I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16203

Managerial Practices and Student Performance: Evidence from Changes in School Principals

Adriana Di Liberto
Ludovica Giua
Fabiano Schivardi
Marco Sideri
Giovanni Sulis

JUNE 2023

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16203

Managerial Practices and Student Performance: Evidence from Changes in School Principals

Adriana Di Liberto
University of Cagliari, IZA and CRENoS

Ludovica Giua
University of Cagliari and CRENoS

Fabiano Schivardi
LUISS, EIEF and CEPR

Marco Sideri
University of Cagliari

Giovanni Sulis
University of Cagliari, IZA and CRENoS

JUNE 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Managerial Practices and Student Performance: Evidence from Changes in School Principals*

We study how managerial practices of school principals affect student performance and aspirations. We link administrative data on secondary Italian students to the management scores of their school principals in 2011 and 2015 based on the World Management Survey methodology. The frequent turnover of school principals over this period allows us to causally interpret school-fixed-effect estimates. We find that management quality positively and substantially impacts standardized math and language tests and student desire to attend college. The comparison to pooled-OLS suggests that fixed effects correct for the downward bias arising from selection of better school principals into more difficult schools.

JEL Classification: L2, I2, M1, O32

Keywords: management, productivity, school principals, student outcomes

Corresponding author:
Adriana Di Liberto
Dipartimento di Scienze Economiche e Aziendali
University of Cagliari
via S. Ignazio 17
09123, Cagliari
Italy

E-mail: diliberto@unica.it

* We thank the World Management Survey team for their training and full support during the first wave of the Italian WMS data collection process. We also thank Fondazione Agnelli, the seminar participants at Brunel University (London) and the 10th Italian Congress of Econometrics and Empirical Economics (Cagliari) for very useful comments and discussions, and INVALSI for providing the 2011 and 2015 data on student outcomes. Anna Maria Maullu (Associazione Nazionale Presidi) and Sergio Repetto (Ufficio Scolastico Regionale Sardegna) gave us crucial information on Italian school principals. We are solely responsible for all the remaining errors. Financial support from the Regione Sardegna Legge 7/2007 (grant CUP F71J09000300002; grant no. CRP 60619; grant CUP F25C20001170002) and from the Fondazione Giovanni Agnelli is gratefully acknowledged. Fabiano Schivardi thanks the ERC for financial support (ERC grant 835201).

1 Introduction

Growth theory has long established that human capital is the main engine of growth (Lucas, 1988), a prediction confirmed by empirical evidence (Mankiw, Romer, and Weil, 1992; Hanushek and Woessmann, 2012). Given that human capital is to a large extent accumulated in schools, it is not surprising that much research effort has been devoted to identifying the attributes of a schooling system that enhance student achievements. However, a clear consensus on these attributes, actionable for policy recommendations, has yet to be reached. For example, the value-added literature has shown that teachers are very important for student achievements, but there is no clear-cut indication of which teacher characteristics matter and, as a consequence, how to increase teacher effectiveness.[1]

In recent years, some encouraging results have emerged from the growing literature on school management. Using RCTs in US schools, Fryer (2014, 2017) finds that management matters, although evidence on the persistence of the treatment is inconclusive. Moreover, external validity remains a concern. Bloom et al. (2012, 2015) and Di Liberto, Schivardi, and Sulis (2015) use the World Management Survey (WMS henceforth), a tool which measures the quality of the managerial practices within organizations (Bloom and Van Reenen, 2007, 2010, 2011), to show that schools that are better managed are also characterized by better performing students.[2] These results, based on different countries, suggest that enforcing good managerial practices could be an effective way to improve student achievements. However, their interpretation in terms of causality is questionable, as they cannot completely rule out unobserved heterogeneity correlated with both the quality of managerial practices and student achievements.

In this paper, we provide evidence that the quality of managerial practices in schools has a causal effect on student educational outcomes. We do so using a panel of Italian high schools that changed the school principal (SP henceforth) between 2011 and 2015 and whose managerial practices have been measured under the old and the new SP. The SP is the key figure in the functioning of the school, responsible for determining the working conditions in which teachers operate. A change in the SP, therefore, is likely to induce a discontinuous

[1] In the value-added literature, the importance of teachers is estimated by regressing student achievements on teacher fixed effects. The typical finding is that the latter explain a large portion of the overall variance of student performance (Hanushek, 2011). However, when the estimated fixed effects are regressed on observable teacher characteristics, such as age, gender, teaching experience, field of study, or certification, no clear correlation emerges (Burgess, 2016).

[2] On this, see also Lemos, Muralidharan, and Scur (2021) who modify the original WMS tool to obtain more granular but yet comparable measures of management quality more suitable for less developed countries analysis, and Leaver, Lemos, and Scur (2019) who show how to scale up the measurement of school management using existing public data.

change in the managerial practices within the school. We exploit this change and estimate the effect of the quality of managerial practices on student achievements, while controlling for school fixed effects in our regressions. In this way, we account for any fixed unobserved heterogeneity at the school level, such as the socio-economic status of the catchment area or the school's reputation, that can influence the quality of the students attracted by the school.

Fixed effects estimates are particularly well-suited to our setting for several reasons. First, the change in the SP can lead to substantial changes in managerial practices, as SPs are responsible for running the school. This limits the potential bias coming from the fact that fixed effects might substantially reduce the signal-to-noise ratio if most of the variation is cross-sectional. Second, we apply the WMS protocol, which provides a precise measure of the quality of managerial practices. Third, in the Italian school system, SPs manage the school in terms of organization but have limited autonomy in other areas such as the mission, the curricula, and teaching styles. This is different from a corporation, where a new CEO might choose to change the product mix or to enter new markets, making it difficult to disentangle changes in managerial practices from other possible changes within the firm. Thus, in our context, any changes in student outcomes can be reasonably attributed to changes in managerial practices introduced by the new SP.

We build on an initial sample of SPs interviewed in 2011 to measure the quality of managerial practices in their school (Di Liberto, Schivardi, and Sulis, 2015). In 2015, we identified schools that had changed their SPs and run a second wave of interviews to the new SPs ("switcher schools"). For schools that did not change SP, we assign the same management score computed in 2011, assuming that the quality of managerial practices is a fixed attribute of the SP. We further extend our sample by assigning the score of the new SPs interviewed in 2015 to the schools they managed in 2011. Finally, we link this panel dataset of SPs and their corresponding schools to administrative data on school characteristics and student performance, aspirations and background information. Our final sample comprises 309 SP-school-year observations and 23,514 student-year observations.

Our goal is to estimate the effect of managerial quality on different student outcomes: standardized test scores in math and (Italian) language, as well as aspirations to pursue tertiary education. The key identification challenge is that SPs might select into schools based on their ability. While we control for a large set of student, SP and school characteristics, there might still be unobserved school attributes related to the SP's managerial ability and to student outcomes. The bias could go either way. For example, more capable

SPs might be assigned to the best schools. Conversely, it is possible that school districts assign the best SPs to the most problematic schools. Our fixed effects strategy controls for any time-invariant unobserved heterogeneity at the school level, addressing concerns of non-random assignment. One potential threat is that the quality of students attending the school changes with changes in the SP's managerial ability. We argue this is not likely to be the case in the short period we consider and offer evidence supporting this claim. Finally, we show that the *change* in the managerial index at the school level has a degree of variability similar to the cross-sectional within-year variability, which indicates that fixed effects are not likely to suffer from low statistical power.

Our results are clear cut: better managerial practices improve student achievements and aspirations. We find that a standard deviation increase in management quality increases average test score in mathematics by 0.09 standard deviations. We find a positive effect, albeit slightly smaller in magnitude, for language test scores too: a standard deviation increase in managerial quality increases Italian test scores by 0.07 standard deviations. As for student aspirations, a standard deviation increase in the managerial score makes it more likely that students wish to achieve at least a college degree and that they aim to obtain a higher educational attainment than their parents by around a third of a standard deviation.

Interestingly, when running a pooled OLS regression, that is, without school fixed effects, we obtain smaller estimates (and not significant, in the case of language test scores and educational aspirations). This indicates that, if anything, the allocation of SPs to schools tends to bias downward the OLS estimates, possibly due to better SPs being assigned to worse schools. However, we find no evidence of selection on observables: for example, there is no correlation between student scores in 2011 and the managerial quality of the SP in 2015. This implies that selection occurs on unobservables, questioning the possibility to identify the effects in cross-sectional regressions.

We check the robustness of our results along several dimensions. Given that our estimates rely on "switcher" schools, that is, schools that change SP between 2011 and 2015, we show that switching itself has no effect on student outcomes. Moreover, switcher schools are not different from the others in terms of observable characteristics of the school and of the students measured in 2011. Our results are also robust to including additional controls, excluding SPs close to retirement, who might receive preferential treatment in the assignment process, and restricting the sample to the 56 balanced effective switchers, that is, schools whose SPs were interviewed in 2011 (old SP) and 2015 (new SP).

Finally, we analyze the heterogeneity of the estimates. First, all components of the

overall managerial index, which assess different aspects of school management (more on this below), have similar effects on student outcomes. Additionally, the results obtained from quantile regression models suggest that the effects are fairly homogeneous across the different percentiles of the student scores.

This paper contributes to the scant literature addressing the causal effect of management quality on student outcomes. While existing works report cross-sectional evidence (Fryer, 2014, 2017; Tavares, 2015; Di Liberto, Schivardi, and Sulis, 2015; Romero, Sandefur, and Sandholtz, 2020), our analysis exploits a panel dataset.[3] This allows us to disentangle the effect of managerial practices implemented by SPs from other attributes of the school. We also add to the body of research that investigates the importance of school leadership using the value-added model to build individual SP effectiveness measures. As for teachers, this model cannot inform us on how to enhance the effectiveness of SPs (Branch, Hanushek, and Rivkin, 2012; Böhlmark, Grönqvist, and Vlachos, 2016; Bartanen and Husain, 2022). Our paper delivers a clear policy implication: student achievements benefit from improvements in managerial practices.

In addition, our analysis contributes to the growing but still limited literature focusing on the role of school effectiveness on outcomes other than student achievements (Angrist, Hull, and Walters, 2022). We here on student educational aspirations. These are identified as important determinants of the individual incentives to invest by recent behavioral models describing typical poverty-trap mechanisms (Dalton, Ghosal, and Mani, 2016; Genicot and Ray, 2017).[4] We show that improving school management quality may be effective in increasing educational aspirations even among students whose parents did not attend college.[5]

Beyond schools, we contribute to the literature that tries to estimate the causal effect of managerial practices on performance. To the best of our knowledge, the only other studies that take advantage of panel data in combination with WMS-based measures of

[3] Fryer (2014, 2017) and Romero, Sandefur, and Sandholtz (2020) are based on RCT field experiments, while Tavares (2015) and Di Liberto, Schivardi, and Sulis (2015) adopt an instrumental variable approach.

[4] Academic ambitions differ substantially by socio-economic background, even among students who are similarly proficient at school (Haveman and Wolfe, 1995; OECD, 2019), while recent evidence finds that lower educational aspirations are associated with poorer school outcomes (Guyon and Huillery, 2021). Using data on Italy, Carlana, La Ferrara, and Pinotti (2022) find that specific school programs may be effective in modifying student aspirations and soft-skills among high-achieving immigrants, while Pagani, Comi, and Origo (2021) show that class rank improves conscientiousness through perceived ability and academic motivation.

[5] A large literature finds that parent's education matters for the student educational ambitions (Haveman and Wolfe, 1995; Björklund and Salvanes, 2011; OECD, 2019). Evidence on Italy shows that the educational career is even more affected by parental education than in other industrialized countries (Checchi and Flabbi, 2013).

management quality are Bloom, Sadun, and Van Reenen (2016) and Bloom et al. (2019). They analyze private corporations and find that the link between management quality and performance is positive and statistically significant when including firm fixed effects. However, the size tends to be smaller than in OLS estimates. We focus on schools that changed SPs, which has two advantages. First, the change in SP is likely to bring about substantial changes in managerial practices, reducing the concern of limited time series variability in the practices. Second, SPs in our setting are markedly constrained in their executive capacity in comparison to managers of private firms, and this allows us to rule out potential time-varying confounders that would not be accounted for by school fixed effects.

The remainder of the paper is organized as follows. Section 2 describes the data. Section 3 introduces the empirical framework and outlines the identification. Section 4 discusses the results, selection issues and heterogeneity. Section 5 concludes.

2 Data and descriptive statistics

We use a rich dataset on upper secondary Italian schools that merges information from three different sources in two school years (2011 and 2015). We first collect longitudinal data on managerial practices adopted in Italian secondary schools using the WMS methodology (Bloom and Van Reenen, 2007). These are then merged with administrative data on student performance and background information drawn from INVALSI (National Institute for the Evaluation of the Educational System of Instruction and Training) and with a second administrative dataset on school characteristics provided by MIUR (the Italian Ministry of Education).

In this section we describe the construction of the school management panel dataset and discuss how these variables relate to the channels identified by the literature through which SPs and their managerial activities influence student outcomes. Then, we present the information drawn from the two administrative datasets. Last, we provide some descriptive evidence of our key variables.

2.1 Management outcomes: the World Management Survey panel data

The WMS project supplies a protocol that enables to collect high-quality data on the managerial practices adopted within different organizations, including schools (Bloom et al., 2015). This methodology is based on a standardized telephone double-blind interview and comprises a set of open-ended questions that are evaluated using a scoring grid (Bloom and Van Reenen, 2011, 2010). The open-ended nature of the questions allows to investigate

aspects that cannot be captured by closed questions and interviewers are trained extensively to ensure a uniform scoring of the answers. The SP answers collected by the interviewers are converted into a score that summarizes the quality of managerial practices and ranges between 1 (worst) and 5 (best).

The WMS covers 23 managerial activities grouped into five specific management areas: operations, monitoring, targets, people (i.e. human resource management), and leadership. In our context, *Operations* (four questions) is concerned with the standardization of the educational processes, the personalization of teaching and the diffusion of best practices within the school. *Monitoring* (five questions) focuses on monitoring performance and reviewing the results at the school level, while *Targets* (five questions) assesses the quality of the process through which quantitative and qualitative targets are set and their interconnection in the short, medium and long run. *People* (five questions) is dedicated to human resource management, such as removing poor performers, rewarding employees based on performance, and hiring and keeping the best teachers or staff in schools. Finally, *Leadership* (three questions) assesses the SP's leadership capacity jointly with a clear definition of roles and responsibilities within the school.[6]

Within the literature on management and educational outcomes, human resource management (*People*) is consistently identified as a critical area of management for schools. However, the level of autonomy that SPs have in this area depends heavily on institutional factors that shape their ability to select and incentivize teachers and staff (Bloom et al., 2015; Bartanen and Grissom, 2023). This is important to note because the WMS measures actual management practices, which are influenced by these institutional constraints. In Italy, SPs have limited autonomy when it comes to teacher allocation and salaries, as these decisions are made at the central level.

Italian SPs have more discretion in other areas of management which also impact student performance. For example, the introduction of organizational innovations that enable teachers to work more effectively is captured by the *Operations* section of the survey. The *Monitoring* and *Targets* sections capture activities such as supervising teachers, monitoring their performance, assigning them to classrooms, and setting specific targets. The *Leadership* section of the survey captures the motivation of the teaching staff, which is also an important aspect of the SP job.[7] In addition, SPs can directly influence student outcomes

[6]The full set of questions on managerial practices is reported in Table C1 in Appendix C. Our management index excludes the last question of the survey, which was not asked in the 2015 follow up survey. See also Bloom et al. (2015) for a discussion concerning the measurement of managerial practices in schools.

[7]The *Operations* and *Leadership* sections broadly cover what the management literature and education scholars identify as instructional leadership activities (Grissom and Loeb, 2011) and transformational

through activities such as allocating teachers and students across classes or addressing discipline and absenteeism (Bartanen, 2020). SPs can coordinate effective school-wide policies, such as increasing communication from school staff to families or implementing data systems to support at-risk students. While the WMS does not directly measure the quality of specific SP actions on these factors, they are likely captured by the other areas of management analyzed, since these policies can only be implemented in a well-organized and monitored environment. Following the literature, in most empirical specifications we use an overall management quality index that is calculated as the average of the scores obtained in each question of the survey. However, we will also investigate its individual components.

Together with the management scores, the WMS also collects data on demographics and background characteristics of the SPs and the school they manage. This set of variables includes information on age, gender, tenure as SP within the school, overall tenure (both as teacher and SP) within the school, whether they had other job experience outside the school, whether they chose to be assigned to that specific school, and if they manage multiple schools.[8]

We exploit a two-wave panel of managerial scores at school level. During the school year 2010-11, we collected the first wave of data on managerial practices of SPs for a representative sample of 341 upper-secondary Italian schools. During the school year 2014-15, we checked if our 2010-11 schools had the same SP or not. In schools where the SP did not change, we assume that the quality of managerial practices also remained the same. Hence, in 2014-15 we assign the same managerial score the SP obtained during the 2011 interview.[9] As for the schools where the SP had changed since 2011 ("switcher" schools), we interviewed the new SP in 2015 following the same WMS protocol.

Finally, to further increase our panel sample size we exploit a "chain" approach (see Appendix A), that still assumes that the quality of managerial practices stays constant

leadership activities (Robinson, Lloyd, and Rowe, 2008).

[8]The few and lengthy national selection processes that have occurred in Italy over the years imply that, during the period analyzed, SPs were fewer than the number of vacancies in Italian schools, and some SPs managed multiple schools in the same year. The management of multiple schools (*reggenze*) is usually allowed in exceptional cases and for limited periods of time. However, since this may affect the quality of managerial practices implemented in a school, in our analysis we include a specific dummy which flags if a SP runs multiple schools.

[9]This assumption is especially reasonable in our context, due to the short time interval between the two waves of data collection. Interviewing SPs twice may also result in recall bias, questioning the comparability of the two measurement exercises. On this, see also Lemos, Muralidharan, and Scur (2021), who exploit panel data on a sample of Indian schools and student outcomes but measure management quality only once in each school at the end of the study period, treating school management as fixed over time. In any case, given that we use school fixed effects, schools that did not experience SPs turnover do not contribute to the estimation of the effect of managerial practices on student performance.

over time. In detail, we asked the new SPs interviewed for the first time in 2015 which school they managed in 2010-11 and assign the same managerial score to the school they were running in 2010-11. As for 2014-15, we collected new data from additional interviews with the new SPs in these schools. With this approach, we also extended our panel with new schools that were not part of the first wave sample.[10]

However, since there may be complementarities between the SP and the school, and SPs may adapt their managerial practices accordingly, in our robustness section we will check if our results are confirmed when we exclude these additional schools and use only the smaller sub-group of "switcher schools". A detailed description of our WMS data collection process and the composition of the sub-samples of schools is in Appendix A.

2.2 Administrative data on students and schools

To obtain our final dataset, we first merge the WMS database with the database provided by INVALSI, a government agency that carries out a yearly evaluation of student attainment in both mathematics and Italian language. This is our main source of information at the student level.

The INVALSI standardized tests are compulsory for all Italian students attending public or private schools in specific grade levels. We focus on tenth-grade upper secondary school students in the 2010-11 and 2014-15 school years. Our outcomes consist of standardized test scores in math and language and aspirations to pursue tertiary education.[11]

The INVALSI questionnaire also collects detailed information about the student's background and family characteristics. In our analysis we include the following additional student demographic information: gender, immigration status, age relative to the student's class cohort, class size and socioeconomic status (SES).[12] The latter is proxied for by the parents' occupational status, their educational attainment and the household's possession

[10] The hypothesis of constant managerial quality is also supported by the literature that uses value-added (VA) models to measure SPs' contribution to improving student outcomes, and which relies on the premise that a SP's effectiveness is constant across any two schools (Bartanen and Husain, 2022).

[11] Educational aspirations are measured by INVALSI with the following question: "Which educational attainment do you wish to achieve?". Respondents can choose among: a) compulsory education only; b) secondary vocational or technical qualification; c) high-school diploma; d) post-secondary vocational or technical qualification; e) bachelor's degree; f) higher education (Masters or PhD).

[12] As for immigration status, we flag whether the student is a first- or second-generation immigrant. We use two dummies to control for the relative age-to-grade cohort. One is for students who are at the expected age for their grade level (or "regular" students, 15 years old), and the other is for students who are one year younger and attending tenth-grade classes ahead of schedule (14 years old). The remaining category encompasses students who are repeating a grade or those (typically immigrant students) who attend a grade lower than their age would imply.

of educational resources.[13]

Finally, we merge a second administrative dataset on public schools ("*La scuola in chiaro*") provided by the Italian Ministry of Education (MIUR), which includes additional information on teacher and staff characteristics at the school level.[14] For each school, the MIUR dataset provides time varying information on the number of teachers and students, the number of permanent teachers, the number of female teachers, the number of administrative staff, and the number of students who are transferred to or from another school.

2.3 Descriptive statistics

Our final sample consists of data on all schools from the WMS panel that could be matched to the administrative datasets from the INVALSI and MIUR. It comprises 309 school-SP-year observations (172 in 2011 and 137 in 2015) and 23,514 student-year observations. This is the sample used in the main analysis. In some specifications, we also rely on different versions of the sample, including a balanced sub-sample of "switcher schools", i.e. the 56 schools that change the SP and where the management quality is measured both in 2011 and 2015. Additional details about these different samples together with potential attrition issues are available in Appendix A. Table B1 in the Appendix reports the main descriptive statistics separately for 2011 and 2015 for all variables used in the analysis: outcome variables, student characteristics, managerial quality indexes, and SP and school characteristics.

Figure 1 describes our four dependent variables. The plots at the top show the kernel distributions of the mathematics and language standardized test scores by year (2011 is in blue and 2015 is in red). The distributions are statistically different across subjects and years. The math scores tend to be distributed along the whole range of skill, with a larger variance in 2015. The distributions of the language scores are positively skewed, with a leftward shift of the 2015 score distribution relative to the 2011 score distribution.

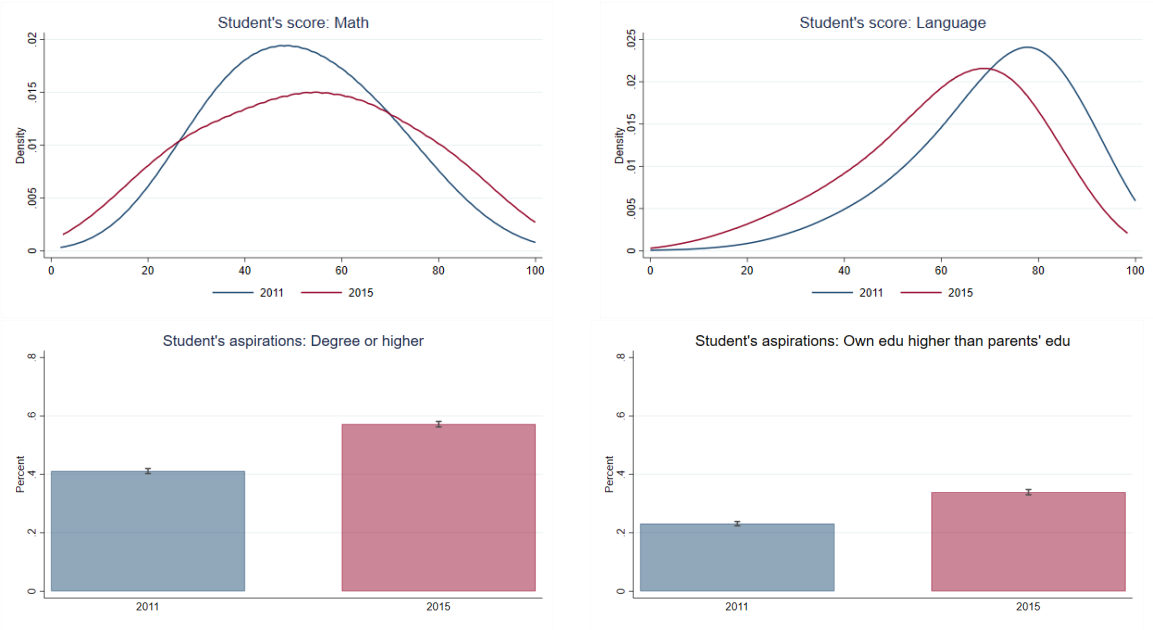
[13] We categorize parents' occupational status as: self-employed, high-SES jobs (such as managers, executives, and civil servants), medium-SES jobs (office workers, teachers, etc.) low-SES jobs (such as construction workers and waiters), and "at home" (houseworkers, the unemployed, or the retired). Parents' educational attainment is divided into three categories: below, equivalent to, or above upper secondary education. Both are computed separately for mothers and fathers. We measure the possession of educational resources in the household by the number of books in the house. In addition, in our robustness exercises we also use a comprehensive SES index, obtained by a principal component analysis, computed by INVALSI and similar to the one used by the OECD for the PISA test.

[14] Given that this information is only available for public schools, we lose observations corresponding to private schools (422 students). The number is small because in the Italian school system the large majority of upper secondary schools are public and private schools account for less than 4% of enrolled students (Ministero dell'Istruzione, 2014).

The two additional outcomes described at the bottom of Figure 1 relate to the educational aspirations of students. We compute a dummy for whether the student aims to obtain at least a bachelor's degree and a dummy for whether the student wishes to achieve a higher level of education than their own parents. This is equal to one if the student intends to obtain at least a bachelor's degree and their parents do not have one. This second variable specifically aims at capturing the student ambitions for intergenerational upward mobility. The two plots show that the share of students with higher educational ambitions increases from 2011 to 2015 by around a third. These differences are statistically significant at conventional levels.

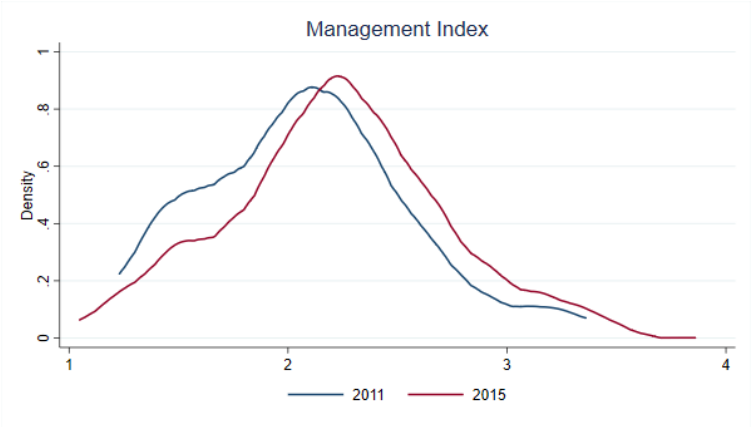
Figure 2 plots the kernel density distribution of the overall management index of SPs in 2011 and 2015. In both years we observe a susbtantial dispersion of the management scores across schools. Compared to 2011, the 2015 distribution is shifted to the right, and suggests a decrease of the share of SPs adopting poor managerial practices. This shift may be explained by the fact that, in 2011, the government held a national competition for SPs where, for the first time, an important element of screening was managerial ability. This

Figure 1: Distribution of the dependent variables, by year



Note: The plots at the top report the kernel distributions of the student standardized test scores in math (left) and language (right), by year. The vertical bars in the plots at the bottom show the average share of students wishing to achieve at least a degree (left), or at least their parents' educational attainment (right), by year. Vertical black lines refer to confidence intervals. Blue lines and bars refer to 2011, red lines and bars refer to 2015. In all cases, we reject the null hypothesis that the distributions in 2011 and 2015 are equal.

Figure 2: Distribution of the overall management index, by year



Note: The plot reports the kernel distribution of the overall management index, by year. The blue line refers to 2011, the red line refers to 2015. We reject the null hypothesis that the distributions in 2011 and 2015 are equal.

implies that newly appointed SPs have on average better managerial skills than before.

Table B1 shows that the increase in the managerial score between 2011 and 2015 holds for all areas but *People*, which has consistently low scores in both years. As discussed above, this is a reflection of the Italian institutional features and the survey design. In fact, the WMS measures the quality of managerial practices actually adopted by SPs in each school, not their managerial abilities *per se*. As argued by Di Liberto, Schivardi, and Sulis (2015), compared to the other dimensions, human resource management is the area with the highest degree of institutional constraints. Also, institutional constraints have not changed between 2011 and 2015.

3 Empirical framework and identification

While there is widespread evidence that managerial practices correlate positively with performance in a variety of settings, moving from correlation to causation has proven difficult. This difficulty is due to the possible presence of unobserved heterogeneity correlated with both performance and managerial practices. In our specific context, it is possible that SPs who implement better managerial practices may self-select into the best schools, which are typically located in catchment areas with high socio-economic status and better-performing students. The opposite could also be true, with better SPs assigned to more challenging schools. Cross-sectional estimates cannot rule out all endogeneity concerns, as finding suitable instruments for managerial practices is also difficult. We follow a different strategy: given the structure of our data, we use school fixed effects to control for all time-invariant

unobserved heterogeneity.

Specifically, we study the effects of managerial practices on student achievements using the following regression framework:

$$Y_{ijt} = \alpha + \beta MI_{jt} + \gamma X_{ijt} + \delta Z_{jt} + \theta_j + \eta_{r(j)t} + \upsilon_{ijt}, \quad (1)$$

where Y_{ijt} is an indicator of performance or aspirations of student i attending school j in year t, MI_{jt} is the managerial quality index for school j in year t, X_{ijt} is a set of individual student controls, Z_{jt} are school and SP controls. The model also includes school and area-year fixed effects (θ_j and $\eta_{r(j)t}$, respectively, where $r(j)$ is the area in which school j is located),[15] while υ_{ijt} is the error term. Standard errors are clustered at the school level. The inclusion of school fixed effects ensures that time-invariant school heterogeneity does not bias our estimates of managerial practices. In fact, we are measuring how student outcomes at the school level change with managerial practices, where changes in the latter are related to changes in the SP.

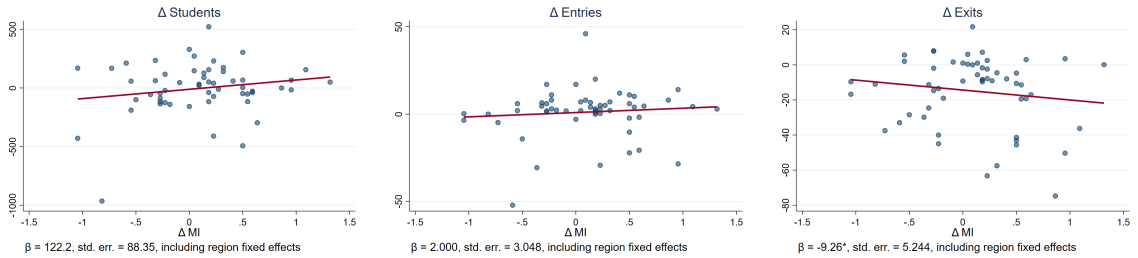
School fixed effects address what we see as the main endogeneity concern. However, other elements can threaten our identification framework. The first is time-varying shocks to student performance potentially related to changes in managerial practices. We see two potential channels: change in the pool of teachers and of students. As far as the teachers are concerned, one of the effects of better managerial practices is exactly to be able to attract and retain the best teachers. The *People* section of the survey investigates exactly such practices. So any improvement of student outcomes coming from this channel can be attributed to managerial practices themselves.

The change in the student pool, instead, could possibly undermine our identification. However, we deem it unlikely that the student pool changes so quickly following a change in the SP. Moreover, we can directly control for the socio-economic status of students in the regressions, as well as check if changes in such variable at the school level is correlated with changes in managerial practices. We will show that this is not the case.

Indeed, when we consider the sub-sample of switcher schools and plot the change in the management index versus the change in the total number of students, we obtain a positive but not significant correlation (left panel, Figure 3). This slight increase is driven by a decrease in the number of students who transfer to another school (ΔExits), rather than an increase in the number of students who transfer from another school (ΔEntries). This

[15] Areas are defined in terms of the 4 Italian macro-regions: North-West, North-East, Centre, South and islands. The full list of control variables and their corresponding descriptive statistics is in Table B1 under the headings "Student characteristics", "School characteristics", and "SP characteristics". If a covariate is included as categorical, the Table reports statistics for each category.

Figure 3: Change in management index and in number of students, entries and exits



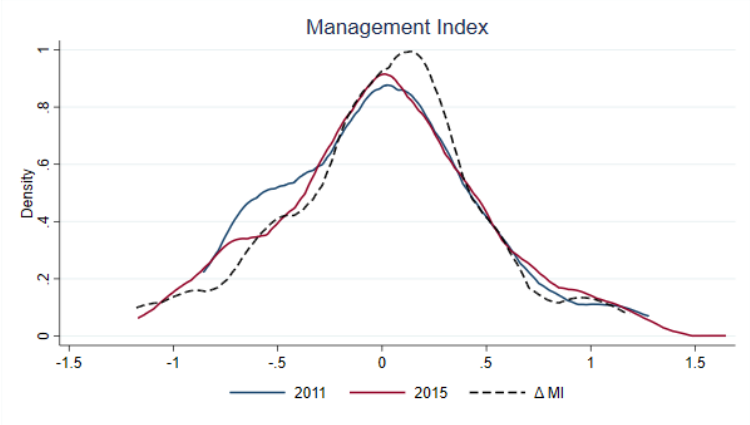
Note: Correlation between the change in management index and the change in the school's total number of students, total number of entries (students who transfer from another school) and total number of exits (students who transfer to another school). Sample of switcher schools only (56 schools). Each plot reports the correlation coefficient and robust standard errors, conditional on region fixed effects. * p<.10.

suggests that, if anything, we can exclude that a positive change in the management index is associated to the SP being able to attract better students from competing schools or by screening students based on ability (e.g., changing the school retention policy and increasing the share of students that are not promoted to the next grade level, thus inducing the worst performing students to transfer to other schools).

Another possibility is that managerial practices are correlated with other policies that SPs introduce. This too is not likely to be the case. Specifically, SPs do not have the ability to influence their school's mission or curricula as these are centrally established by the Ministry of Education, which also provides guidance on the specific skills and knowledge that students are expected to acquire. Therefore, any improvements in student outcomes can be reasonably attributed to changes in managerial practices, which represent the key area where SPs can have an impact.

One last potential problem with fixed effects estimates is that the time series variability might be limited, decreasing the signal-to-noise ratio of the explanatory variable and leading to imprecise estimates. Again, this is not likely to be the case in our setting. First, the WMS is based on long and detailed interviews that accurately measure the quality of managerial practices. Second, as explained above, SPs can deeply affect the school organization, so the change of a SP can bring about substantial changes in managerial practices. To test this conjecture, in Figure 4 we plot the distributions of the management index in 2011 and in 2015, and the change in the management index for the sample of switchers (ΔMI). To ease comparability, we center the three distributions at zero. Reassuringly, they are very similar. In particular, the degree of variability that characterizes the change in the management index is similar to the cross-sectional variability within each year, suggesting

Figure 4: Demeaned distribution of the management index and of the change in the management index



Note: The solid lines refer to the distribution of the demeaned management index in 2011 (blue) and 2015 (red). The dashed black line shows the distribution of the change in management index (ΔMI) in the sample of switchers only (i.e. 56 schools where the SP has changed between 2011 and 2015).

that our fixed effects estimation should not suffer from low variability problems.

4 Results

In this section, we first describe the main results and then discuss selection, robustness and heterogeneity of the effects.

4.1 Main results

In Table 1 we present the main results on the effect of managerial practices, measured by the overall management index, on student performance. The dependent variables are the standardized test scores in mathematics (Panel A) and in Italian language (Panel B).

In Column 1 we begin with a parsimonious specification that only includes our management index and a base set of school and area-by-year fixed effects. In subsequent columns, we augment our set of controls and include the characteristics of the SP, the school, and the student and their family. Our saturated model in Column 4 thus includes a large set of potential determinants of student outcomes. Including additional controls is important, as the adoption of good managerial practices may be correlated with the SP's observable characteristics, such as age, tenure and experience, or with other school or student characteristics. Not including such controls would result in attributing other sources of variation to managerial practices.

Table 1: Effect on student scores

Panel A	(1)	(2)	(3)	(4)	(5)
Dep. variable	Score: Math				
Management index	2.628	3.360**	3.490**	3.729**	2.489**
	(1.719)	(1.675)	(1.610)	(1.648)	(1.212)
Observations	23,514	23,514	23,514	23,514	23,514
School-year clusters	309	309	309	309	309
R-squared	0.331	0.332	0.335	0.384	0.258
School FE	✓	✓	✓	✓	
Area × Year FE	✓	✓	✓	✓	✓
SP characteristics		✓	✓	✓	✓
School characteristics			✓	✓	✓
Student characteristics				✓	✓
Panel B	(1)	(2)	(3)	(4)	(5)
Dep. variable	Score: Language				
Management index	1.441	2.084*	2.233*	2.420**	0.734
	(1.187)	(1.248)	(1.186)	(1.191)	(0.811)
Observations	23,436	23,436	23,436	23,436	23,436
School-year clusters	309	309	309	309	309
R-squared	0.411	0.412	0.413	0.451	0.361
School FE	✓	✓	✓	✓	
Area × Year FE	✓	✓	✓	✓	✓
SP characteristics		✓	✓	✓	✓
School characteristics			✓	✓	✓
Student characteristics				✓	✓

Note: * p<.10 ** p<.05 *** p<.01. Robust standard errors clustered at school level. All regressions include area-year fixed effects, where areas are North-West, North-East, Centre, and South. Student characteristics comprise gender, immigration status, class size, whether the student is regular or younger than regular, the number of books in the house, mother's and father's education and occupational status. SP characteristics encompass age, gender, overall tenure within the school, tenure as SP within the school, additional job experience outside the school, whether the school was chosen, and whether the SP manages multiple schools. School characteristics are the number of pupils, teachers, administrative staff, female teachers, and permanent teachers.

For both mathematics and language, in Column 1 we obtain a positive but marginally statistically insignificant coefficient. However, as we add more controls the estimated coefficient increases and becomes statistically significant at conventional levels for both outcomes. Our estimates are also quantitatively important. Our preferred specification in Column 4 implies that a standard deviation increase in the management index increases average student test score results by 0.091 standard deviations in mathematics and 0.066 standard deviations in language. The magnitude is in line with previous evidence. Exploiting an IV strategy on the first wave of the data used here, Di Liberto, Schivardi, and Sulis (2015) find that a one-standard deviation increase in the management score causes a 0.10 standard deviations increase in student math achievement. Tavares (2015) finds that participation to a school management program increases math performance of Brazilian eighth graders

by approximately 0.14-0.22 standard deviations.[16]

As mentioned in Section 3, our fixed effects model should capture an important source of heterogeneity across schools which may potentially bias our estimates of the managerial index. In order to verify the direction of the bias, in Column 5 we also estimate our model excluding school fixed effects (but including all other controls). The comparison between the two specifications suggests that the inclusion of school fixed effects allows to correct for the presence of a downward bias in the pooled cross-sectional estimates, consistent with the hypothesis that SPs with better managerial capabilities tend to be assigned to schools with low performing students.[17]

Next, we consider the effect of managerial practices on student educational aspirations towards pursuing tertiary education. Results are reported in Table 2. The first dependent variable is a dummy for whether the student aims to obtain at least a bachelor's degree (Panel A), while the second one measures whether the student wishes to achieve a higher level of education with respect to their own parents (Panel B). Results again suggest a positive and statistically significant effect of managerial quality, which emerges more clearly when including SP, school and student controls. Estimates in Column 4 of Table 2 indicate that a standard deviation increase in management index increases average student aspirations by a third of a standard deviation. As for the previous two outcomes, estimates would be considerably biased towards zero in the absence of fixed effects (Column 5).

The comparison between Columns 4 and 5 in both Tables 1 and 2 suggests the presence of a negative selection of SPs in terms of managerial capabilities into the highest performing schools. The process through which SPs are assigned to schools in the Italian system can help us explain this result. First, actual assignments are made by the Regional School Authorities (RSAs) which try to accommodate SPs requests but have to fill in the positions for the schools that were not chosen by any SP. These are likely to be the most 'difficult' schools, for which the RSA might try to allocate more capable SPs. Second, anecdotal evidence suggests that the most prestigious schools are assigned by the RSA to older SPs as a sort of 'end-of-career benefit'. While older SPs may be more experienced, they tend

[16] However, correlations estimated in existing works tend to be larger: Bloom et al. (2015) find substantial cross-country heterogeneity, such that a one-standard deviation increase in the managerial index is associated with an increase in pupil outcomes between 0.2 and 0.4 standard deviations. Additionally, the literature that uses value-added models to estimate the SP contribution to improving student outcomes finds heterogeneous results: a one-SD increase in principal value-added increases student achievement by 0.05 to 0.20 SD (Branch, Hanushek, and Rivkin, 2012; Dhuey and Smith, 2018; Bartanen and Husain, 2022).

[17] Here, estimates imply that a standard deviation increase in management index increases average student test score results by 0.061 standard deviations in mathematics (statistically significant at 5% level) and 0.020 standard deviations in language (not statistically different from zero).

to be less trained and rely less on formal managerial procedures than younger SPs. Third, cohort effects may play a role, since newly appointed SPs have a stronger background on management due to recent national competitions putting more emphasis on managerial capabilities during the selection process. The newly selected SPs might tend to be assigned to more difficult schools.

4.2 Selection

We further investigate the presence of potential selection and confounding issues that might generate bias in our estimates regardless of the inclusion of the school fixed effects.

First, in Table 3 we use the balanced sample of switcher schools and show that the managerial quality of the new SPs in 2015 is uncorrelated to the scores of the students enrolled in the same school in 2011, both for math and language, nor to their aspirations. Moreover, we also find that the socio-economic indicator SES is unrelated to the current managerial quality (Column 1 of Table 4).

Table 2: Effect on student aspirations

Panel A	(1)	(2)	(3)	(4)	(5)
Dep. variable	Aspirations: Degree or higher				
Management index	0.016	0.031	0.033*	0.036**	-0.002
	(0.022)	(0.020)	(0.019)	(0.018)	(0.018)
Observations	23,514	23,514	23,514	23,514	23,514
School-year clusters	309	309	309	309	309
R-squared	0.285	0.286	0.287	0.317	0.262
School FE	✓	✓	✓	✓	
Area × Year FE	✓	✓	✓	✓	✓
SP characteristics		✓	✓	✓	✓
School characteristics			✓	✓	✓
Student characteristics				✓	✓
Panel B	(1)	(2)	(3)	(4)	(5)
Dep. variable	Aspirations: Own edu > parents' edu				
Management index	0.032	0.047***	0.048***	0.035**	-0.006
	(0.020)	(0.018)	(0.017)	(0.015)	(0.014)
Observations	23,514	23,514	23,514	23,514	23,514
School-year clusters	309	309	309	309	309
R-squared	0.100	0.101	0.101	0.327	0.288
School FE	✓	✓	✓	✓	
Area × Year FE	✓	✓	✓	✓	✓
SP characteristics		✓	✓	✓	✓
School characteristics			✓	✓	✓
Student characteristics				✓	✓

Note: * p<.10 ** p<.05 *** p<.01. Robust standard errors clustered at school level. All regressions include area-year fixed effects, where areas are North-West, North-East, Centre, South. The set of additional controls is identical to that included in Table 1.

Table 3: 2011 student scores and managerial skills of SP in 2015

Dep. variable	(1)	(2)	(3)	(4)
	Management index in 2015			
Avg score in 2011: Math	0.005			
	(0.007)			
Avg score in 2011: Language		0.003		
		(0.006)		
Avg aspirations in 2011: Degree or higher			-0.072	
			(0.263)	
Avg aspirations in 2011: Own edu > parents' edu				-0.074
				(0.592)
Observations	56	56	56	56
R-squared	0.096	0.089	0.088	0.087

Note: * p<.10 ** p<.05 *** p<.01. Robust standard errors. All regressions include area fixed effects, where areas are North-West, North-East, Centre, South. Dependent variable is the SP's management index in 2015. Average scores are measured in 2011. Balanced sample of switcher schools (n=56).

Second, given that our effect is identified by the "switcher" schools, i.e. those where the SP has changed between 2011 and 2015, we need to rule out that our results are driven by additional potential unobservable confounders that influence both SP turnover and student performance. In columns 2-5 of Table 4 we show the absence of a relationship between a school's switcher status and the student outcomes.

In Appendix Table B2 we assess the impact of various school-specific attributes measured in 2011 on the probability that the SP changes between 2011 and 2015. The coefficients,

Table 4: Effect on student scores and aspirations, selection of SP

Dep. variable	(1)	(2)	(3)	(4)	(5)
		Scores		Aspirations	
	SES	Math	Language	Degree or higher	Own edu > parents' edu
Management index	-0.009				
	(0.025)				
SP has changed		-0.337	-0.341	0.007	0.004
		(1.112)	(0.796)	(0.018)	(0.014)
Observations	23,318	23,514	23,436	23,514	23,514
R-squared	0.783	0.255	0.360	0.262	0.288
School-year clusters	308	309	309	309	309
Sample	Whole	Whole	Whole	Whole	Whole
School FE	✓				
Area × Year FE	✓	✓	✓	✓	✓
SP characteristics	✓	✓	✓	✓	✓
School characteristics	✓	✓	✓	✓	✓
Student characteristics	✓	✓	✓	✓	✓

Note: * p<.10 ** p<.05 *** p<.01. Robust standard errors clustered at school level. All regressions include area-year fixed effects, where areas are North-West, North-East, Centre, South. *SP has changed* identifies the school's switcher status. The set of additional controls is identical to that included in Table 1.

which are all estimated in separate regressions, imply that turnover is associated to the SP's age and tenure and by managing multiple schools. In all cases the positive relation is expected, because older or more experienced SPs are more likely to retire, and the management of multiple schools is temporary by definition as it is usually allowed in exceptional cases and for limited periods of time. A positive coefficient associated with the number of administrative staff might also suggest that bigger schools are more likely to change SP, but this is not confirmed by the coefficients referred to the number of students and teachers. Importantly, the probability of changing SP is not statistically related to the student characteristics.

4.3 Robustness

Next, we test the robustness of our results to different model specifications and sample selections. We report the results in Appendix Tables B3 for the standardized test scores and B4 for aspirations. In the first three columns we allow for differences in the model specification. In Column 1 we replace the information on the number of books in the house and the parents' education and job status with the standardized student's socioeconomic status (SES) index. In Column 2 we allow for non-linearity in the SP's tenure within the school. Column 3 includes time-varying fixed effects at a more geographically disaggregated level, i.e. province-year fixed effects. Overall, this evidence confirms that management quality is an important input of our estimated education production function.

In Columns 4 and 5 we further check if results are confirmed within different sample selections. In Column 4 we exclude the few observations related to temporary SP posts (namely, when the SP manages multiple schools) and find that the coefficient for the overall managerial index remains almost the same. In Column 5 we verify that the estimates are robust to excluding SPs close to retirement (namely, those in the 4th quartile of the age distribution). As previously mentioned, in Italy SPs are often assigned the more prestigious or preferred schools at the end of their career as a form of recognition. Therefore, it is possible that the observed main effect is primarily driven by the more tenured and experienced SPs being assigned to the best performing schools. This is unlikely to be the case, as our OLS estimates, if anything, indicate that the bias goes in the opposite direction, and Column 5 results confirm that the estimates are robust to excluding older SPs.

Finally, in Appendix Table B5 we specifically examine if considering only the balanced sub-samples of schools affects our results. In Columns 1 and 4 we restrict our main sample of 309 school-year observations to a balanced panel of 236 school-year observations (i.e. 118

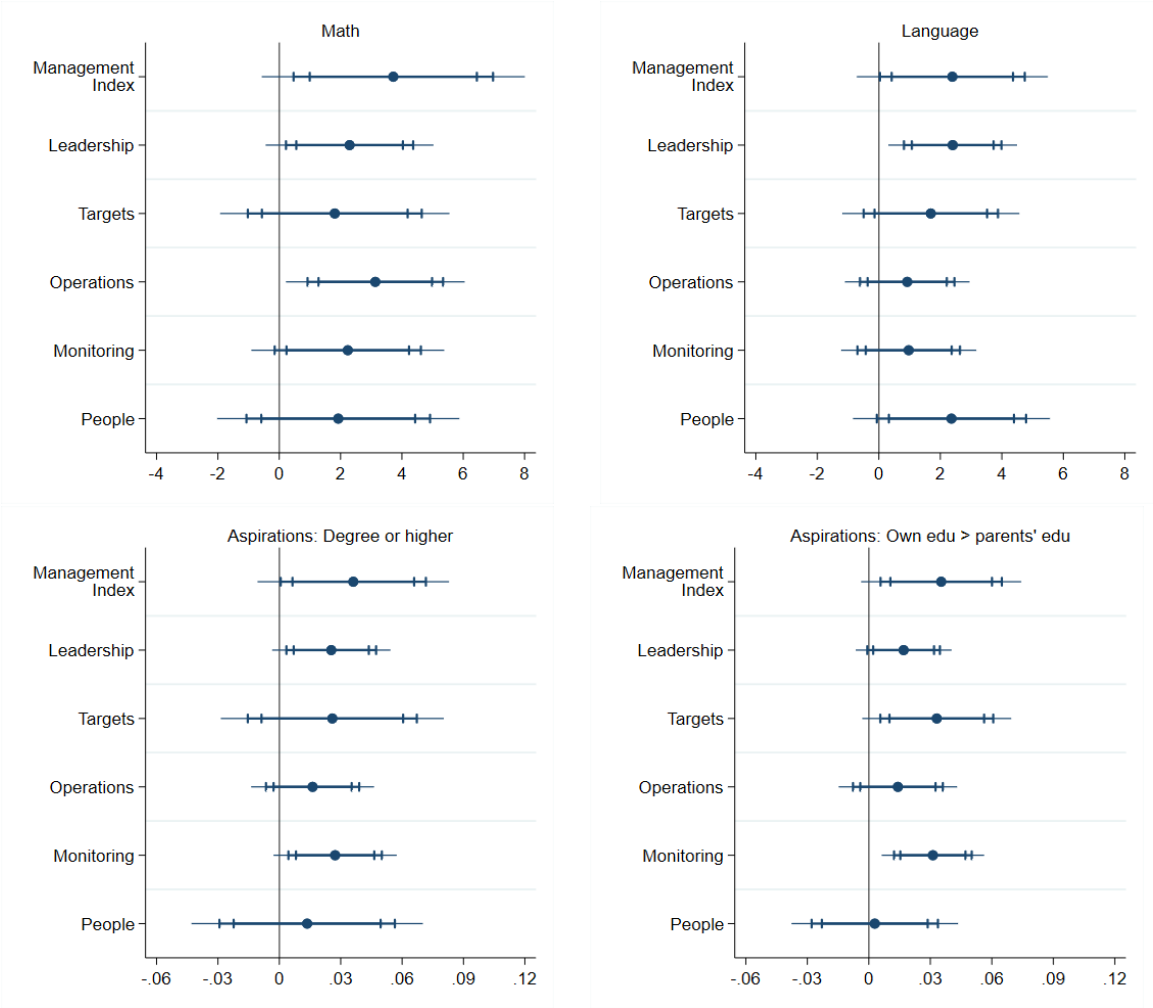
schools observed both in 2011 and 2015). In this case, we exclude schools participating in the first wave of data collection, but whose new SP did not answer the interview in 2015, and also schools that could not be matched with administrative data in either 2011 or 2015. Here, we could expect some sample selection bias, as these schools may be worse in terms of student characteristics or managerial practices, or both. However, this source of attrition does not appear to impact our analysis, as our main results are confirmed across all our four outcome variables. In Columns 2 and 5 we exclude schools that did not change their SP between the two waves, leaving us with 140 school-year observations. As our coefficient of interest is always identified by these schools, as expected, we find consistent evidence that management is positively and significantly related to our outcomes of interest, with larger coefficients for math and student aspirations. Last, Columns 3 and 6 further restrict the sample to the schools where the SP has changed between 2011 and 2015 and the management scores were collected when each SP was running the school (i.e. the 56 switcher schools). In other words, as described in Section 2.1, in this sample we are not exploiting the assumption that managerial quality is time-invariant, since we use only those schools where we observe SP turnover between the two waves and where the old SP has been interviewed in 2011 and the new in 2015. Even in this case, despite the loss in statistical power due to a significant drop in sample size, the estimated coefficients are all in line with the main results.

4.4 Heterogeneity

Our overall management index summarizes different dimensions of managerial skills, namely Leadership, Targets, Operations, Monitoring and People. As mentioned in Section 2.3, from 2011 to 2015 the quality of managerial practices increases for all dimensions, with the relevant exception of People, which remains substantially stable across years. In Figure 5 we show the effect of each component on the four outcomes estimated separately. Overall, all dimensions contribute fairly homogeneously to the increase in student scores and aspirations. The main exception are People, where estimates, despite always positive, are not statistically significant at conventional levels and, to a lesser extent, Targets.

Finally, we investigate the issue of which students (based on their academic performance) benefit more from effective managerial practices. Figure 6 reports coefficients estimated via quantile regression models for the 10th, 25th, median, 75th and 90th percentiles. Results suggest that the effect of managerial practices is homogeneous across the distribution of the student math score, while the effect on language scores is stronger at the left tail of the

Figure 5: Effects by management dimension



Note: All coefficients are estimated in separate regressions, based on the model as from equation 1. All regressions include school and area-year fixed effects, as well as controls for the characteristics of SP, school and students. Robust standard errors clustered at school level. Confidence intervals at 90, 95 and 99 percent level are shown.

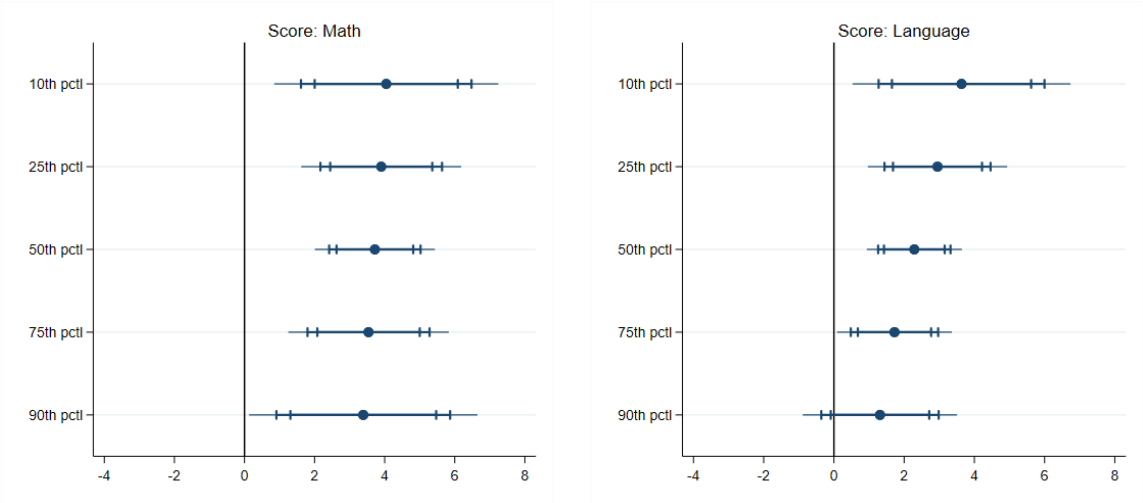
student outcome distribution, that is, low performing students benefit the most.[18]

5 Conclusions

In this paper we contribute to the still scant literature on the effects of SP managerial practices on student performance. Our key methodological contribution is to account for

[18] Moreover, when we interact our main regressor MI with a relevant set of observables from the different controls, we find that the effect of the management index is fairly homogeneous with respect to the SP, the school or the student characteristics. If anything, the positive effect of managerial practices on the student scores are reinforced the higher the SP's tenure, suggesting a role for experience within the same school.

Figure 6: Effects on scores by percentiles



Regressions estimated using the Stata command `xtqreg` by Machado and Santos Silva (2019). They include area-year fixed effects and controls for SP, school and average student characteristics. Confidence intervals at 90, 95 and 99 percent level are shown.

unobserved heterogeneity and selection through a school fixed effects regression framework, using the fact that, for a number of schools, we exploit the change in the SP and the resulting variation in managerial practices. This is an important step forward in terms of interpreting the results in causal terms.

We find that managerial practices positively and substantially impact student performance and educational aspirations. Our estimates imply that a standard deviation increase in the management index increases average student test score results by 0.09 standard deviations in mathematics and 0.07 standard deviations in language. As for aspirations, the probability that students wish to achieve at least a bachelor's degree or intend to obtain a higher educational attainment than their parents increases by roughly a third of a standard deviation. We show that our results are robust to several modifications of the empirical framework and to controlling for student, school and SP characteristics.

Our findings imply that policies directed at improving the quality of managerial practices in schools will positively affect student academic success. Our results on student educational aspirations further suggest that well-run high schools not only improve cognitive skills but also have positive effects on psychological traits considered important drivers of student further educational attainments and, in particular, on tertiary education ambitions.

Overall, our evidence indicates that increasing the quality of managerial practices in schools is a promising area of intervention to improve student outcomes. We show that such

practices depend primarily on SPs. This analysis also implicitly suggests that an accurate initial screen on managerial capabilities, as done in the Italian SPs selection system, can eliminate potential low-performing SPs and improve the school's learning environment. Moreover, given that the literature shows that managerial practices can be taught (Bloom et al., 2013; Bruhn, Karlan, and Schoar, 2018) and that SPs are a population substantially smaller than that of teachers, interventions aimed at improving their managerial capabilities might be a cost-effective way to improve student achievements.

References

Angrist, Joshua, Peter Hull, and Christopher R Walters. 2022. "Methods for Measuring School Effectiveness." Working Paper 30803, National Bureau of Economic Research.

Bartanen, Brendan. 2020. "Principal quality and student attendance." *Educational Researcher* 49 (2):101–113.

Bartanen, Brendan and Jason A Grissom. 2023. "School principal race, teacher racial diversity, and student achievement." *Journal of Human Resources* 58 (2):666–712.

Bartanen, Brendan and Aliza N. Husain. 2022. "Connected networks in principal value-added models." *Economics of Education Review* 90:102292.

Björklund, Anders and Kjell G Salvanes. 2011. "Education and family background: Mechanisms and policies." In *Handbook of the Economics of Education*, vol. 3. Elsevier, 201–247.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen. 2019. "What Drives Differences in Management Practices?" *American Economic Review* 109 (5):1648–83.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does management matter? Evidence from India." *The Quarterly Journal of Economics* 128 (1):1–51.

Bloom, Nicholas, Christos Genakos, Raffaella Sadun, and John Van Reenen. 2012. "Management practices across firms and countries." *The Academy of Management Perspectives* 26 (1):12–33.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, and John Van Reenen. 2015. "Does management matter in schools?" *The Economic Journal* 125 (584):647–674.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2016. "Management as a Technology?" Working Paper 22327, National Bureau of Economic Research.

Bloom, Nicholas and John Van Reenen. 2007. "Measuring and Explaining Management Practices Across Firms and Countries." *The Quarterly Journal of Economics* 122 (4):1351–1408.

———. 2010. "New Approaches to Surveying Organizations." *American Economic Review* 100 (2):105–09.

———. 2011. *Human Resource Management and Productivity, Handbook of Labor Economics*, vol. 4, chap. 19. Elsevier, 1697–1767.

Böhlmark, Anders, Erik Grönqvist, and Jonas Vlachos. 2016. "The headmaster ritual: The importance of management for school outcomes." *The Scandinavian Journal of Economics* 118 (4):912–940.

Branch, Gregory F, Eric A Hanushek, and Steven G Rivkin. 2012. "Estimating the effect of leaders on public sector productivity: The case of school principals." Working Paper 17803, National Bureau of Economic Research.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar. 2018. "The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in Mexico." *Journal of Political Economy* 126 (2):635–687.

Burgess, Simon. 2016. "Human capital and education: The state of the art in the economics of education." *IZA Discussion Paper* 9885.

Carlana, Michela, Eliana La Ferrara, and Paolo Pinotti. 2022. "Goals and gaps: Educational careers of immigrant children." *Econometrica* 90 (1):1–29.

Checchi, Daniele and Luca Flabbi. 2013. "Intergenerational Mobility and Schooling Decisions in Germany and Italy: The Impact of Secondary School Tracks." *Rivista di Politica Economica* (3):7–57.

Dalton, Patricio S, Sayantan Ghosal, and Anandi Mani. 2016. "Poverty and aspirations failure." *The Economic Journal* 126 (590):165–188.

Dhuey, Elizabeth and Justin Smith. 2018. "How school principals influence student learning." *Empirical Economics* 54:851–882.

Di Liberto, Adriana, Fabiano Schivardi, and Giovanni Sulis. 2015. "Managerial practices and student performance." *Economic Policy* 30 (84):683–728.

Fryer, Jr, Roland G. 2017. "Management and Student Achievement: Evidence from a Randomized Field Experiment." Working Paper 23437, National Bureau of Economic Research.

Fryer, Roland G. Jr. 2014. "Injecting Charter School Best Practices into Traditional Public Schools: Evidence from Field Experiments." *The Quarterly Journal of Economics* 129 (3):1355–1407.

Genicot, Garance and Debraj Ray. 2017. "Aspirations and inequality." *Econometrica* 85 (2):489–519.

Grissom, Jason and Susanna Loeb. 2011. "Triangulating Principal Effectiveness: How Perspectives of Parents, Teachers, and Assistant Principals Identify the Central Importance of Managerial Skills." *American Educational Research Journal* 48 (5):1091–1123.

Guyon, Nina and Elise Huillery. 2021. "Biased aspirations and social inequality at school: Evidence from french teenagers." *The Economic Journal* 131 (634):745–796.

Hanushek, Eric A. 2011. "The economic value of higher teacher quality." *Economics of Education Review* 30 (3):466–479.

Hanushek, Eric A and Ludger Woessmann. 2012. "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation." *Journal of Economic Growth* 17:267–321.

Haveman, Robert and Barbara Wolfe. 1995. "The determinants of children's attainments: A review of methods and findings." *Journal of Economic Literature* 33 (4):1829–1878.

Leaver, Clare, Renata Freitas Lemos, and Daniela Scur. 2019. "Measuring and explaining management in schools: New approaches using public data." *World Bank Policy Research Working Paper* (9053).

Lemos, Renata, Karthik Muralidharan, and Daniela Scur. 2021. "Personnel Management and School Productivity: Evidence from India." Working Paper 28336, National Bureau of Economic Research.

Lucas, Robert E. Jr. 1988. "On the mechanics of economic development." *Journal of Monetary Economics* 22 (1):3–42.

Machado, José AF and JMC Santos Silva. 2019. "Quantiles via moments." *Journal of Econometrics* 213 (1):145–173.

Mankiw, N Gregory, David Romer, and David N Weil. 1992. "A contribution to the empirics of economic growth." *The Quarterly Journal of Economics* 107 (2):407–437.

Ministero dell'Istruzione, dell'Università e della Ricerca. 2014. *The Italian Education System.*

OECD. 2019. *PISA 2018 Results (Volume II).*

Pagani, Laura, Simona Comi, and Federica Origo. 2021. "The effect of school rank on personality traits." *Journal of Human Resources* 56 (4):1187–1225.

Robinson, Viviane M. J., Claire A. Lloyd, and Kenneth J. Rowe. 2008. "The Impact of Leadership on Student Outcomes: An Analysis of the Differential Effects of Leadership Types." *Educational Administration Quarterly* 44:635–674.

Romero, Mauricio, Justin Sandefur, and Wayne Aaron Sandholtz. 2020. "Outsourcing education: Experimental evidence from Liberia." *American Economic Review* 110 (2):364–400.

Tavares, Priscilla. 2015. "The impact of school management practices on educational performance: Evidence from public schools in São Paulo." *Economics of Education Review* 48 (C):1–15.

A Data Appendix

Between February and May 2011, we carried out the initial round of 341 double-blind interviews with Italian SPs. The interviews were conducted by the Italian team of five analysts and two senior managers who received training and monitoring from the international WMS team. The 2011 sample of 341 schools-SPs is representative of the population of Italian upper secondary schools. It was randomly drawn from the population of Italian upper secondary schools, and stratified by type of school and geographical location.[19]

The follow-up in 2015 strictly followed the rules and procedures of the first wave of the WMS team, and it was characterized by a large degree of coordination across the two waves. The 2015 team consisted of six new interviewers who were trained following the same WMS protocol, and worked under the supervision of the same two senior managers of 2011. The data collection process started in January and ended in June 2015.

In 2015, we first detected which schools of the 2011 sample changed the SP in 2015 or not. Out of the initial 2011 sample of 341 schools, we identified 127 schools managed by the same SP, while in 190 schools we observed SP turnover (turnover schools). For 24 schools we could not collect any information.

We did not conduct a second interview with the 127 SPs who had already answered the WMS in 2011 and did not change the school they managed. We assume that managerial quality is constant, and assign the same information collected during the 2011 interviews in 2015. However, the time-varying characteristics (such as SPs age, tenure etc.) of the dataset were adjusted accordingly.

In 2015 we conducted the WMS double-blind interviews only with the new SPs of the schools that changed SPs between the two waves. We collected information on managerial practices for 114 of them. Thus, for each school in this group, we have two independent measures of management collected in 2011 and 2015, interviewing the two different SPs who were in charge of running the same school. We label this sub-group as "switcher schools".

In order to increase the size of our sample, in 2015 we also select a new sample of 28 schools that were not previously interviewed. To this new sample we applied what we call the "chain" method. Table A1 below helps to illustrate how it works. We proceeded in steps: i) we interviewed the SP of school A in 2015 (SP Y in Table A1) and assigned the managerial score to school A in 2015; ii) during the interview, we asked SP Y about the school they were in 2011 (call it school B) and assigned the managerial score of SP Y to school B in 2011; iii) we then called school B, asked the new SP to participate in the

Table A1: The chain method

School managed in	2011	2015	
School A	SP X	**SP Y**	(2015 interview)
School B	**SP Y**	SP W	(2015 interview)

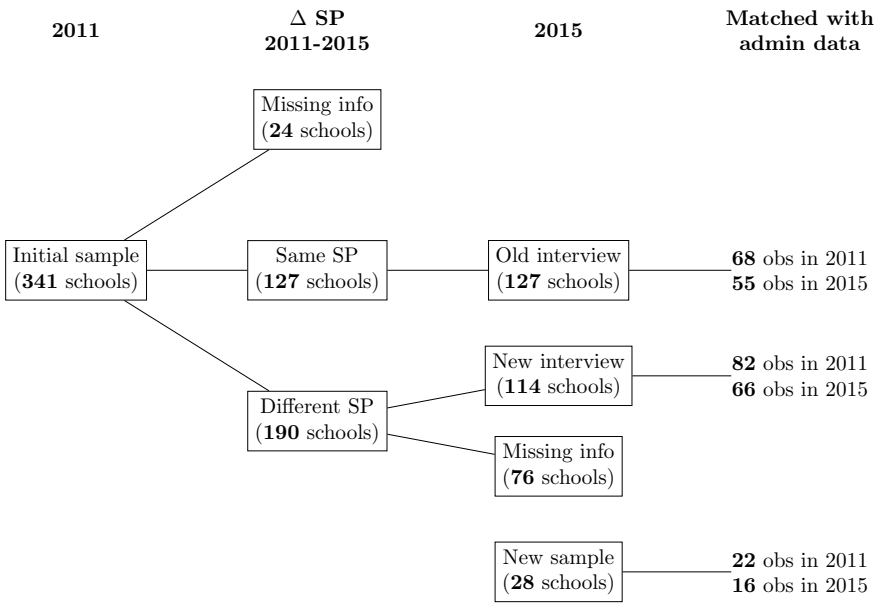
[19] Italy has a tracked upper secondary school system that consists of different school types offering different curricula. For additional information on the data collection procedures of the 2011 interviews see Di Liberto, Schivardi, and Sulis (2015).

project, interviewed them (SP W in Table A1), and assigned the managerial score of SP W to school B in 2015. Again, these are schools that changed SP between 2011 and 2015, or turnover schools, but we assume that managerial quality is constant between 2011 and 2015. In other words, although for this group of schools the WMS double-blind interviews were both conducted in 2015, we still obtain two measures of management from two different SPs who managed the school in 2011 and 2015. This is different from the group of 114 schools mentioned above, where interviews were conducted separately in 2011 and 2015, and assigned to each school accordingly.

To sum up, the WMS interviews conducted during the two waves yield a panel of schools such that, out of the initial 2011 sample of 341 schools, in the 114 turnover schools in 2015 we interviewed again SP, while 127 that did not change SP were assigned the same management score as 2011. Finally, 28 schools are added to this final panel sample, but their SPs were interviewed in 2015. Thus, at this stage, we drop from the sample 24 schools for which we did not recover information, plus 76 turnover schools we were not able to interview in 2015. In total, with our two waves of WMS interviews, we end up with 269 schools observed twice. The construction of the WMS panel is outlined in Figure A1.

The merge with the administrative INVALSI student data and the MIUR administrative information on schools resulted in the loss of some observations. This was mainly due to discrepancies in the school administrative identifier or to missing or inconsistent data in the administrative records from INVALSI or MIUR. Of the 114 that changed SP between 2011 and 2015, we are left with 68 observations in 2011 and 55 in 2015; of the 127 that did not change SP, we observe 82 in 2011 and 66 in 2015; of the additional sample of 28 schools, we

Figure A1: WMS interviews: construction of panel of schools



observe 22 in 2011 and 16 in 2015 (see last column of Figure A1).

In sum, in our analysis we mainly use the largest sample we obtain from the merge with the two administrative datasets which consists of an unbalanced panel of 309 schools (172 schools/SPs observed in 2011, 137 in 2015) and 23,514 observations at the student level. We also use an unbalanced panel sub-sample that only includes turnover schools where the different SPs have been interviewed in each wave when they were running the school or "switcher schools". This comprises 148 school-year observations (10,626 student-year observations).

Throughout the paper, we also present further evidence based on the balanced versions of the panel. First, we restrict our main sample of 309 school-year observations to a balanced panel of 236 school-year observations (i.e. 118 schools observed both in 2011 and 2015). Second, we exclude schools that did not change their SP between the two waves (96 school-year and 7,808 student-year observations), leaving us with only those schools that experienced an SP turnover and 140 school-year observations. This sub-sample consists of both schools recovered using the "chain method" as described above (comprising 28 school-year units and 2,776 student-year observations) and the "switcher schools" which comprises 56 schools (i.e. 112 SPs and 8,404 students). Finally, since the former are schools where the SP changed between 2011 and 2015 but for which we assign the 2011 managerial quality index based on interviews carried out in 2015, we also relax this assumption and follow a conservative approach replicating the analysis using only the sub-group of 56 "switcher schools".[20] All details are summarized in Table A2.

Table A2: Final sample of schools by sub-group and year

	Main sample			Balanced sample		
	2011	2015	Total	2011	2015	Total
SP changed ("switcher schools")	82	66	148	56	56	112
SP changed ("chain method")	22	16	38	14	14	28
Same SP in 2011 and 2015	68	55	123	48	48	96
Total	172	137	309	118	118	236

A.1 Attrition

As described above, from the first and second wave of the data collection and from the merge with the two additional administrative datasets we lose observations and in this subsection we analyze if this attrition implies also sample selection. To this end, we check if the schools we lose are similar or not along some key characteristics measured in 2011, namely, the management index, the average socio-economic status of students, their score in math and Italian language, the area where the school is located and the type of school.

The first three columns of Table A3 divide our initial representative sample of 341 schools/SPs interviewed in 2011 into three groups: the 127 schools that at the moment of

[20] In this last analysis we are excluding the so-called "chain method" schools. The different balanced samples results are in Table B5.

Table A3: Attrition, following the initial sample over time

	SP changed			Interviewed in 2015			Matched w admin data			First vs last
	No	Yes	Missing	No	Yes	Δ	No	Yes	Δ	Δ
Management Index	2.14	1.91	1.82	1.83	1.97	-.13**	1.96	1.97	-.01	.17**
SES	0.08	0.02	0.27	-0.08	0.10	-.18**	0.37	-0.00	.37***	.08
Score in Math	47.83	46.62	55.49	43.84	48.30	-4.46***	49.21	47.95	1.26	-.12
Score in Italian	67.08	65.68	59.07	63.57	66.95	-3.38**	70.08	65.73	4.36*	1.35
Region: North-West	0.26	0.25	0.29	0.16	0.32	-.16**	0.44	0.27	.17*	-.01
Region: North-East	0.24	0.16	0.03	0.07	0.23	-.16***	0.19	0.24	-.06	-.01
Region: Centre	0.13	0.17	0.33	0.17	0.18	.00	0.22	0.16	.06	-.02
Region: South	0.37	0.41	0.33	0.61	0.28	.32***	0.16	0.33	-.17*	.04
Inst. sup.	0.27	0.37	0.08	0.37	0.37	.00	0.34	0.38	-.03	-.11*
Lyceum	0.44	0.33	0.46	0.28	0.36	-.08	0.47	0.32	.15	.12*
Technical	0.20	0.16	0.38	0.20	0.14	.06	0.09	0.16	-.06	.05
Vocational	0.09	0.13	0.08	0.16	0.13	.03	0.09	0.15	-.05	-.06
Number of schools	127	190	24	76	114		32	82		

Note: Asterisks denote that the difference between the two samples is significant at *10%, **5%, ***1%*. First vs last Δ refers to the difference between the schools that did not change SP (n=127, first column) and the schools that were matched to the INVALSI and MIUR administrative data (n=82).

recall in 2015 had the same SP in charge, the 190 that have instead changed the SP, and the residual group of 24 schools for which we were not able to recover any information. In this case, our initial 24 missing schools do not seem to identify a clear pattern in terms of sample selection.

We then consider the sub-group of schools that changed the SP (the following three columns) and compare the characteristics of schools that we have been able to interview in 2015 (114) with those that went missing (76). The third column identifies whether the difference in the mean values of each variable in the two samples is significant.

In the next three columns of Table A3, we further take into account the missing observations due to the match with the administrative INVALSI and MIUR: in fact, of the initial 114 schools, after the match we lose 32 schools. Thus, we compare the characteristics of the matched schools (82 schools) with those that were not (32). Again, the third column identifies the differences in terms of initial conditions between the two samples.

Overall, Table A3 confirms the presence of a typical selection of better students/schools between the two waves of data collection. Missing schools had in 2011 on average students with a lower background and educational achievements, while we also lose schools mainly from the southern and less developed regions.

Finally, the last column of Table A3 compares the 2011 average characteristics of the final sample of the original switching schools to those that did not change their SP that we use in our empirical analysis. In particular, we evaluate the difference between the 127 that did not change SP and those that changed SP, were interviewed, and matched with administrative data (82 schools). Unlike the previous evidence, here we observe no differences in terms of student background and test scores, and only a statistically significant but negligible discrepancy in terms of managerial quality. If any, we find some asymmetries in terms of composition by school type.

B Additional tables and figures

Table B1: Descriptive statistics

	2011				2015				Δ
Variable	Obs	Mean	Median	Std. dev.	Obs	Mean	Median	Std. dev.	
Outcome variables									
Score in Mathematics	12,894	51.321	50.943	17.364	10,620	53.089	52.381	22.141	-1.768***
Score in Italian	12,816	71.303	73.750	15.876	10,620	61.219	63.934	17.691	10.084***
Aspirations: Degree or higher	12,894	0.411	0.000	0.492	10,620	0.572	1.000	0.495	-0.161***
Aspirations: Own edu > parents' edu	12,894	0.231	0.000	0.422	10,620	0.339	0.000	0.473	-0.108***
Students' characteristics									
Female	12,894	0.511	1.000	0.500	10,620	0.535	1.000	0.499	-0.024***
1st or 2nd gen immigrant	12,894	0.119	0.000	0.323	10,620	0.152	0.000	0.359	-0.033***
Books in the house: 0-10	12,894	0.077	0.000	0.266	10,620	0.089	0.000	0.284	-0.012**
Books in the house: 11-25	12,894	0.196	0.000	0.397	10,620	0.144	0.000	0.351	0.052***
Books in the house: 26-100	12,894	0.315	0.000	0.465	10,620	0.296	0.000	0.457	0.019**
Books in the house: 101-200	12,894	0.196	0.000	0.397	10,620	0.253	0.000	0.434	-0.056***
Books in the house: 201-500	12,894	0.216	0.000	0.411	10,620	0.219	0.000	0.413	-0.003
Mother's edu: below secondary	12,894	0.301	0.000	0.459	10,620	0.265	0.000	0.441	0.036***
Mother's edu: secondary	12,894	0.475	0.000	0.499	10,620	0.490	0.000	0.500	-0.016*
Mother's edu: tertiary	12,894	0.225	0.000	0.417	10,620	0.245	0.000	0.430	-0.021***
Father's edu: below secondary	12,894	0.317	0.000	0.465	10,620	0.293	0.000	0.455	0.024***
Father's edu: secondary	12,894	0.462	0.000	0.499	10,620	0.471	0.000	0.499	-0.009
Father's edu: tertiary	12,894	0.221	0.000	0.415	10,620	0.236	0.000	0.425	-0.015**
Mother's job: at home	12,894	0.363	0.000	0.481	10,620	0.299	0.000	0.458	0.064***
Mother's job: self-employed	12,894	0.108	0.000	0.310	10,620	0.125	0.000	0.330	-0.017***
Mother's job: high-skilled	12,894	0.132	0.000	0.338	10,620	0.136	0.000	0.343	-0.005
Mother's job: medium-skilled	12,894	0.223	0.000	0.416	10,620	0.228	0.000	0.419	-0.005
Mother's job: low-skilled	12,894	0.174	0.000	0.379	10,620	0.212	0.000	0.409	-0.037***
Father's job: at home	12,894	0.041	0.000	0.198	10,620	0.053	0.000	0.224	-0.012***
Father's job: self-employed	12,894	0.303	0.000	0.459	10,620	0.290	0.000	0.454	0.013*
Father's job: high-skilled	12,894	0.259	0.000	0.438	10,620	0.216	0.000	0.412	0.043***
Father's job: medium-skilled	12,894	0.152	0.000	0.359	10,620	0.156	0.000	0.363	-0.004
Father's job: low-skilled	12,894	0.245	0.000	0.430	10,620	0.285	0.000	0.452	-0.040***
Student is ahead	12,894	0.029	0.000	0.168	10,620	0.005	0.000	0.067	0.024***
Student is regular	12,894	0.792	1.000	0.406	10,620	0.845	1.000	0.362	-0.053***
Class size	12,894	21.843	22.000	4.132	10,620	20.211	21.000	4.844	1.632***
SES	12,727	0.209	0.192	0.963	10,591	0.121	0.173	0.965	0.088***
Region: Centre	12,894	0.167	0.000	0.373	10,620	0.190	0.000	0.392	-0.023***
Region: North-East	12,894	0.213	0.000	0.410	10,620	0.312	0.000	0.463	-0.099***
Region: North-West	12,894	0.301	0.000	0.459	10,620	0.410	0.000	0.492	-0.109***
Region: South and Islands	12,894	0.319	0.000	0.466	10,620	0.088	0.000	0.284	0.230***
Managerial quality									
Management Index	172	2.049	2.023	0.459	137	2.152	2.182	0.506	-0.103
ΔMI					135	0.043	0.000	0.439	
MI: Leadership	172	2.236	2.333	0.532	137	2.365	2.333	0.639	-0.129
MI: Targets	172	1.870	1.800	0.513	137	1.949	2.000	0.544	-0.079
MI: Operations	172	2.350	2.250	0.677	137	2.522	2.500	0.718	-0.172*
MI: Monitoring	172	2.236	2.200	0.630	137	2.416	2.400	0.625	-0.180*
MI: People	172	1.690	1.600	0.367	137	1.669	1.600	0.412	0.021
SP characteristics									
Age	172	57.837	59.000	4.855	137	57.526	58.000	5.467	0.312
Tenure as SP within school	172	4.831	4.000	4.505	137	4.569	3.000	3.523	0.262
Overall tenure (as SP or teacher) within school	172	6.785	4.000	7.222	137	7.022	4.000	7.334	-0.237
Female	172	0.355	0.000	0.480	137	0.431	0.000	0.497	-0.076
Job experience outside school	172	0.419	0.000	0.495	137	0.467	0.000	0.501	-0.049
School was chosen	172	0.843	1.000	0.365	137	0.825	1.000	0.382	0.018
SP manages multiple schools	172	0.035	0.000	0.184	137	0.088	0.000	0.284	-0.053*
School characteristics									
Number of students	172	793.512	773.000	273.653	137	876.350	872.000	288.807	-82.839*
Number of teachers	172	81.657	78.000	24.044	137	93.109	85.000	34.089	-11.453***
Number of admin staff	172	26.052	24.000	8.869	137	26.679	26.000	9.216	-0.627
Number of female teachers	172	53.820	51.500	19.429	137	62.190	58.000	23.486	-8.370***
Number of permanent teachers	172	65.401	64.000	20.314	137	72.847	69.000	25.864	-7.446**

Note: Descriptive statistics referred to the full sample used in the main analysis. Δ is the average change between 2011 and 2015, where differences and corresponding p-values are reported. * p<.10 ** p<.05 *** p<.01. Sources: INVALSI (outcome variables and student characteristics), WMS (managerial quality and SP characteristics), MIUR (school chatacteristics).

Table B2: Prob that SP changes by 2015, by 2011 school-specific attribute (estimated one by one)

Dep. variable	(1) SP has changed
SES	0.103
	(0.098)
Share of foreign students	-0.316
	(0.635)
Share of students ahead	-0.646
	(0.909)
Share of regular students	0.297
	(0.331)
Average class size	-0.030*
	(0.017)
Number of students	0.000
	(0.000)
Number of teachers	0.000
	(0.002)
Number of admin staff	0.010***
	(0.004)
Number of female teachers	0.000
	(0.002)
Number of permanent teachers	0.002
	(0.002)
SP's age	0.030***
	(0.009)
SP's tenure as SP within school	0.013*
	(0.007)
SP's overall tenure within school (as SP or teacher)	-0.003
	(0.006)
SP's gender	0.083
	(0.093)
SP has had experience outside	-0.037
	(0.093)
School was chosen	0.017
	(0.125)
SP manages multiple schools	0.425***
	(0.047)
Observations	118

Note: * p<.10 ** p<.05 *** p<.01. All coefficients are estimated in separate regressions. Dependent variable is a dummy taking value one if the SP has changed in between 2011 and 2015. School-level characteristics are measured in 2011.

Table B3: Effect on student scores, robustness checks

Panel A	(1)	(2)	(3)	(4)	(5)
Dep. variable	Score: Maths				
Management Index	3.741**	4.334**	5.808***	3.716**	3.997**
	(1.587)	(1.696)	(1.955)	(1.722)	(1.881)
Observations	23,318	23,514	23,514	22,254	16,095
R-squared	0.372	0.384	0.392	0.385	0.392
School-year clusters	308	309	309	291	210
Sample	Whole	Whole	Whole	Single school	Younger SP
School FE	✓	✓	✓	✓	✓
Area × Year FE	✓	✓		✓	✓
Province × Year FE			✓		
SP characteristics	✓	✓	✓	✓	✓
School characteristics	✓	✓	✓	✓	✓
Student characteristics	✓	✓	✓	✓	✓
Controls	SES	Non-linear			
Panel B	(1)	(2)	(3)	(4)	(5)
Dep. variable	Score: Language				
Management Index	2.566**	2.673**	3.115**	3.147**	1.366
	(1.180)	(1.292)	(1.540)	(1.247)	(1.776)
Observations	23,318	23,436	23,436	22,176	16,036
R-squared	0.437	0.451	0.460	0.447	0.469
School-year clusters	308	309	309	291	210
Sample	Whole	Whole	Whole	Single school	Younger SP
School FE	✓	✓	✓	✓	✓
Area × Year FE	✓	✓		✓	✓
Province × Year FE			✓		
SP characteristics	✓	✓	✓	✓	✓
School characteristics	✓	✓	✓	✓	✓
Student characteristics	✓	✓	✓	✓	✓
Controls	SES	Non-linear			

Note: * p<.10 ** p<.05 *** p<.01. Robust standard errors clustered at school level. All regressions include area-year fixed effects, where areas are North-West, North-East, Centre, South. The set of additional controls is identical to that included in Table 1. In Column 1 the SES index replaces information on the number of books in the house and the parents' education and job status. Column 2 also controls for the square of tenure within the school. Column 3 includes province × year fixed effects. Column 4 excludes cases where the SP manages multiple schools. Column 5 only considers SPs up to 60 years of age (below the 75th percentile).

Table B4: Effect on student aspirations, robustness checks

Panel A	(1)	(2)	(3)	(4)	(5)
Dep. variable	Aspirations: Degree or higher				
Management Index	0.041**	0.041**	0.068***	0.044**	0.054***
	(0.018)	(0.019)	(0.024)	(0.018)	(0.020)
Observations	23,318	23,514	23,514	22,254	16,095
R-squared	0.309	0.317	0.321	0.322	0.321
School-year clusters	308	309	309	291	210
Sample	Whole	Whole	Whole	Single school	Younger SP
School FE	✓	✓	✓	✓	✓
Area × Year FE	✓	✓		✓	✓
Province × Year FE			✓		
SP characteristics	✓	✓	✓	✓	✓
School characteristics	✓	✓	✓	✓	✓
Student characteristics	✓	✓	✓	✓	✓
Controls	SES	Non-linear			
Panel B	(1)	(2)	(3)	(4)	(5)
Dep. variable	Aspirations: Own edu > parents' edu				
Management Index	0.049***	0.036**	0.051**	0.044***	0.046***
	(0.018)	(0.015)	(0.021)	(0.015)	(0.011)
Observations	23,318	23,514	23,514	22,254	16,095
R-squared	0.160	0.327	0.329	0.333	0.321
School-year clusters	308	309	309	291	210
Sample	Whole	Whole	Whole	Single school	Younger SP
School FE	✓	✓	✓	✓	✓
Area × Year FE	✓	✓		✓	✓
Province × Year FE			✓		
SP characteristics	✓	✓	✓	✓	✓
School characteristics	✓	✓	✓	✓	✓
Student characteristics	✓	✓	✓	✓	✓
Controls	SES	Non-linear			

Note: * p<.10 ** p<.05 *** p<.01. Robust standard errors clustered at school level. All regressions include area-year fixed effects, where areas are North-West, North-East, Centre, South. The set of additional controls is identical to that included in Table 1. In Column 1 the SES index replaces information on the number of books in the house and the parents' education and job status. Column 2 also controls for the square of tenure within the school. Column 3 includes province × year fixed effects. Column 4 excludes cases where the SP manages multiple schools. Column 5 only considers SPs up to 60 years of age (below the 75th percentile).

Table B5: Effect on student scores and aspirations, balanced samples of schools

Panel A	(1)	(2)	(3)	(4)	(5)	(6)
Dep. variable	Score: Maths			Score: Italian		
Management Index	3.829**	4.524**	3.575*	2.484**	2.505*	2.916
	(1.663)	(1.740)	(2.107)	(1.197)	(1.263)	(1.863)
Observations	18,988	11,180	8,404	18,915	11,109	8,351
School-year clusters	236	140	112	236	140	112
R-squared	0.374	0.386	0.410	0.459	0.469	0.476
School FE	✓	✓	✓	✓	✓	✓
Area-Year FE	✓	✓	✓	✓	✓	✓
SP characteristics	✓	✓	✓	✓	✓	✓
School characteristics	✓	✓	✓	✓	✓	✓
Student characteristics	✓	✓	✓	✓	✓	✓
Panel B	(1)	(2)	(3)	(4)	(5)	(6)
Dep. variable	Aspirations: Degree or higher			Aspirations: Own edu > parents' edu		
Management Index	0.036**	0.046**	0.034	0.035**	0.036**	0.033
	(0.018)	(0.019)	(0.030)	(0.015)	(0.018)	(0.025)
Observations	18,988	11,180	8,404	18,988	11,180	8,404
School-year clusters	236	140	112	236	140	112
R-squared	0.311	0.326	0.348	0.339	0.358	0.379
School FE	✓	✓	✓	✓	✓	✓
Area-Year FE	✓	✓	✓	✓	✓	✓
SP characteristics	✓	✓	✓	✓	✓	✓
School characteristics	✓	✓	✓	✓	✓	✓
Student characteristics	✓	✓	✓	✓	✓	✓

Note: * p<.10 ** p<.05 *** p<.01. Robust standard errors clustered at school level. All regressions include area-year fixed effects, where areas are North-West, North-East, Centre, South. The set of additional controls is identical to that included in Table 1.

C The WMS questionnaire

Table C1: Questions of the World Management Survey

Section: Definition	Questions
Leadership: "Leadership vision"	Q1. A) What is the school's vision for the next five years? Do teachers/ staff know and understand the vision? B) Who does your school consider to be your key stakeholders? How is this vision communicated to the overall school community? C) Who is involved in setting this vision/ strategy? When there is disagreement, how does the school leader build alignment?
Leadership: "Clearly defined accountability for leaders"	Q15. A) Who is accountable for delivering on school targets? B) How are individual school leaders held responsible for the delivery of targets? Does this apply to equity and cost targets as well as quality targets? C) What authority do you have to impact factors that would allow them to meet those targets (e.g., budgetary authority, hiring and firing)? Is this sufficient?
Leadership: "Clearly def. leadership and teacher roles"	Q16. A) How are the roles and responsibilities of the school leader defined? How are they linked to student outcomes/ performance? B) How are leadership responsibilities distributed across individuals and teams within the school? C) How are the roles and responsibilities of the teachers defined? How clearly are required teaching competences defined and communicated? D) How are these linked to student outcomes/ performance?
Operations: "Standardisation of instructional processes"	Q2. A) How structured or standardized are the instructional planning processes across the school? B) What tools and resources are provided to teachers (e.g., standards-based lesson plans and textbooks) to ensure consistent level of quality in delivery across classrooms? C) What are the expectations for the use of these resources and techniques? D) How does the school leader monitor and ensure consistency in quality across classrooms?

Continuation of Table C1

Operations: "Personalization of Instruction and Learning"	Q3. A) How much does the school attempt to identify individual student needs? B) How are these needs accommodated for within the classroom? How do you as a school leader ensure that teachers are effective in personalising instruction in each classroom across the school? C) What about students, how does the school ensure they are engaged in their own learning? How are parents incorporated in this process?
Operations: "Data-Driven Planning and Student Transitions"	Q4. A) Is data used to inform planning and strategies? If so how is it used – especially in regards to student transitions through grades/ levels? B) What drove the move towards more data-driven planning/ tracking?
Operations: "Adopting Educational Best Practices"	Q5. A) How does the school encourage incorporating new teaching practices into the classroom? B) How are these learning or new teaching practices shared across teachers? What about across grades or subjects? How does sharing happen across schools (community, state-wide etc), if at all? C) How does the school ensure that teachers are utilising these new practices in the classroom? How often does this happen?
Monitoring: "Continuous Improvement"	Q6. A) When problems (e.g., within school/ teaching tactics/ etc.) do occur, how do they typically get exposed and fixed? B) Can you talk me through the process for a recent problem that you faced? C) Who within the school gets involved in changing or improving process? How do the different staff groups get involved in this? D) Does the staff ever suggest process improvements?
Monitoring: "Performance Tracking"	Q7. A) What kind of main indicators do you use to track school performance? What sources of information are used to inform this tracking? B) How frequently are these measured? Who gets to see this performance data? C) If I were to walk through your school, how could I tell how it was doing against these main indicators?
Monitoring: "Performance Review"	Q8. A) How often do you review (school) performance – formally or informally– with teachers and staff? B) Could you walk me through the steps you go through in a process review? C) Who is involved in these meetings? Who gets to see the results of this review? D) What sort of follow-up plan would you leave these meetings with? Is there an individual performance plan?

Continuation of Table C1

Monitoring: "Performance Dialogue"	Q9. A) How are these review meetings structured? B) Do you generally feel that you do have enough data for a fact-based review? C) What type of feedback occurs during these meetings?
Monitoring: "Consequence Management"	Q10. A) Let's say you've agreed to a follow-up plan at one of your meetings, what would happen if the plan was not enacted? B) How long does it typically go between when a problem is identified to when it is solved? Can you give me a recent example? C) How do you deal with repeated failures in a specific department or area of process?
Targets: "Target Balance"	Q11. A) What types of targets are set for the school to improve student outcomes? Which staff levels are held accountable to achieve these stated goals? B) How much are these targets determined by external factors? Can you tell me about goals that are not externally set for the school (e.g., by the government or regulators)?
Targets: "Target Inter-Connection"	Q12. A) How are these goals cascaded down to the different staff groups or to individual staff members? B) How are your targets linked to the overall school-system performance and its goals?
Targets: "Time Horizon of Targets"	Q13. A) What kind of time scale are you looking at with your targets? B) Which goals receive the most emphasis? C) Are the long-term and short-term goals set independently? D) Could you meet all your short-run goals but miss your long-run goals?
Targets: "Target Stretch"	Q14. A) How tough are your targets? How pushed are you by the targets? B) On average, how often would you say that you and your school meet its targets? How are your targets benchmarked? C) Do you feel that on targets all departments/ areas receive the same degree of difficulty? Do some departments/ areas get easier targets?
Targets: "Clarity and Comparability of Targets"	Q.17 A) If I asked one of your staff members directly about individual targets, what would they tell me? B) Does anyone complain that the targets are too complex? Could every staff member employed by the school tell me what they are responsible for and how it will be assessed? C) How do people know about their own performance compared to other people's performance?

Continuation of Table C1

People: "Rewarding High Performers"	Q18. A) How does your evaluation system work? What proportion of your employee pay is related to the results of this review? B) Are there any non-financial or financial bonuses/ rewards for the best performers across all staff groups? How does the bonus system work (for staff and teachers)? C) How does your reward system compare to that of other schools?
People: "Removing Poor Performers"	Q19. A) If you had a teacher who was struggling or who could not do his/ her job, what would you do? Can you give me a recent example? B) How long is underperformance tolerated? How difficult is it to terminate a teacher? C) Do you find staff members/ teachers who lead a sort of charmed life? Do some individuals always just manage to avoid being fired?
People: "Promoting High Performers"	Q20. A) Can you tell me about your career progression/ promotion system? B) How do you identify and develop your star performers? C) What types of professional development opportunities are provided? How are these opportunities personalised to meet individual teacher needs? D) How do you make decisions about promotion/ progression and additional opportunities within the school, such as performance, tenure, other? Are better performers likely to be promoted faster, or are promotions given on the basis of tenure/ seniority?
People: "Managing Talent"	Q21. A) How do school leaders show that attracting talented individuals and developing their skills is a top priority? B) How do you ensure you have enough teachers of the right type in the school? C) Where do you seek out and source teachers? D) What hiring criteria do you use?
People: "Retaining Talent"	Q22. A) If you had a top performing teacher who wanted to leave, what would the school do? B) Could you give me an example of a star performer being persuaded to stay after wanting to leave? C) Could you give me an example of a star performer who left the school without anyone trying to keep him?
People: "Creating a Distinctive Employee Value Proposition"	Q23. A) What makes it distinctive to teach at your school, as opposed to other similar schools? If you were to ask the last three candidates would they agree? Why? B) How do you monitor how effectively you communicate your value proposition and the following recruitment process?