), present in secondary schools, works in the school laboratories and requires a full diploma, with a specific curriculum depending on the typology of the laboratory; cook (CU), only in boarding schools, requires a diploma from an hospitality training institute; nurse (IF), only in boarding schools, requires a bachelor degree in nursing sciences; cloakroom attendant (GU), only in boarding schools, requires a diploma in the fashion sector;
- Area D: in this area we find the apical profile of school administration, the director of general and administrative services (DSGA); each school has its own DSGA and it has a degree in law, in political sciences or in economics.

ATA workers can be hired from two different rankings at region level. The first is a permanent ranking: people can enter this ranking through a selection based only on qualifications for profiles of Areas A and B.[24] Candidates may access this selection (published yearly from the Schooling regional office) only if they can demonstrate a previous experience of, at least, 24 months in public schools in the same profile for which the candidate wants to apply. Permanent rankings are used to hire ATA with long-term contracts on the basis of the annual needs of the school system. The second ranking, instead, is temporary (it lasts 24 months) and is used to hire short-term workers.[25] A specific procedure is dedicate for hiring the DSGA: rankings for this position are made by ordinary selections based on qualifications and tests and are used both for short-term and long-term contracts.

Past selection processes for ATA were regulated by the Decree of the President of the Republic 31st of May 1974, number 420.[26] This decree established a set of rules for school

[24] This selection process is regulated by the Legislative Decree 16th of April 1994, number 297, article 554.

[25] It is regulated by the Ministerial Decrees 75/2001 and 35/2004.

[26] Decreto presidente della Repubblica 31 Maggio 1974, n. 420: *Norme sullo stato giuridico del personale non insegnante statale delle scuole materne, elementari, secondarie ed artistice*

"non-teaching" workers (since 1990s they are defined ATA). This Decree distinguished between administrative workers and workers with more practical and manual tasks. The first were hired evaluating their qualifications and through a public selection process at provincial level; to access those selections, the law required an upper secondary school diploma. Second kind of non-teaching workers, instead, were selected through a simple process evaluating their qualifications. Candidates with a lower secondary diploma could access those selections, but they had to demonstrate at least an experience of two years in the same role. The same decree regulated also non-teaching workers' turnover: in particular, they were able to ask for moving to other schools in the same regions; after their requests, the provincial school officer was able to establish a ranking based on tenure and other qualifications. In this way, non-teaching workers were assigned taking into account their requests but with a priority related to their ranking position and upon vacancies availability. Workers willing to move from other regions were assigned with a lower priority.

Teachers, instead, are selected, obviously, with higher requirements, due to the intellectual nature of their job. Their selection is distinguished between primary school and secondary school level. For the first, it is required a bachelor degree in Education or a Diploma in Education (the latter only for those that achieved it before the schooling year 2001-2002). Secondary school teachers, instead, must have a master degree in the field corresponding to the teaching discipline. The Ministerial Decree 10th of September 2010, number 249, has established a new set of rules, introducing the TFA (active training internship), a compulsory training of 1,500 hours to obtain the teaching qualification to be undertaken after the achievement of a master degree.

Past selection processes for teachers were regulated by the Decree of the President of the Republic 31st of May 1974, number 417.[27] Primary school teachers were selected by a public selection at provincial level, while secondary school teachers through a public selection at regional level. Both of them required an academic degree and were selected after a successful achievement of a written test, a course lasting four months and a final oral examination, and considering their qualifications. Teachers were assigned to each school on the basis of their ranking position, resulting from the selection process. Successful candidates were appointed for a probationary period lasting one school year. The same decree regulated also teachers' turnover, with no differences compared to ATA.

Several laws has been issued during 1980s and 1990s aimed at modifying the ATA and teachers' selection process but none of them has modified the entry requirements.

[27] Decreto Presidente Repubblica 31 maggio 1974, n. 417: *Norme sullo stato giuridico del personale docente, direttivo ed ispettivo della scuola materna, elementare, secondaria ed artistica dello Stato*

B Figures and Tables

B.1 Figures

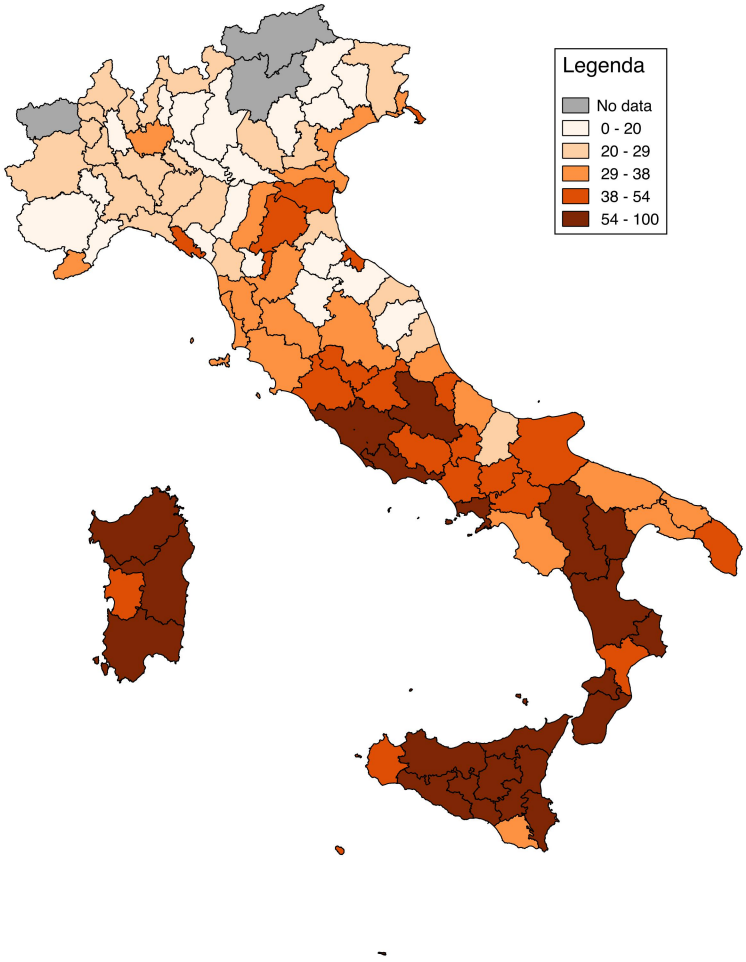

Figure 1: Absenteeism rate distribution across Italian regions - teachers

Notes: Territorial distribution across the 103 Italian regions of our indicator of absenteeism rate for teachers. Description of data sources in Appendix C.

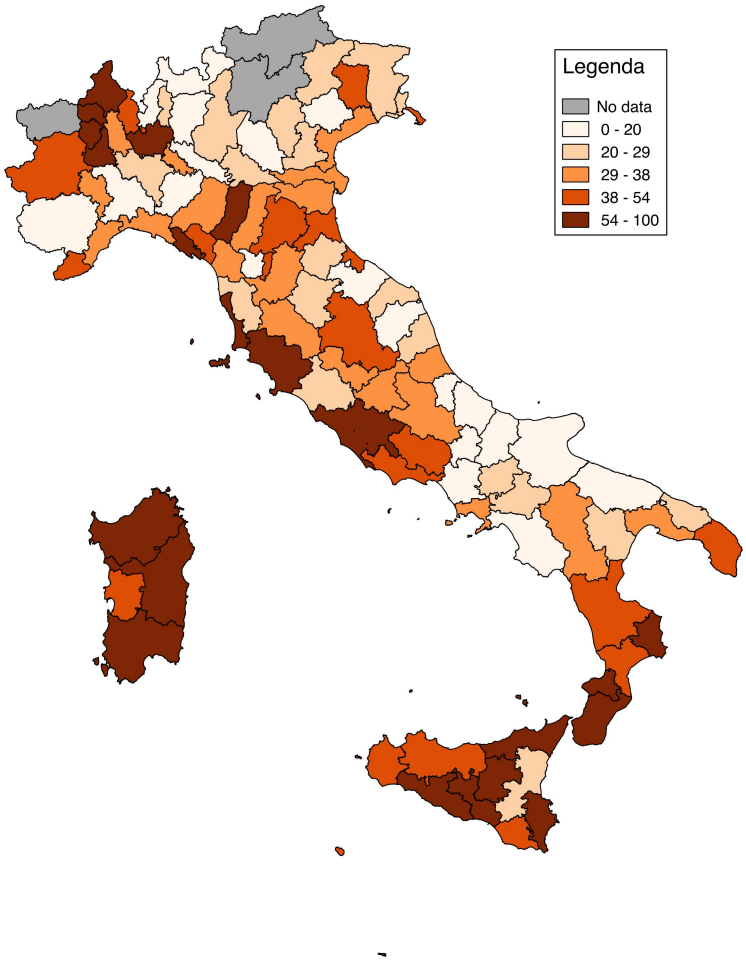

Figure 2: Absenteeism rate distribution across Italian regions - ATA

Notes: Territorial distribution across the 103 Italian regions of our indicator of absenteeism rate for ATA. Description of data sources in Appendix C.

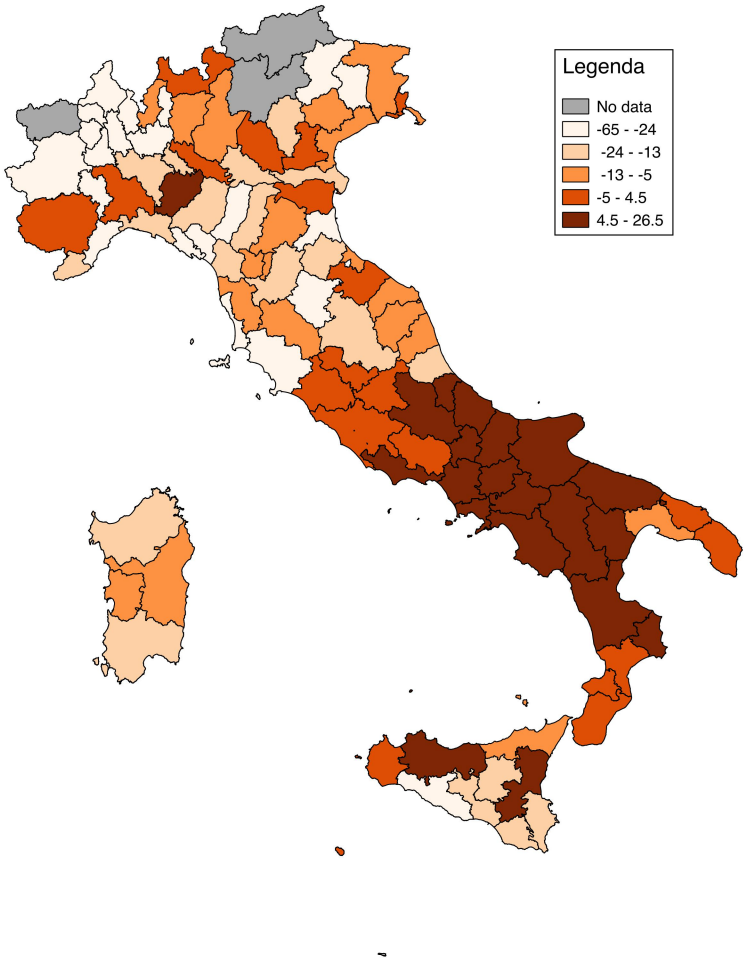

Figure 3: Absenteeism differential across Italian regions - teachers vs ATA

Notes: Territorial distribution across the 103 Italian regions of the differential between the absenteeism rate for teachers and the absenteeism rate for ATA. Description of data sources in Appendix C.

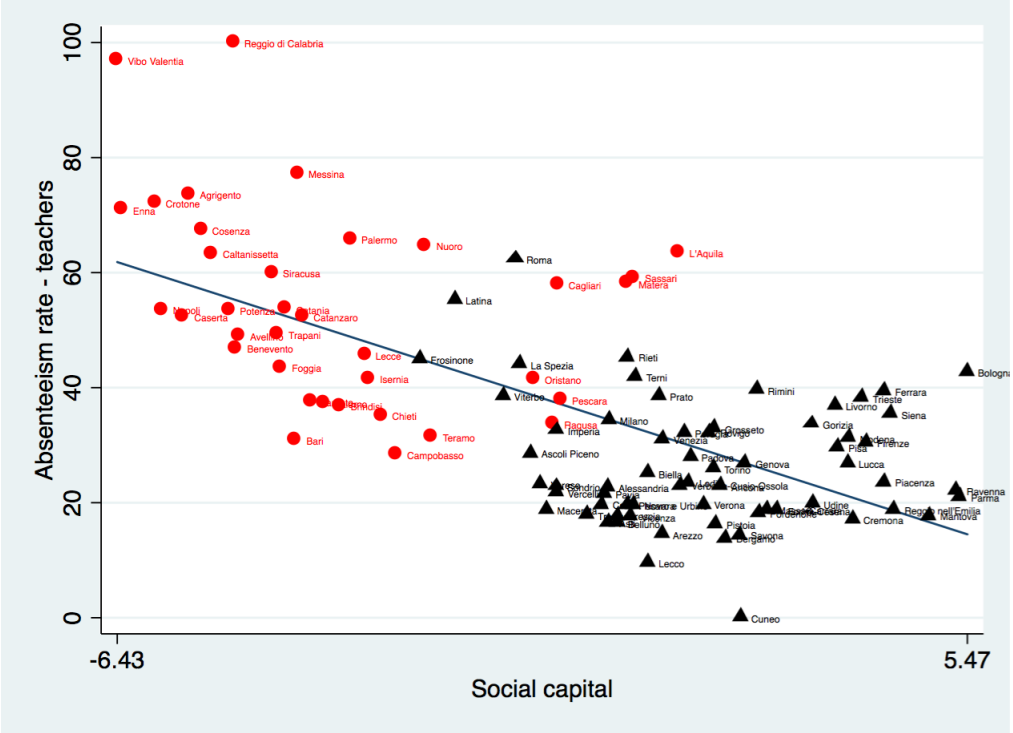

Figure 4: Teachers' absenteeism rate and social capital

Notes: Teachers absenteeism rate (vertical axis) and our social capital indicator (horizontal axis). Red dots identify Southern regions, black triangles identify Cental and Northern regions. Data sources are described in Appendix C.

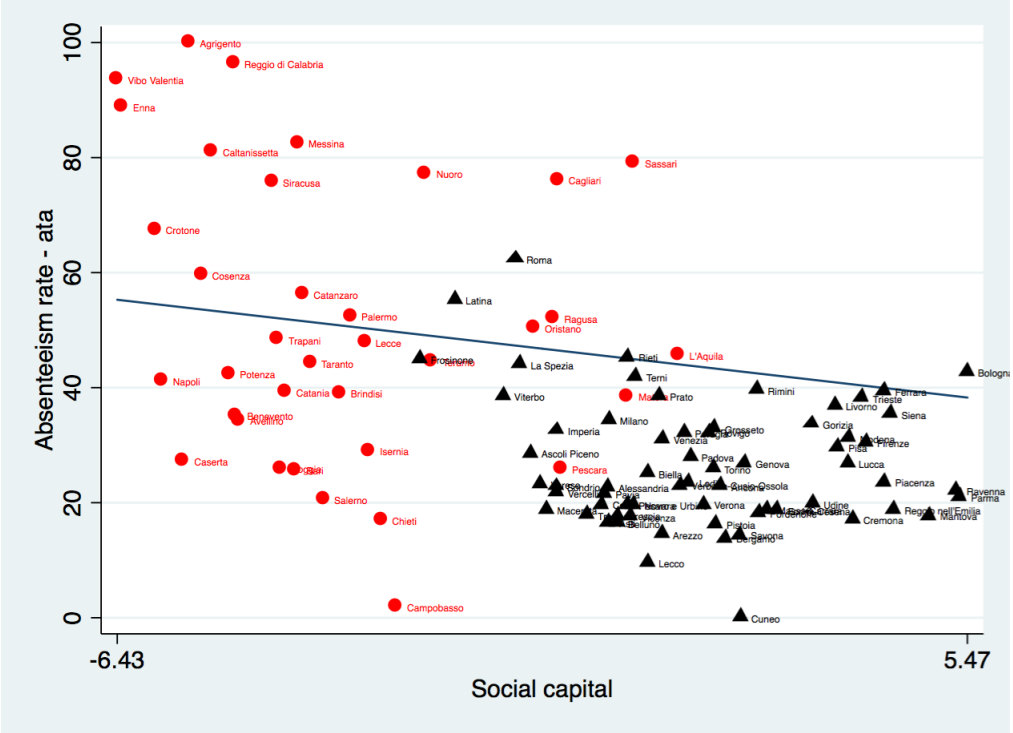

Figure 5: ATA absenteeism rate and social capital

Notes: ATA absenteeism rate (vertical axis) and our social capital indicator (horizontal axis). Red dots identify Southern regions, black triangles identify Centre and Northern regions. Description of data sources in Appendix C.

B.2 Tables

Table 1: Descriptive statistics

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Teachers' absenteeism rate	10197	8.187	3.572	0	25
ATA absenteeism rate	10197	15.063	8.463	0	56
Teachers' absenteeism rate (prov)	100	7.818	1.566	4.78	13.09
ATA absenteeism rate (prov)	100	15.044	2.095	10.05	20.86
Social capital	100	-0.075	3.136	-6.43	5.47
Log of Number of teachers	10147	4.122	0.407	1.39	5.63
Teachers short term contract (%)	10197	10.899	10.178	0	100
Female teachers (%)	10147	81.416	14.705	22.5	100
Teachers over 55 (%)	10136	33.061	11.745	2.2	87.433
Teachers turnover (%)	10145	6.299	5.824	0	100
Mortality rate	100	103.403	14.236	81.055	142.062
Rural area	10197	0.235	0.424	0	1
Mountain area	10197	0.361	0.480	0	1
Unemployment rate, 2009	100	7.956	3.662	2.09	17.64
Added value per capita, 2010	100	21857.13	5611.00	13097.14	45839.81
Average temperature 00-09	100	13.718	2.501	6.1	18.3
Extortion rate	100	6.362	3.697	1.71	19.45
Alcohol consumption rate	100	23.686	5.864	13.15	37.3
Educational infrastructures	100	2.017	0.516	0.78	2.91
Retention rate	7918	0.458	1.848	0	31.28
Early leavers	9942	3.446	4.453	0	70.56

Table 2: Correlation matrix

	Teac. abs.	ATA abs.	Social capit.	N. of teac.	Short contr.	Fem. teac.	Over 55	Teac. turn.	Av. temp.	Mort. rate	Rural area	Mount. area	Unem. rate	Added value	Ext. rate	Alc. cons.	Educ. infr.	Ret. rate	Earl leav
Teachers' abs. rate	1																		
ATA abs. rate	0.1975	1																	
Social capital	-0.2859	-0.0315	1																
N. of teachers	-0.078	0.0621	0.2757	1															
Teachers short term	-0.1169	-0.058	0.2463	0.0143	1														
Female teac.	0.0648	0.0315	0.121	-0.1342	-0.2327	1													
Over 55	0.1667	0.0553	-0.3268	-0.1709	-0.4316	-0.0725	1												
Teachers turnover	0.1075	-0.0197	-0.1247	-0.1665	0.152	0.0128	-0.0857	1											
Average temp.	0.3093	0.0831	-0.6146	-0.1908	-0.2211	-0.0876	0.2828	0.0936	1										
Mortality rate	-0.0856	0.0148	0.4502	-0.0028	0.1086	0.044	-0.0231	-0.0285	-0.2836	1									
Rural area	0.0327	-0.0421	-0.0594	-0.3638	0.1268	0.1022	-0.0893	0.0954	0.0499	0.1361	1								
Mountain area	0.075	-0.0029	-0.048	-0.219	0.086	-0.0019	0.0186	0.0371	-0.0443	0.127	0.3006	1							
Unemp. rate	0.2956	0.0686	-0.7577	-0.2688	-0.2407	-0.1319	0.3102	0.108	0.6753	-0.3113	0.0659	0.0625	1						
Added value p. c.	-0.1825	0.0109	0.6333	0.3381	0.1939	0.1463	-0.2601	-0.1049	-0.4842	0.0318	-0.2086	-0.096	-0.6671	1					
Ext. rate	0.2201	0.0388	-0.5659	-0.199	-0.1666	-0.0891	0.2526	0.1059	0.4824	-0.1569	0.0783	0.0072	0.5059	-0.5163	1				
Alcohol cons. rate	-0.2295	-0.0119	0.7082	0.2028	0.2093	0.0805	-0.321	-0.1438	-0.5446	0.2395	-0.0489	-0.0604	-0.6572	0.5123	-0.5408	1			
Educ. infrastr.	-0.2541	-0.0315	0.8064	0.251	0.2587	0.1198	-0.2554	-0.0951	-0.6355	0.4793	-0.0153	-0.0095	-0.8319	0.6498	-0.5403	0.6616	1		
Retention rate	0.0366	-0.0026	-0.1133	0.0276	0.0887	-0.2455	-0.0047	0.0257	0.1384	-0.0628	-0.0157	-0.0038	0.1495	-0.1148	0.0936	-0.0759	-0.1368	1	
Early leavers	0.0104	0.003	-0.0598	0.1371	0.1708	-0.5118	0.0131	0.0167	0.1194	-0.047	-0.136	-0.0423	0.1408	-0.0645	0.033	-0.0417	-0.1088	0.3266	1

Table 3: Absenteeism rate, by macro-region

Variable	Obs	Mean	Std. Dev.	Min	Max
			North		
Teachers' absenteeism rate	3602	6.806774	2.817927	0	22
ATA absenteeism rate	3602	14.63937	7.48799	0	50
			Center		
Teachers' absenteeism rate	1858	8.267313	3.341308	0	24
ATA absenteeism rate	1858	15.41362	7.884136	0	54
			South		
Teachers' absenteeism rate	4737	9.20456	3.817461	0	25
ATA absenteeism rate	4737	15.24805	9.325534	0	56

Table 4: Absenteeism rate, by school type

Variable	Obs	Mean	Std. Dev.	Min	Max
			Istituti comprensivi		
Teachers' absenteeism rate	5057	8.078011	3.474121	0	25
ATA absenteeism rate	5057	15.09719	8.344254	0	56
			Primary schools		
Teachers' absenteeism rate	1996	9.423848	3.640996	0	24
ATA absenteeism rate	1996	15.5481	8.576543	0	52
			Lower secondary schools		
Teachers' absenteeism rate	1071	7.737628	3.622267	0	22
ATA absenteeism rate	1071	14.64986	9.507483	0	52
			Upper secondary schools		
Teachers' absenteeism rate	2073	7.493086	3.423531	0	25
ATA absenteeism rate	2073	14.72697	8.037495	0	53

Table 5: Absenteeism rate, by upper secondary school type

Variable	Obs	Mean	Std. Dev.	Min	Max
			Lyceum		
Teachers' absenteeism rate	822	6.954988	3.185177	0	22
ATA absenteeism rate	822	14.96229	8.395654	0	53
			Professional schools		
Teachers' absenteeism rate	895	7.76648	3.458316	0	22
ATA absenteeism rate	895	14.60559	7.725996	0	48
			Other schools		
Teachers' absenteeism rate	356	8.048221	3.695779	0	25
ATA absenteeism rate	356	14.48876	7.968504	0	47

Table 6: OLS regressions - Dep. var.: Absenteeism rate

VARIABLES	(1) TEAC	(2) TEAC	(3) TEAC	(4) TEAC	(5) TEAC	(6) TEAC	(7) ATA	(8) ATA	(9) ATA	(10) ATA	(11) ATA	(12) ATA
Social capital	-1.2764***	-1.2474***	-1.1445***	-0.6699***	-0.7445***	-0.7935***	-0.1205	-0.1193	-0.1515	0.0036	-0.1696	-0.2905
	(0.168)	(0.168)	(0.168)	(0.228)	(0.249)	(0.267)	(0.137)	(0.136)	(0.137)	(0.201)	(0.210)	(0.224)
Log of Number of teachers			1.3660	1.8552***	1.9179***	0.7200			3.2168***	3.1661***	3.3133***	3.0272***
			(0.923)	(0.644)	(0.620)	(0.685)			(0.629)	(0.492)	(0.479)	(0.982)
Teachers short term contract (%)			0.0152	0.0199	0.0178	0.0436			-0.0124	-0.0100	-0.0111	-0.0585*
			(0.029)	(0.022)	(0.022)	(0.029)			(0.024)	(0.024)	(0.024)	(0.035)
Female teachers (%)			0.0849***	0.0892***	0.0910***	0.0824***			0.0426***	0.0421***	0.0457***	0.0634***
			(0.019)	(0.012)	(0.012)	(0.027)			(0.016)	(0.015)	(0.015)	(0.022)
Teachers over 55			0.1387***	0.1078***	0.1060***	0.1318***			0.0829***	0.0597***	0.0620***	-0.0000
			(0.027)	(0.022)	(0.021)	(0.028)			(0.022)	(0.022)	(0.020)	(0.029)
Teachers turnover (%)			0.2118***	0.1970***	0.2006***	0.1526**			-0.0056	-0.0145	-0.0037	-0.0379
			(0.043)	(0.041)	(0.040)	(0.064)			(0.035)	(0.034)	(0.033)	(0.058)
Average temperature 00-09				0.9182***	0.9159***	0.7419***				0.4957*	0.5216**	0.1784
				(0.257)	(0.245)	(0.259)				(0.254)	(0.237)	(0.273)
Mortality rate, 09-11				0.0699**	0.0682*	0.0483				0.0780**	0.0838**	0.1071**
				(0.034)	(0.036)	(0.037)				(0.032)	(0.036)	(0.041)
Rural area				-0.7366	-0.7623*	-0.6785				-0.9001	-0.8897	-1.4812*
				(0.449)	(0.428)	(1.007)				(0.572)	(0.550)	(0.793)
Mountain area				2.1389***	2.2181***	2.8054***				0.9069	1.0078*	0.5881
				(0.727)	(0.725)	(0.699)				(0.573)	(0.533)	(0.695)
Unemployment rate, 2009				0.5565***	0.6220***	0.4333**				0.4306**	0.5077***	0.4117*
				(0.199)	(0.217)	(0.183)				(0.197)	(0.178)	(0.238)
Added value per capita, 2010				0.0002*	0.0002*	0.0001				0.0002***	0.0002***	0.0002***
				(0.000)	(0.000)	(0.000)				(0.000)	(0.000)	(0.000)
Extortion rate					0.1501	0.0749					0.1233	0.0281
					(0.102)	(0.106)					(0.125)	(0.158)
Alcohol consumption rate					0.1112	0.1122					0.2124***	0.0994
					(0.099)	(0.099)					(0.080)	(0.087)
Educational infrastructures					0.4909	0.7629					0.0888	-1.0286
					(1.499)	(1.368)					(1.361)	(1.578)
Retention rate						0.2093***						0.1070***
						(0.032)						(0.040)
Early leavers						-0.0517						-0.0215
						(0.115)						(0.121)
Dummy School level		Yes	Yes	Yes	Yes			Yes	Yes	Yes	Yes	
Observations	10,197	10,197	10,135	10,135	10,135	2,931	10,197	10,197	10,135	10,135	10,135	2,931

Notes: Standard errors in parentheses, clustered at region level: *** p<0.01, ** p<0.05, * p<0.1.

Table 7: OLS - Dep. Var.: Teachers absenteeism rate

VARIABLES	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Social capital	-1.2542***	-1.2262***	-1.1187***	-0.6705***	-0.7192***	-0.7527***
	(0.149)	(0.149)	(0.150)	(0.210)	(0.228)	(0.248)
ATA absenteeism rate	0.1836***	0.1784***	0.1702***	0.1508***	0.1492***	0.1405***
	(0.014)	(0.014)	(0.013)	(0.013)	(0.012)	(0.023)
Log of Number of teachers			0.8184	1.3776**	1.4236**	0.2947
			(0.862)	(0.632)	(0.606)	(0.683)
Teachers short term contract (%)			0.0173	0.0214	0.0194	0.0518*
			(0.028)	(0.022)	(0.021)	(0.028)
Female teachers (%)			0.0776***	0.0829***	0.0842***	0.0735***
			(0.019)	(0.012)	(0.012)	(0.026)
Teachers over 55			0.1246***	0.0988***	0.0967***	0.1319***
			(0.026)	(0.021)	(0.020)	(0.027)
Teachers turnover (%)			0.2127***	0.1992***	0.2011***	0.1579**
			(0.039)	(0.038)	(0.038)	(0.061)
Average temperature 00-09				0.8435***	0.8381***	0.7168***
				(0.224)	(0.215)	(0.233)
Mortality rate, 09-11				0.0581*	0.0557*	0.0332
				(0.030)	(0.032)	(0.033)
Rural area				-0.6008	-0.6296	-0.4704
				(0.415)	(0.398)	(0.958)
Mountain area				2.0021***	2.0678***	2.7227***
				(0.702)	(0.696)	(0.655)
Unemployment rate, 2009				0.4915***	0.5463***	0.3754**
				(0.181)	(0.202)	(0.172)
Added value per capita, 2010				0.0002	0.0002	0.0001
				(0.000)	(0.000)	(0.000)
Extortion rate					0.1317	0.0709
					(0.090)	(0.097)
Alcohol consumption rate					0.0795	0.0982
					(0.090)	(0.091)
Educational infrastructures					0.4777	0.9074
					(1.353)	(1.250)
Retention rate						0.1943***
						(0.031)
Early leavers						-0.0487
						(0.113)
Dummy School level		Yes	Yes	Yes	Yes	
Observations	10,197	10,197	10,135	10,135	10,135	2,931

Notes: Standard errors in parentheses, clustered at region level: *** p<0.01, ** p<0.05, * p<0.1.

Table 8: IV regressions - Dep. Var.: Teachers absenteeism rate

VARIABLES	(1) I_Stage	(2) IV	(3) I_Stage	(4) IV	(5) I_Stage	(6) IV	(7) I_Stage	(8) IV	(9) I_Stage	(10) IV
Normans	-0.0397***		-0.0392***		-0.0228***		-0.0195***		-0.0232***	
	(0.003)		(0.003)		(0.003)		(0.003)		(0.005)	
Swabians	-0.0038***		-0.0039***		-0.0059***		-0.0041***		-0.0046***	
	(0.001)		(0.001)		(0.001)		(0.001)		(0.001)	
Anjou	-0.0064***		-0.0064***		-0.0045***		-0.0071***		-0.0058***	
	(0.001)		(0.001)		(0.001)		(0.001)		(0.001)	
Spain	-0.0099***		-0.0098***		-0.0056***		-0.0063***		-0.0060***	
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
Bourbons	0.0205***		0.0204***		-0.0021		0.0104***		0.0101	
	(0.004)		(0.004)		(0.003)		(0.004)		(0.007)	
Papal state	-0.0050***		-0.0050***		-0.0038***		-0.0033***		-0.0033***	
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
Venice	-0.0039***		-0.0039***		-0.0019***		-0.0031***		-0.0030***	
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
Austria	0.0004		0.0004		-0.0007*		-0.0016***		-0.0007	
	(0.000)		(0.000)		(0.000)		(0.000)		(0.001)	
Savoy	-0.0020***		-0.0019***		-0.0015***		-0.0011***		-0.0012***	
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
Social capital (TSLS)		-1.4563***		-1.3323***		-1.0029***		-0.8989**		-1.2917***
		(0.162)		(0.176)		(0.340)		(0.398)		(0.493)
Social capital (LIML)		-1.4077***		-1.3327***		-0.9448***		-1.1000***		-1.6660***
CLR (95%)		[-1.506; -1.324]		[-1.447; -1.236]		[-1.187; -0.717]		[-1.442; -0.793]		[-2.258; -1.129]
ATA absenteeism rate	-0.0031***	0.1784***	-0.0028***	0.1703***	-0.0026***	0.1539***	-0.0033***	0.1515***	-0.0045**	0.1344***
	(0.001)	(0.013)	(0.001)	(0.012)	(0.001)	(0.011)	(0.001)	(0.010)	(0.002)	(0.016)
Log of Number of teachers			0.0407	0.9816*	0.1739***	1.4282***	0.1515***	1.3873***	0.1721**	0.3237
			(0.039)	(0.567)	(0.040)	(0.489)	(0.039)	(0.510)	(0.068)	(0.631)
Teachers short term contract (%)			0.0062***	0.0329	0.0007	0.0061	-0.0012	0.0053	-0.0013	0.0302
			(0.002)	(0.021)	(0.002)	(0.019)	(0.002)	(0.019)	(0.003)	(0.027)
Female teachers (%)			0.0029**	0.0688***	0.0021*	0.0777***	0.0019	0.0778***	-0.0005	0.0604**
			(0.001)	(0.012)	(0.001)	(0.011)	(0.001)	(0.011)	(0.002)	(0.024)
Teachers over 55 (%)			-0.0026*	0.1212***	-0.0076***	0.0781***	-0.0086***	0.0912***	-0.0047	0.0995***
			(0.001)	(0.025)	(0.001)	(0.019)	(0.001)	(0.020)	(0.003)	(0.025)
Teachers turnover (%)			0.0047*	0.2177***	0.0015	0.1887***	0.0019	0.1922***	-0.0058	0.1189**
			(0.003)	(0.034)	(0.002)	(0.033)	(0.002)	(0.032)	(0.004)	(0.048)
Average temperature 00-09					-0.0091	0.8217***	0.0057	0.8213***	0.0031	0.6055***
					(0.011)	(0.204)	(0.011)	(0.186)	(0.021)	(0.185)
Mortality rate, 09-11					0.0409***	0.0625**	0.0260***	0.0617*	0.0237***	0.0693
					(0.001)	(0.032)	(0.002)	(0.034)	(0.003)	(0.045)
Rural area					0.2200***	-0.7557*	0.1514***	-0.7028*	0.0997	-0.0598
					(0.037)	(0.412)	(0.036)	(0.388)	(0.080)	(0.838)
Mountain area					-0.0519*	1.6442***	-0.0617**	1.2083**	-0.1379**	2.1648***
					(0.031)	(0.497)	(0.030)	(0.500)	(0.058)	(0.610)
Unemployment rate, 2009					-0.0760***	0.3541*	-0.0314***	0.5343***	-0.0243*	0.3040*
					(0.007)	(0.183)	(0.007)	(0.181)	(0.014)	(0.156)
Added value per capita, 2010					0.0000***	0.0001	0.0000***	0.0001	0.0000**	0.0001
					(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Extortion rate							-0.0062	0.1171	-0.0013	0.0352
							(0.004)	(0.092)	(0.009)	(0.096)
Alcohol consumption rate							0.0610***	0.1201*	0.0512***	0.1212
							(0.006)	(0.068)	(0.011)	(0.079)
Educational infrastructures							0.8412***	0.4527	0.9461***	2.1047
							(0.063)	(1.354)	(0.118)	(1.497)
Retention rate									-0.0007	0.1963***
									(0.004)	(0.030)
Early leavers									0.0078	-0.0760
									(0.009)	(0.096)
Dummy School level		Yes		Yes		Yes		Yes		
Observations	10,197	10,197	10,135	10,135	10,135	10,135	10,135	10,135	2,931	2,931
R-squared	0.803	0.140	0.804	0.155	0.826	0.186	0.833	0.187	0.828	0.153
first-stage F statistics		46.42		40.51		7.590		4.561		4.570
Kleibergen-Paap P-value		8.61e-07		6.10e-06		0.000782		0.00700		0.00889
Hansen J statistic		16.99		16.54		9.146		10.30		7.325
Chi-sq(3) P-val		0.0302		0.0353		0.330		0.245		0.502
Pagan-Hall P-value		0		0		0		0		0

Notes: Standard errors in parentheses, clustered at region level: *** p<0.01, ** p<0.05, * p<0.1. Social capital assumed endogenous. Excluded instruments are the variables Normans, Swabians, Anjou, Bourbons, Austria, Papal State, Savoy, Spain and Venice (Independent states not included) that identify the number of years during which each regime has persisted in a specific territory during 1100-1800.

Table 9: Robustness checks

VARIABLES	(1) IV	(2) IV	(3) IV	(4) IV	(5) IV	(6) IV	(7) IV	(8) IV	(9) IV	(10) IV	(11) IV	(12) IV	(13) IV
	Istituti comprensivi			Primary schools			Lower secondary schools			Upper secondary schools			
Social capital (TSLS)	-1.6686***	-1.6849***	-0.9079**	-1.1965***	-1.1427***	-1.1236*	-1.3249***	-1.2804***	-0.9874	-1.2493***	-1.2567***	-1.6049***	-1.5592***
	(0.177)	(0.181)	(0.389)	(0.214)	(0.294)	(0.592)	(0.237)	(0.244)	(0.654)	(0.166)	(0.163)	(0.514)	(0.467)
Social capital (LIML)	-1.6123***	-1.6367***	-1.1057***	-1.0897***	-1.1507***	-0.8984*	-1.2058***	-1.1166***	-1.0888*	-1.3069***	-1.2316***	-1.4763***	-1.3040***
CLR (95%)	[-1.747; -1.496]	[-1.809; -1.493]	[-1.536; -0.707]	[-1.308; -0.877]	[-1.418; -0.893]	[-1.847; 0.043]	[-1.494; -0.924]	[-1.448; -0.792]	[-2.346; 0.168]	[-1.515; -1.113]	[-1.450; -1.027]	[-2.200; -0.814]	[-2.070; -0.600]
ATA absenteeism rate	0.1794***	0.1725***	0.1573***	0.1617***	0.1609***	0.1481***	0.1979***	0.1646***	0.1496***	0.1578***	0.1477***	0.1300***	0.1277***
	(0.013)	(0.013)	(0.014)	(0.021)	(0.020)	(0.019)	(0.028)	(0.027)	(0.022)	(0.026)	(0.024)	(0.023)	(0.023)
Log of N. of teachers		1.4623**	1.4059**		0.8357	1.5113		3.3481**	3.9506***		-0.2642	0.4154	0.0354
		(0.738)	(0.604)		(1.384)	(1.211)		(1.342)	(1.491)		(0.844)	(0.835)	(0.799)
Teachers short cont. (%)		0.0769***	0.0654***		0.1123	0.0921		0.0385	0.0398		0.0276	-0.0050	-0.0032
		(0.025)	(0.021)		(0.108)	(0.126)		(0.049)	(0.046)		(0.035)	(0.034)	(0.035)
Female teachers (%)		0.1239***	0.1404***		-0.1417**	-0.0998		-0.0234	-0.0686		0.0940***	0.1017***	0.1135***
		(0.015)	(0.013)		(0.072)	(0.084)		(0.079)	(0.070)		(0.027)	(0.026)	(0.027)
Teachers over 55 (%)		0.1105***	0.0859***		0.0594	0.0460		0.2123***	0.2022***		0.1575***	0.1416***	0.1154***
		(0.028)	(0.024)		(0.038)	(0.033)		(0.052)	(0.046)		(0.032)	(0.034)	(0.036)
Teachers turnover (%)		0.2254***	0.1817***		0.3046***	0.2974***		0.1371*	0.1254		0.1758***	0.1907***	0.1912***
		(0.032)	(0.030)		(0.093)	(0.084)		(0.079)	(0.078)		(0.067)	(0.064)	(0.064)
Average temp. 00-09			1.1331***			0.5879***			0.2197			0.4610*	0.6528***
			(0.161)			(0.186)			(0.256)			(0.248)	(0.211)
Mortality rate, 09-11			0.0544			0.1385***			0.0819			0.0846	0.0739
			(0.035)			(0.047)			(0.057)			(0.052)	(0.045)
Rural area			-0.8110*			-1.4912*			-0.1938			-1.0310	-1.0826
			(0.422)			(0.823)			(1.446)			(1.259)	(1.448)
Mountain area			1.3252***			0.5226			1.6786			2.0930***	1.9151**
			(0.496)			(1.042)			(1.051)			(0.733)	(0.749)
Unemployment rate, 2009			0.3854**			0.7409***			0.7522**			0.5148***	0.4112**
			(0.165)			(0.267)			(0.296)			(0.187)	(0.179)
Added value p. c., 2010			0.0001			0.0003**			0.0000			0.0001	0.0001
			(0.000)			(0.000)			(0.000)			(0.000)	(0.000)
Extortion rate			0.2390**			0.0117			-0.2920**			-0.0436	-0.0547
			(0.099)			(0.129)			(0.146)			(0.119)	(0.117)
Alcohol consumption rate			0.1144			0.1890**			-0.0197			0.2120**	0.2246***
			(0.075)			(0.096)			(0.138)			(0.086)	(0.080)
Educational infrastr.			0.5566			-1.2900			0.8253			2.9375*	2.2675
			(1.138)			(2.401)			(2.644)			(1.669)	(1.614)
Retention rate													0.0977**
													(0.047)
Early leavers													-0.2262
													(0.149)
Dummy Profes. School										4.9066***	4.6775***	4.7327***	5.1433***
										(0.898)	(1.093)	(0.977)	(1.276)
Dummy Other Schools										3.3241***	4.2229***	4.4881***	4.0149***
										(0.631)	(0.786)	(0.774)	(1.017)
Observations	5,057	5,056	5,056	1,996	1,996	1,996	1,071	1,068	1,068	2,073	2,015	2,015	1,789
R-squared	0.138	0.162	0.208	0.083	0.089	0.131	0.101	0.125	0.147	0.102	0.121	0.142	0.146
first-stage F statistics	46.15	39.82	4.466	44.16	39.20	4.603	44.10	36.18	4.168	55.44	53.32	5.168	5.456
Kleibergen-Paap P-value	0.00	0.00	0.0183	0.00	0.00	0.00	0.00	0.000177	0.00777	0.00	0.00	0.00596	0.00506
Hansen J statistic	16.89	16.92	9.980	10.85	11.75	11.33	14.24	12.21	9.590	11.81	9.088	6.387	6.602
Chi-sq(3) P-val	0.0313	0.0309	0.266	0.210	0.163	0.183	0.0757	0.142	0.295	0.160	0.335	0.604	0.580
Pagan-Hall P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Notes: Standard errors in parentheses, clustered at region level: *** p<0.01, ** p<0.05, * p<0.1. Social capital assumed endogenous. Excluded instruments are the variables Normans, Swabians, Anjou, Bourbons, Austria, Papal State, Savoy, Spain and Venice (Independent states not included) that identify the number of years during which each regime has persisted in a specific territory during 1100-1800. First stage results omitted.