Managerial Practices and Student Performance*

Adriana Di Liberto
University of Cagliari,
IZA and CRENoS

Fabiano Schivardi
Bocconi, IGIER, EIEF
and CEPR

Giovanni Sulis
University of Cagliari
and CRENoS

December 15, 2014

Abstract

We study the effects of managerial practices in schools on student outcomes. We measure managerial practices using the World Management Survey, a methodology that enables us to construct robust measures of management quality comparable across countries. We find substantial heterogeneity in managerial practices across six industrialized countries, with more centralized systems (Italy and Germany) lagging behind the more autonomous ones (Canada, Sweden, the UK, the US). For Italy, we are able to match organizational practices at the school level with student outcomes in a math standardized test. We find that managerial practices are positively related to student outcomes. The estimates imply that if Italy had the same managerial practices as the UK (the best performer), it would close the gap in the math OECD-PISA test with respect to the OECD average. We argue that our results are robust to selection issues and show that they are confirmed by a set of IV estimates and by a large number of robustness checks. Overall, our results suggest that policies directed at improving student cognitive achievements should take into account principals selection and training in terms of managerial capabilities.

Keywords: management, productivity, school principals, cognitive skills.
J.E.L. Classification: L2, I2, M1, O32.

*Contacts: Adriana Di Liberto, diliberto@unica.it; Fabiano Schivardi, fabiano.schivardi@unibocconi.it; Giovanni Sulis, gsulis@unica.it. We thank Daniele Checchi, Gianfranco De Simone, Andrea Gavosto, Erik Hanushek, Victor Lavy, Angelo Paletta, Paolo Sestito, Marco Sideri, John Van Reenen and seminar participants at the University of Cagliari, University of Padova, 2013 AIEL Conference (Rome), 2013 Brucchi Luchino Conference (Rome), 2014 RES Conference (Manchester), 2014 ESPE Conference (Braga), *10 Years of the World Management Survey* Conference (Cambridge, MA) for very useful comments and discussions. We also thank the editors, the referees, the participants and the discussants, Margherita Fort and Sasha Becker, at the 60th Panel Meeting in Rome for their constructive comments. Sulis acknowledges hospitality at the Center for Labor Economics at Berkeley University. We thank Renata Lemos and Raffaella Sadun for their support in the World Management Survey data collection process and Barbara Dettori for help with the PISA data. We also thank INVALSI and in particular Patrizia Falzetti for providing the data on student outcomes. We are solely responsible for all the remaining errors. Financial support from the Regione Sardegna, Legge 7/2007 - grant CUP F71J09000300002, and from the Fondazione Giovanni Agnelli is gratefully acknowledged.

https://doi.org/10.1093/epolic/eiv015

1 Introduction

The importance of human capital for economic growth is one of the most uncontroversial facts in economics (Lucas, 1988; Barro and Lee, 1994). And human capital is in a large part "produced" in school. It is therefore not surprising that the debate on the determinants of student performance, and on policies that can improve them, is very lively. However, despite a large amount of work, this debate is far from having reached robust conclusions. Hanushek and Woessmann (2011) review the available empirical studies and conclude that "evidence from both within and across countries points to the positive impact of competition among schools, of accountability and student testing, and of local school autonomy in decision making." However, these factors are conducive to better student achievement only in well-developed school systems (Hanushek, Link, and Woessmann, 2013). In fact, only when schools are well managed students can benefits from decentralization, while giving autonomy to badly run institutions can indeed worsen student outcomes.

Indeed, there is a growing attention regarding the role of school principals (SPs in what follows) as managers in charge of running the school (Bloom et al., 2014). However, we still know relatively little on this issue, because assessing the role of managerial practices on student outcomes is a difficult task, mostly due to the challenges of measuring such practices. Thus, while there is a large *qualitative* literature stressing the importance of the role of SPs and leadership on school outcomes, only few recent studies have attempted to *quantify* the role played by the SPs on student outcomes. This paper addresses this question.

We collect data on school managerial practices through extensive phone interviews of around 400 SPs. The sample is representative of the population of Italian upper secondary schools. The interviews are based on the World Management Survey (Bloom and Van Reenen, 2007, 2010b) data collection method, that allows to score the managerial practices adopted in a given institution. The survey covers 23 specific managerial activities that can be combined to obtain a synthetic measure of management quality and also grouped into five specific management areas: operations, monitoring, targets, incentives and leadership. The double blind and open questions techniques implemented in the World Management Survey enables to obtain high quality data that control for typical problems of self-assessment bias. It has been applied to a large number of both private firms and public institutions in health and education (Bloom et al., 2012, 2014). The data collection method is standardized and allows for meaningful comparisons across countries. This type of data is increasingly used in academic research (Bloom and Van Reenen, 2010a). The work

of Bloom et al. (2014) is particularly relevant for our paper. They study the relationship between school managerial practices and student performance in a cross-country setting, finding that management quality is positively associated with pupil outcomes. We see the two studies as complementary. On one side, the cross-country dimension shows that the correlation we find extends behind the case of Italy and is of general nature. At the same time, the cross-country nature of the data does not allow them to determine whether this correlation is causal. Our within country exercise, on the contrary, enables us to exploit plausible identification strategies to investigate the causal effect of management on student performance.

We first compare the managerial practices for six countries for which data are available (Canada, Germany, Italy, Sweden, the UK and the US). We find substantial heterogeneity in managerial practices across countries, with more centralized systems (Italy and Germany) lagging behind the more autonomous ones, as also found by Bloom et al. (2014). We also show that these cross country differences are not simply due to different institutional constraints that each national schooling system imposes on SPs. For example, we find that the gap of Italian schools practices is not lower in areas where the institutional constraints are less of an issue, such as in planning and monitoring the school objectives, compared to areas in which they are much more binding in some countries than in others, such as in terms of hiring and firing teachers. This suggests that the observed cross country heterogeneity in the quality of managerial practices is at least partially due to difference in underlying SP abilities. This in turn can be attributed to the different selection and/or the training mechanisms in place in different countries.

For the Italian data we are able to match the indicators of managerial practices with tenth grade student results in a standardized math test administrated by the INVALSI, the Italian institute in charge of evaluating school performance. The Italian case is an interesting one to study the effects of SPs managerial practices on student outcomes. First, there is substantial geographical heterogeneity in both quantitative (educational attainments) and qualitative (cognitive skill tests results) educational outcomes. Second, as we argue in detail in the paper, the process of assignment of SPs to schools greatly reduces endogeneity concerns, according to which the most capable SPs are assigned to the best schools. In fact, we have access to a rich set of covariates at the school, SP, and individual student level, that should control for the most likely selection issues. Moreover, SPs are assigned through an informal process based mostly on seniority, a characteristic we can control for. Finally, we use the reforms of the Italian school system to construct an IV regression.

Our baseline model is an OLS regression of student performance on the indicator of overall managerial practices (obtained as the mean of all the areas surveyed in the interview), controlling for a large number of school, SPs and student characteristics. Data on student performance are the test scores for Maths expressed as percentage of right answers. We find that the indicator of managerial practices has a positive coefficient of 2.24, significant at 10%. Given that the test results are between 0 and 100, with a sample average of 49.04, the estimated coefficient can be readily interpreted in terms of increased test score results. It implies that a unit increase in the indicator of managerial practices (which has mean 2.01 and standard deviation of .5) would improve the students average test score by 4.6%, approximately a 12 positions increase for Italy, which ranked 32nd in the 2012 OECD PISA country rankings. While a unit increase in managerial practices is clearly substantial, it is also approximately the distance from the average value in Italy and in the UK, the country with the highest score.

In terms of specific areas of managerial practices, we find that the effects are positive in all categories, although statistically significant only for leadership and monitoring activities of school processes. In particular, the fact that we find no significant effect on the incentives section (People), that includes human resource management, is consistent with the high degree of institutional constraints that Italian SPs face on this subject.

We perform several robustness checks. First, we use the fact that, starting in 2006, a new national competition was introduced that also explicitly assesses managerial skills. This reform should have an impact on managerial skills of those that became SPs afterwards, while being unrelated to the assignment to specific schools (conditional on controls). We use a dummy for those who became SP after 2006. The IV results confirm the OLS ones, with the effect becoming substantially larger. Second, since student sorting is more likely to arise in more densely populated areas where schools are in competition, we run a separate analysis for isolated vs non isolated schools. The results are at odds with the sorting explanation. We also replicate our analysis on different sub-samples, namely, of low and high socioeconomic background students, finding that managerial practices impact students performance more in the first category of schools. We consider how the effect of managerial practices varies depending on the importance of institutional constraints, finding that this is indeed the case, and check if cheating behaviour may be driving our main results, ruling out this possibility. We also show that managerial practices affect the distribution of test scores within the school uniformly, shifting the whole distribution to the right. Finally we consider an alternative student outcome variable and check if good practices play a role on the

probability of students lagging behind in the age/grade ladder. All these exercises confirm that management quality is an important input of our estimated education production function.

Our last exercise is to compare the management indicators of the WMS with those that can be obtained from the OECD PISA survey, which contains a section where SPs self assess the quality of management and the degree of autonomy of the school. This is a less demanding way to assess practices and it is available for a large set of countries. It is therefore useful to check if the two surveys supply a similar picture. Unfortunately, this does not seem to be the case. Both the direct comparison between PISA and WMS indicators and the regression analysis based on PISA management variables reveal marked differences between the two surveys. These disparities are likely to be due to both methodological differences in data collection and possible mis-measurement of the self-reported PISA managerial indices. They suggest that self assessment cannot substitute the direct assessment of SPs managerial capabilities by a third party.

The rest of the paper is organized as follows. In Section 2 we review the related literature. Section 3 describes the World Management Survey and discusses the channels of action of SPs. Section 4 compares the survey results for the six countries. We describe the student data and the additional controls in Section 5 and the identification and the empirical design in the following Section 6. Section 7 discusses the main results and Section 8 the extensions and robustness checks. We also compare our management data with the existing measures of school principals leadership provided by the OECD PISA project in Section 9 and finally draw the main policy implications and conclude in Section 10.

2 Literature review

The role of the SPs on student learning is increasingly identified as crucial. Most existing analysis on the role of the SPs on student performance has been qualitative and only few recent studies have attempted to quantify the role played by the SPs in the results obtained by the students during their school career.[1] This is due, at least in part, by measurement problems since the identification of SPs efficiency is a difficult empirical issue and results may change significantly depending on the methodology adopted.[2] Most quantitative studies use the value added approach, a methodology already introduced to estimate the effect

[1] One of the first studies is Brewer (1993).

[2] Loeb, Kalogrides, and Béteille (2012) compare three measures of SP quality calculated with standard value added methodologies with alternative survey measures of SP performance and find low correlation across the different indicators.

of individual teachers on student performance but employed also outside the education framework to identify the role of CEOs in firms productivity.[3]

Using data collected between 1995 and 2001 for a sample of Texas schools, Branch, Hanushek, and Rivkin (2012) estimate the role of principals on student academic achievement using the semi-parametric approach provided by Bertrand and Schoar (2003) in their study of corporate management styles. They find significant variation in principal quality and identify a large effect, similar to that found for teachers, of principal leadership on student outcomes.[4] Moreover, SPs quality variance appears to be larger for more disadvantaged schools suggesting that the leadership skills have larger effects in these schools. Significant effects of SPs on student test results have also been found in Coelli and Green (2012) for Canada and Böhlmark, Grönqvist, and Vlachos (2012) for Sweden. Together with student test outcomes, the latter study also finds that SPs quality significantly affects alternative school outcomes variables and find that SPs in smaller schools have a larger effect on student test results.[5]

Li (2012) focuses on the labor market dynamics of the principals and shows as test-based accountability systems may significantly change SPs incentives and, through that, their allocation decisions, with unintended consequences on disadvantaged students. In particular, she uses data from the No Child Left Behind (NCLB) policy as implemented in North Carolina that introduced formal sanctions for schools and principals missing specific student performance targets. This analysis suggests that the relative change in the risk-reward structure of low versus high-performing schools introduced by the new test-based accountability system decreased the average quality of principals serving disadvantaged schools. In fact, principals' pay does not fully adjust to compensate the risks, inducing more able SPs, who are more likely to have the option of working elsewhere, to depart these schools.[6] Thus, even if one goal of the NCLB policy was to increase the competencies of most disadvantaged students, this study shows that this induced allocation effect may

[3]See Chetty, Friedman, and Rockoff (2011), Rivkin, Hanushek, and Kain (2005) and Hanushek and Rivkin (2012) for teachers and Bertrand and Schoar (2003) for CEOs efficiency estimates.

[4]Their lower bound results imply that a principal in the top 16 percent of the quality distribution (or one standard deviation above average) would lead each year to student gains that are 0.05 s.d. or more higher than average for all students in the school. For results on the impact of teachers see Rockoff (2004) and Rivkin, Hanushek, and Kain (2005) among the others.

[5]Specifically, these alternative outcome variables are grade inflation, wage dispersion, the presence of a gender balanced teaching staff, teachers retention rates and teachers on long-term sick leave.

[6]Unlike other states, North Carolina already had an accountability program in place before the introduction of NCLB but performance targets and sanctions were less binding for principals/schools. Using a different sample, Branch, Hanushek, and Rivkin (2012) do not find strong evidence of more effective leaders having higher probability of exiting more disadvantaged schools.

produce exactly the opposite result. The importance of SPs incentives is also investigated by Lavy (2008) who finds that the increase in the salary of high schools principals in Israel led to significant improvements in student academic achievements (Cullen and Mazzeo, 2008; Brewer, 1993).

Unlike the first, the second strand of literature goes more in depth about the specific SPs managerial practices and activities using both teachers/parents survey responses based upon personal perceptions of the principal or SPs self-assessment surveys. In fact, recent studies criticize the excessive attention paid by education scholars to the role played by the SP in supporting teaching activities and conversely stress the importance of more managerial activities. Grissom and Loeb (2011) exploit the answers given to a questionnaire submitted to 314 SPs in the district of Miami who were asked to provide a self-evaluation on a scale from 1 to 4 for the effectiveness in leading the school in 42 specific tasks and find that the more strictly managerial and organizational skills have the greatest impact on educational attainment. Their results are also compared with the answers given by the assistant principals or vice-principals to the same questionnaire as well as alternative indicators of teachers and parents satisfaction with school quality.[7]

In general, both the value added and the survey responses approaches to estimate the SPs effectiveness may be subject to criticisms. In particular, the former approach exploits SPs turnover across schools and requires large longitudinal data sets to observe a sufficient number of principals switches to convincingly identify their quality. This is done to reduce concerns about conflating principal and other school effects that would be present including stayers, that is, principals who are only observed in one school. However, even when long panel data are available, self selection problems may still arise since SPs are not not likely to be randomly assigned to schools: if SPs systematically move in best performing schools (in terms of student test achievement gains) value-added measures of principal efficiency are still biased.[8] Second, value added measures do not control for the possibility that SPs quality change over time with tenure and experience. Third, they produce an overall measure of the SPs impact but they do not tell much about what SPs actually do to influence student learning.[9] As seen above, the survey approach overcome the latter criticism. However, the

[7]Robinson, Lloyd, and Rowe (2008) conduct two separate meta-analyses and find significant effect of both *instructional leadership* activities and of more specific management activities on student academic performance.

[8]In this case, such problems conflate the true SP effect with other factors for test results change, but other authors dispute this view. See for example Li (2012) who argues that compared to teachers value-added measures "...these concerns are less of a problem in the context of studying principals." Li (2012), p.17.

[9]This approach enables Branch, Hanushek, and Rivkin (2012) to investigate the relationship between the observed patterns of teacher exits and principals quality.

use of these type of indices raises the concerns over mis-measurement since they suffer from being based either on (teachers/parents) perceptions or suffer from the typical problems of self-assessment bias.[10]

3 The role of SPs managerial practices in schools

In this section we first discuss the channels through which managerial practices can affect student outcomes and then we describe how we measures them.

3.1 How do managerial practices affect student outcomes?

Education scholars recognize an important but often indirect role of the SPs through their influence on teachers. Schools are complex organizations, and SPs are seen as the leaders who set the conditions through which teachers make a more direct impact on student performance. That is, their contribution to student learning is done by shaping the conditions and climate in which teaching and learning occur. A notable advantage of the WMS tool is that it enables to distinguish the specific channels identified by the literature through which SPs affect student outcomes. We next describe such channels, how they relate to the measures of practices obtained through the WMS and discuss how they are likely to play out in the Italian institutional environment.

The quality of teachers is considered as one of the primary channels (Hanushek and Rivkin, 2006, 2012). The degree of autonomy in the recruitment of new staff and dismissal of the existing one varies by country, since collective bargaining agreements may prevent principals from engaging in firing/hiring low/high performing teachers.[11] The six questions on the WMS survey section on incentives and human resource management (see the People section, Table A10 in the online Appendix) ask if rewards or punishments are awarded as a consequence of well-defined and monitored individual achievements and how the school actively controls the number and types of teachers, staff and leadership needed to meet goals. In terms of hirings and firings and wage determination, Italian SPs have very limited autonomy since both teachers allocation and salaries are set at the central level. The WMS survey by design measures actual practices rather than the intrinsic ability of the SP and this implies that institutional constraints will impact on the quality of the managerial practices

[10] "...on average principals rated themselves highly on most tasks, a pattern consistent with other principal self-assessment tools." Grissom and Loeb (2011), p. 1100.

[11] The schooling system is largely dominated by the public sector and collective bargaining agreements are usually binding in most countries. See Loeb, Kalogrides, and Béteille (2012) and Bloom et al. (2012, 2014). See also Jacob (2011) for an example of the effect that a new collective bargaining agreement giving SPs more flexibility in the recruitment of new staff has on school teachers composition.

adopted and on the score obtained by the SPs. Thus, we expect all Italian SPs to obtain low scores in this area of management, while differences across Italian SPs would reflect their ability to apply more informal rather than formal incentives to both select and incentivate teachers and staff.

A second channel focuses on the role of SPs in promoting the introduction of organizational innovations that enable teachers to work more effectively. Indeed, Italian SPs have certainly some discretion in designing the organizational structure of the school they lead, and the impact of institutional constraints on the managerial practices adopted is lower that for human resource management. The set of questions included in both the Targets and Monitoring areas of the WMS focus on these issues. In details, the Targets section of the survey examines mainly the type of targets set by the school in terms of student outcomes. As also stressed by Branch, Hanushek, and Rivkin (2012), public sector CEOs in general and, thus, also SPs, do not necessarily have a well-defined objective function. This is certainly the case for Italian SPs, since they do not have any direct incentive to maximize schools test results and, even at aggregate school level, test results are not made public. The five questions on Monitoring included in our survey focus on the tracking of school performance, reviewing performance with teachers and staff and acting accordingly (e.g., making sure that, if a problem is identified, the appropriate actions to solve it are adopted). It also includes questions on whether school performance data are regularly tracked, reviewed with appropriate frequency, and communicated to the staff (see Table A10 for details).

Organizational innovations also include specific activities that facilitate and improve the quality of teaching and learning. These are also called by education scholars *instructional leadership* activities (Robinson, Lloyd, and Rowe, 2008; Grissom and Loeb, 2011) and the WMS section on Operations is the one that best identifies these practices. It includes four questions that focus on how the SP deals with different aspects of instructional planning process designed in a school. The main ones are the alignment of instructional strategies across teachers, the capacity to meet specific student needs and how the school provides information and connects students and parents with adequate resources to support student learning. In order to obtain a high score in this area the SP has to specify that when implementing these managerial practices she/he makes use of data and of comprehensive monitoring. For example, if we focus on the fourth question (see "Adopting Educational Best Practices" in Table A10), the maximum score is obtained by SPs that provide the school staff with specific opportunities to collaborate and share best practice techniques and also supports their monitored implementation in the classroom, while the minimum is

obtained when SP answers reveal only minimal understanding or monitoring of improved practices and learning techniques.

Fourth, motivation of the teaching staff, or *transformational leadership*, is also considered an important aspect of the SP work and consists of "...the ability of some leaders...to engage with staff in ways that inspired them to new levels of energy, commitment, and moral purpose." (Robinson, Lloyd, and Rowe, 2008, p. 639). The WMS questions on Leadership capture some of these aspects. In general, they are aimed to capture something that is difficult to measure, that is, all activities that SPs perform to informally stimulate/incentivize teacher work. They also ask whether the SP have clearly identified roles and responsibilities within the school and if there is any internal formal accountability system in place.

Finally, there are other possible paths through which SPs may affect student outcomes that are not directly captured by the WMS questions (Hallinger and Heck, 1998). In particular, SPs may play an important role in determining student discipline and/or in allocating teachers and students more or less effectively across classes, two activities that are identified by the literature as affecting student performance. For example, in Lazear (2001) more discipline is usually associated with less disruptions and better student results, and it represents an important factor to explain the better outcomes obtained by Catholic schools students. Second, consider a setting in which the SPs objective function is to maximize the overall school test score results and the way students are allocated to teachers is a choice variable of the principals, who may assign teachers to students in a way that maximizes average student performance in the school, i.e. better teachers to a more 'difficult' pool of students. In this case, we should still observe effects of managerial practices on student performance. Indeed, even if these factors are not specifically measured by the WMS questions, they are nevertheless indirectly captured by them. Both discipline and an efficient allocation of teachers across classes can only be implemented in a well-organized and monitored environment, since it involves guidelines and rules and the monitoring of student behavior.

3.2 The World Management Survey

As highlighted in section 2, the main difficulty encountered when analyzing the effect of SPs on school outcomes is to provide a reliable quantitative measure of SP abilities in terms of leadership capacity and organizational skills. While there is an established literature that suggests that such components are important determinants of firms productivity differentials across countries and sectors (Bloom and Van Reenen, 2007, 2010b, 2011), good data on managerial practices in the public sector are hardly available (Bloom et al., 2012, 2014).

Obtaining a robust measure of managerial practices that doesn't suffer from such problems is not a simple task. Bloom and Van Reenen (2007) discuss in detail the main challenges. First, measuring management requires a definition of "good" and "bad" managerial practices which is possibly not contingent on the specific production environment (firms, hospitals, schools) and applicable to different units. Second, the responses of the managers to survey questions should be unbiased and there should be no preconceptions of interviewers about the performance of the production unit analyzed. Finally, when collecting data on managers operating in the public sector, additional problems related to the institutional constraints limiting their activity should be, to some extent, taken into account.

In this paper we use the survey tool proposed within the international project World Management Survey (WMS henceforth) to obtain quantitative measures of managerial practices adopted by SPs operating in the Italian upper secondary school system. Such innovative tool, initially developed by Bloom and Van Reenen (2007) for the manufacturing sector and subsequently adapted for the service and public sector, is based on a telephone double-blind survey technique and comprises a set of open ended questions that are subsequently evaluated using a scoring grid (Bloom and Van Reenen, 2011, 2010a). Qualitative answers of SPs are then recoded into quantitative measures with a score ranging between 1 (worst) to 5 (best managerial practices). The aim of the questionnaire is not that of measuring the intrinsic abilities of the SP, or practices that are too contingent to the specific environment, but the quality of managerial practices adopted. In fact, such approach identifies managerial practices that are common across units, such as schools or firms, and focuses on the solutions adopted by principals/managers to solve specific problems.[12]

The questionnaire, that is reported in Table A10 in the online Appendix, comprises five sections that consider different key areas of management practices.[13] The first section is Leadership (three questions) and measures the leadership capacity of the SP jointly with a clear definition of roles and responsibilities within the school. The second section is Operations (four questions) and is concerned with the standardization of instructional processes, personalization of teaching and adoption of best practices within the school. The third dimension is Monitoring (five questions) and focuses on the monitoring of performance and reviewing the results, the dialogue between components within the school and the consequences of anomalies in the processes. The fourth section is Targets (five questions) and has

[12] Although the survey is very similar across different sectors, there are differences in terms of specific sections and questions included. We will discuss this issue in more detail in next sections.

[13] The overall management index that we use in the empirical application is calculated as the average of scores obtained in each question. Moreover we calculate an average score for each section of the questionnaire. See sections 4 and 7 below for further details.

the objective to assess the managerial capacity of SPs to identify quantitative and qualitative targets, their interconnection and their temporal cascade. Finally, the fifth dimension is People (six questions) and it is specifically concerned with human resource management, ranging from promoting and rewarding employees based on performance, removing poor performers, hiring best teachers, and trying to keep the best ones. The remaining part of the questionnaire collects data on the main principal and school characteristics.

4 Cross country comparisons

In addition to Italy, the same survey on the education sector has been previously run in Canada, Germany, Sweden, the UK and the US (Bloom et al., 2012). Given the standardized data collection process, these indicators are fully comparable across countries.[14] Given that the econometric analysis of the effects of managerial practices on student outcomes will be performed for Italy, that is the only country for which we have student outcomes data, in what follows we benchmark the discussion on this country. Figure 1 shows that with 341 interviews conducted, the Italian sample is the largest in the group and in terms of interviews completed it shows the highest response rate.[15]

Cross country differences in the quality of managerial practices can stem from two main sources. First, as seen above managerial practices are clearly influenced by institutional constraints, especially by school legislation and regulations regarding the employment contracts in the public sector. There is indeed a large degree of cross-country heterogeneity in autonomy and accountability of SPs (Pont et al., 2008). For example, in terms of hirings and firings and wage determination, the countries in our sample can be divided into three groups: in the US and the UK SPs have a large degree of autonomy, in Sweden and Canada they have a good degree of autonomy but are subject to some restrictions, particularly on the firing side, and in Germany and Italy they have very limited autonomy. In particular, in Italy teachers are allocated at the central level and cannot be removed by SPs. This will impact on the quality of the managerial practices adopted by the SPs. An institutional framework that greatly constraints human resource management therefore generates low scores in such areas, independently from the intrinsic ability of the principals. A second

[14] A more extensive cross-country comparison that also includes data from two emerging economies, namely, India and Brazil, can be found in Bloom et al. (2014).

[15] The response rate ranges from 57% for Italy to 36%, 26%, 20%, 19% and 8% of eligible schools in Sweden, Germany, the US, Canada and UK. See Bloom et al. (2014). In the online Appendix C we explain in detail the selection of the sample, the response rates and we show that there are no significant differences in terms of observable characteristics between the principals/schools that granted the interview and those who refused to participate.

possibility is that managerial practices are on average (say) of lower quality in a certain country because SPs are intrinsically less capable, depending both on the selection process and on training. It is important to shed light on the relative weight of these possible sources of heterogeneity. The policy implications for improving managerial practices are in fact very different depending on the answer to this question. In the first case, it is to review the institutional framework within which SPs operate. In the second, one should question the selection and training process of SPs.

We begin the analysis with the overall indicator of the quality of managerial practices, Management, obtained as the average of all the questions and reported in Figure 1. With an average of 2.01, the managerial skills of Italian school leaders are significantly lower than those of other countries. The British SPs achieve the highest score, just below 3, followed by Sweden (2.79), Canada (2.80), the US (2.74) and Germany (2.56).

In addition to the average, it is useful to analyze its distribution to evaluate the heterogeneity of managerial skills of the SPs within each country. To explore this aspect, following Bloom et al. (2012, 2014), in Figure 2 we report the distribution of the variable for each country. In Italy it is highly concentrated on low values: a significant proportion of Italian SPs implement low-quality management practices. In contrast, in other countries the distribution indicates a significantly higher "minimum" level of management skills, with few SPs with values less than 2, especially in the case of Sweden and Great Britain. Table 1 reveals a similar trend in terms of percentiles. Italy displays the highest interquartile range: the ratio between the ninetieth and the tenth percentile is equal to 1.98 against an average value for all the other countries of approximately 1.5. This is consistent with the idea that in this country there is little control in establishing a minimum level of managerial ability to become a SP.

One way to determine the relative importance of institutional constraints vis-a-vis intrinsic differences in ability is to analyze the results of the survey for the individual subsections. The basic idea is that institutional constraints are likely to be differently binding for different areas of school management. For example, as argued above, there are substantial differences in the constraints in the hiring/firing process of teachers across countries. On the contrary, in all countries there is ample autonomy in terms of monitoring and organization of school processes. If differences in institutional constraints are a major driver of the Italian low performance, we should find that this is mostly concentrated in the areas in which such constraints are more binding.

When we consider the five macro indicators of managerial practices described in the

previous section, we find that Italy ranks last in all of them. Moreover, the distance from the other countries tend to be similar across areas. This is a first indication of the fact that the differences in managerial practices cannot be simply attributed to the institutional setting in which SPs operate. This indication is confirmed by a more disaggregated analysis. We take each one of the questions and assign a score from 1 (low) to 3 (high) measuring the degree of institutional constraints.[16] We then correlate the degree of constraints with the delay of Italy with respect to the other countries. Contrary to the institutional constraints assumption, we find a negative correlation, meaning that the distance between Italy and other countries is higher in areas where SPs have a greater degree of freedom. We reach a similar conclusion when comparing public to private schools within country: we do not find that SPs in private schools do relatively better in activities in which the institutional constraints are more stringent for public schools.

All in all, this section indicates that Italian SPs score substantially below those of the other five countries in terms of managerial practices. Moreover, such delay cannot be explained simply by differences in the constraints that SPs face in their activities. Rather, they can be attributed to an overall lower quality of managerial practices, due in particular to a large share of SPs with very low scores. This signals that, to improve the managerial quality of Italian SPs, reforming the institutional setting granting schools more autonomy will not be enough: it will also be important to devote specific attention to the selection and training process of SPs.

5 Student performance measures and additional individual-level data

We constructed a database with rich information on student, school and area characteristics. Our main source of additional data is the database provided by the National Institute for the Evaluation of the Educational System of Instruction and Training (INVALSI henceforth), a government agency that carries out a yearly evaluation of student attainment in both Mathematics and Language. The INVALSI standardized tests are compulsory for all Italian schools and students, both public and private, attending specific grades of schooling. In our analysis we focus on the 2010-11 school-year data for tenth grade upper secondary school students.[17]

[16] To save on space, here we only report the main results, referring the interested reader to the online Appendix B, and Table A9 for all the details.

[17] Tests are carried out also by students attending the second and fifth grade (in primary schools) and the sixth and eighth grade (in lower secondary). The Italian school system starts at age six with five years

The 2010-11 was the first school year that these evaluation tests were performed by upper secondary school students. For this reason, the Language test has been intentionally designed by INVALSI to be easier than normal while, conversely, the Math test has been left to a standard level of difficulty in order to precisely measure all skill levels, including the highest.[18] Thus, we exclude the Language test data from our analysis and focus only on the normalized test scores values in Mathematics, obtained as percentage of right answers.

The INVALSI questionnaire is also designed in order to collect detailed information about the student background and family characteristics.[19] In our analysis we include the following additional students demographic information: gender, citizenship (native, first and second generation immigrant students), grade retention and an index of socioeconomic background.[20] The latter is calculated based on the parental occupational status, their educational attainment levels and different measures of household possessions including cultural possessions such as home educational resources and the number of books. This ESCS index for student socioeconomic background is analogous to the same one computed by OECD for the PISA test. The individual scores of this index are obtained by a principal component analysis, with normalized zero mean and unit standard deviation.[21] Our WMS survey dataset on principals and schools is therefore matched to the INVALSI dataset through an anonymous school identifier. Table A1 in the online Appendix sums up the major characteristics of these additional variables for our overall sample, while here we report in Table 2 the specific descriptive statistics for the main sample used in regressions that, for reasons that we explain below, excludes SPs that have not chosen the school. Both averages and standard deviations are very similar to those of the overall sample. We find that our sample is not significantly different from the INVALSI 10th grade students census

of primary school (grades 1 to 5) followed by three years of lower secondary school (grades 6 to 8). Upper secondary education lasts three to five years depending on the type of school chosen.

[18] On this see INVALSI (2011), p.25. Language and Mathematics test scores are very differently distributed and only math is distributed along the whole scale of skill. Similarly to the OECD PISA, the INVALSI standardized tests aim to measure how far students have acquired some of the knowledge and skills essential for full participation in the knowledge society.

[19] Information is collected through a "Family Questionnaire" sent to each family before the test, a "Student Questionnaire" filled by each student the first day of the test and, finally a student general information part compiled from school administrative staff.

[20] First generation are students born abroad of foreign-born parents, while second generation students are native-born children of foreign-born parents. Our dummy retained student is equal to one when students are older than "regular" ones, that is, if, at the end of 2011, he/she is older than 16 years old. Thus, this dummy includes also non-native students who are allocated to a lower grade on the basis of their language skills and not on the basis of a simple age rule.

[21] They are the scores for the first principal component. The index is calculated considering the whole sample of tenth grade upper secondary school Italian students. See also INVALSI (2011) and OECD (2012) for details.

one (data in brackets) in terms of observable characteristics. As expected, the figure for female students is 50%, while the percent of correct answers our sample students get on the math test is 49 (it is 48 for census data). Moreover, first and second generation immigrant students represent respectively, 6% (5.2%) and 2% (2.4%) of our sample (population) and the percentage of retained students is 22% (22%). Finally, both the mean value and the standard deviation of the ESCS index confirm that, even in terms of student socio-economic background, our sample is not biased.

Finally, in order to control for other catchment area characteristics we also use additional data at municipal level provided by the Italian National Institute of Statistics (ISTAT).[22] This dataset enables us to construct a proxy for the wealth level of the school catchment area using data on per capita bank deposit, and to control whether a school is located in densely, intermediate density or sparsely populated areas. As expected for grade 10 students, data in Table 2 show that most upper secondary schools are located in densely populated areas, with only 12% located in rural areas.

6 Empirical framework and identification

We study the effects of managerial practices on student outcomes using a simple regression setting of the form:

$$y_{ij} = \alpha + \beta Manag_j + \gamma X_{ij} + \delta Z_j + \upsilon_{ij} \tag{1}$$

where y_{ij} is an indicator of performance of student i attending school j, $Manag_j$ is the indicator of managerial quality for school j, X_{ij} is a set of individual student controls, Z_j are school, SP and local controls.

Our basic analysis will use OLS regressions. The main problem with this approach is the potential endogeneity of managerial practices. In particular, three selection issues can invalidate the causal interpretation of the β coefficient. First, it might be that more capable SPs self select into schools with better students; second, better teachers might self-select in schools with better managerial practices; third, better students might also do so. Despite the use of a pure cross-sectional dataset, we take particular care in addressing the endogeneity concerns.[23]

[22] These data are provided by ISTAT in the *Atlante statistico dei comuni* dataset.

[23] Of course, the panel dimension would supply useful additional data variability. In our framework, a school fixed effects would take into account selection issues that are dictated by fixed schools attributes, while the presence of a sufficient number of principal transitions would enable us to investigate the pattern of these transitions (i.e. if the least/most effective SPs tend to leave specific type of schools). We plan to investigate these issues in future work.

In terms of SPs self-selection, our data include an extremely rich set of student, SPs and schools controls. We include in all regressions controls for socio-economic characteristics of the students and of the municipality where the school is located, for school types, for SPs demographics. This allows us to control for the most likely sources of endogeneity. For example, assume that more capable SPs self-select into the more prestigious *"liceo classico"*, where students tend to perform better in the standardized test. This is not a problem in our setting. In fact, given that we include a full set of school type dummies, we only use within school type variability of student outcomes to identify the effects of managerial capabilities on student achievements. The same occurs if one is concerned that more capable SPs self-select in schools with students from high income families, as we control for student socioeconomic background.

A second reasons that reduces endogeneity concerns is related to the process through which principals are assigned to schools. For their first assignment, school principals express up to three preferences, choosing among the vacant schools. Afterwards, they can ask to be transferred to a different school. Actual assignments are made by the Regional School Authority (RSA). There is no formal procedure that the RSA must follow. In practice, RSA try to accommodate SPs requests, but have to fill in the positions for the schools that were not chosen by any principal.

Selection issues are unlikely to be at play for SPs that are assigned a school that they have signalled as preferred. In general, the Italian school accountability system does not promote any obvious self-selection process of better SPs in schools with higher ability students. First, there is no reward explicitly linked to student outcomes in standardized tests. Formal self-evaluation procedures are just starting to be introduced in the Italian system and student achievements and results cannot be used for external evaluation and are not even made public at the school level. Second, SP salaries are set by nationwide schedules and depend almost completely on experience and seniority. Indeed, the main concern of the SP is school attractiveness, as the number of students determines the budget, the number of professors etc. that the SP administrates. If enrollment falls below a certain threshold, the school may be merged with another one. Third, even if SPs would actually want to manage schools with better performing students, there is little reason to believe that assignments reflect managerial abilities. In fact, informal conversations with RSA representatives suggest that assignments are based mostly on seniority, i.e., more senior SPs are more likely to get their preferred school. Given that we control for seniority, and that seniority does not seem to be a major determinant of managerial ability, even in this case endogeneity concerns are

not likely to be crucial. Finally, given the generally low mobility rate of Italian workers, a major issue in the school choice decision is likely to be distance from residence.[24]

Selection problems might be more relevant for the schools that are not chosen by the SP but assigned by the RSA. These might indeed be exceptionally "difficult" schools, for which the RSA exerts some informal pressure on SPs that have signalled to be particularly social responsible to take up those schools. In this case, the RSA might try to allocate more capable SPs to more problematic schools. To control for this possibility, in the survey we asked SPs if they were assigned to a school they requested, i.e., they have chosen. In 80 percent of cases this is indeed the case. We use this as our basic sample, excluding SPs that manage a school they did not request.

Although we argue that the selection mechanism is unlikely to give rise to serious endogeneity concerns, we cannot completely rule out this possibility. We will also perform a set of IV regressions and focus on the institutional changes that should determine exogenous shifts in SPs managerial abilities. Since 2000, Italian schools have enjoyed greater organizational autonomy and SPs became also school managers with full responsibility over the school budget. Before that, the main role of SPs was that of "*instructional leader*" rather than of manager and the national competition to acquire the status of upper secondary SP was school-type specific, that is, one could not transfer from Lyceums to a Vocational school. However, it was only in 2004 that a new national competition was introduced to become SP, while the newly selected SPs have been appointed from 2006 onward. Thus, in principle, managerial skills and specific training in human resource management were required to become SP and we can exploit this institutional change in our identification strategy.

A second source of selection is that teachers might choose to move to better managed schools. However, this is not a problem in our framework: rather, it is exactly a key channel through which good managerial practices should operate. Indeed, a SP implementing good managerial practices can affect teachers behavior both through a selection and a treatment effect. Both of these channel are measured by our survey. For example, there are a series of questions on attracting/retaining talents: a good practice is one that operates along this dimension. The same is true in terms of motivating and monitoring behavior: a good practice is one that stimulate effort from teachers. Attracting and retaining better teachers

[24]There is widespread evidence of low willingness to move in the Italian labor market. For example, according to a 1995 survey of the National Institute of Statistics, more than 40% of *unemployed workers* were unwilling to take a job outside the municipality of residence and only 22% were ready to move anywhere Faini, Galli, and Rossi (1996). Specific evidence on teacher mobility can be found in Barbieri, Rossetti, and Sestito (2011).

by offering them a well-functioning working environment is one of the key effects of good managerial practices.

A more serious concern is that the best students might choose to attend schools with better managerial practices. Again, this concern is alleviated by the large set of individual controls we insert in the regressions. In addition, we will address this concerns in Section 8 exploiting the heterogeneity of students by comparing the results for isolated school vs. those with other schools in competition within the school catchment area.

7 Main results

To investigate the relationship between SP managerial skills and student outcomes we estimate a standard education production function where student test performance in mathematics is modeled as a function of SP managerial skills and a set of additional variables that control for personal characteristics (age, gender and tenure in school), student characteristics (gender, socio-economic background, native/I or II generation immigrants, grade retention), school and teachers characteristics (size, type, competition, private/public, teachers turnover) and catchment area characteristics (urbanization, per capita bank deposit and regional dummies). We first discuss the results for the management variables and then move on to the other controls.

7.1 Management variables

We set the scene in Table 3 where we show the OLS results when we alternatively use our six different measures of managerial practices, in addition to the full set of controls. This sample includes 27775 students that are attending schools with principals that lead a school they requested/chosen. Robust standard errors, clustered at principal/school level, are reported in parenthesis. We start in column 1 with our main management practices indicator that represents an overall measure of managerial ability. Results on Management implies that, holding all the other independent variables constant, Italian SPs management practices play a positive role for student performance. In particular, since our dependent variable ranges from 0 to 100, the estimated coefficient value implies that a one unit increase in our management score is expected to increase our average student test score results by more than 2 points, on a mean of 49 out of 100. To give a sense of the magnitude of a unit increase in management, Table 2 shows that the standard deviation of this variable is around .5. More interesting, a unit increase in our management score almost corresponds to the difference between the mean score of all Italian SPs (2.01) and the mean score obtained

by the UK SPs sample (2.98). This meas that, if Italian SPs would have on average the same managerial ability of the UK SPs, student test scores would increase by around 4.6%. When we replicate the model transforming all variables in z-score, we consistently find that a one standard deviation increase in the management score would cause a 0.10 standard deviation increase in student achievement. This result is similar to those found by Bloom et al. (2014) for two of their largest within-country samples, namely, Brazil (0.10) and the US (0.17).[25]

In columns 2 to 6 we investigate the effects of the various dimensions of managerial practices. Column 2 focuses on Leadership, the area of management identified by the education literature as an important channel through which SPs may influence teachers behaviour and, thus, student outcomes. In general, leadership skills are described as the ability of some SPs to engage with the school's staff in ways that motivate them to high levels of efforts and commitment. As discussed above, questions on Leadership ask whether the SP have clearly identified roles and responsibilities within the school and if there is any internal formal accountability system in place. Our analysis confirms that Leadership represents an important area of management for SPs: the estimated coefficient is 2.72, significant at the 5% level.

A second area of management that seems to matter for student outcome is monitoring of school processes. Column 4 suggests that tracking of school performance, reviewing performance with teachers and staff, and the remaining specific managerial skills identified by Monitoring implies a positive and significant coefficient.

Conversely, the remaining individual managerial practices deliver less precise indications and all coefficients are positive but not statistically significant. Column 3 includes our measures on Operations, a management practice which concerns the standardization of the teaching process, the customization of teaching, the use of data in school management and the adoption of best practices. As discussed in Subsection 3.1, compared to other areas of management considered, this is perhaps the one that is closest to the so called *instructional leadership* whose importance in the education literature is highlighted in several works. Column 5 includes results on Targets that examines mainly the type of targets set by the school in terms of student outcomes, their realism and whether they are given consistently

[25]Despite the use of the same methodology to measure SPs managerial practices, when comparing our results with those found in Bloom et al. (2014) some *caveats* need to be taken into account. In general, cross-country measures of school-level pupil outcomes come from different examination results across countries and this makes the single country results difficult to compare. Second, unlike this study, the dependent variables in Bloom et al. (2014) are school-level pupil outcomes measures (rather than individual measures as in our data). As a consequence, they cannot include individual controls and also have fewer controls at the school level.

throughout the school organization. Compared to other countries, Targets represents the individual management area where Italian SPs obtain both the lowest score (1.77) and the largest gap with the best performer country (UK, 2.97 points). Our last management area includes People (column 6) or the incentives section that focus on promotion criteria, pay and bonuses, and fixing or firing bad performers. This is an area in which institutional constraints are likely to be particularly binding, so it is not surprising to find the lowest coefficient of all areas.

Finally, for each area of management we have also replicated the previous analysis introducing an additional control calculated as the average of the remaining managerial domains. In fact, we may have complementarity across managerial practices areas: i.e. the institutional constraints on hiring/firing might therefore affect also the Italian SPs incentives to do monitoring.[26] Results, reported in Table 4, fully confirm the previous ones, with positive effects for Leadership and Monitoring and negative but not statistically significant effects for the remaining areas.[27]

7.2 Other controls

We now turn our attention to examining the results obtained in Table 3 on our additional regressors. In general, we do hardly observe any difference on each of these regressors among our six different specifications. First of all, we focus on three additional SP characteristics, that is, age, gender (women make up 37% of our sample) and tenure in school. Previous studies that quantitatively examine the role of school leaders show mixed evidence regarding the relationship between school performance and principal characteristics (Li, 2012; Clark, Martorell, and Rockoff, 2009). Age is important since, as said above, anecdotal evidence suggests that the more prestigious schools are assigned by the RSA to older SPs as a sort of end of career benefit. Thus, even if we are focusing on the specific sample of SPs that have chosen the school where they operate, it is still possible that younger SPs end up in less prestigious ones. Tenure should control for the possibility that the overall impact of SPs on a specific school increases over time as it takes time for SPs to implement new policies. Thus, without this control our management variables could also capture variation in SPs quality related to differences in length of tenure. Overall, it seems difficult to identify specific

[26] Even if this is not necessarily the case. For example, the cross country PISA survey data on school leadership do not show any obvious complementarity, with countries ranking low on SPs autonomy in hiring and firing, and high on teachers monitoring.

[27] In the online Appendix A we also provide the whole set of results when we include this additional control. Overall, the main results are fully robust, but since the single indexes are highly correlated with the average of the others, when adding interactions or splitting the sample, multicollinearity makes some coefficients somehow volatile across specifications.

SPs demographic characteristics that are good or bad in terms of student performance. In fact, additional principal characteristics, including age, do not seem to matter for student educational outcomes.

Results found in Table 3 for the remaining indicators are largely consistent with the literature. One important exception is the private schools dummy that shows a negative and significant coefficient, implying that private school student performance on standardized tests is worse than that of public Italian school students. This should come as no surprise. Unlike most industrialized countries, private schooling in Italy is associated with poorer rather than better outcomes since the large part of Italian private schools focus more on the recovery of less able students than on across the board high quality education.[28]

Second, as found in previous studies on the Italian case, our results on schools characteristics reveal a significant role for the educational track. Italian upper secondary school tracking is determined by the presence of differentiated curricula rather than by a formal assignment process to academic or vocational courses depending on student past performance or on alternative selection processes. Therefore, Italian students choose schools that specialize in each of the three main curricula: Lyceum, Technical and Vocational. In particular, the vocational/academic intensity is at its lowest/highest level in the Lyceum (with almost no vocational component) and at its highest/lowest level in Vocational schools. In sum, there are significant differences in terms of programs and curricula as well as in the average test scores of students across the three types of schools.[29]

We introduce another dummy variable, "Istituto Superiore", that identifies if a school offers different types of curricula.[30] The coefficient is always negative and significant. This could be an indication that this type of schools, that constitute 37% of our sample, are more complex and difficult to manage and this may affect student outcomes.[31] Further, we find that school size, school competition and teacher turnover do not significantly affect student test results. Italian principals have almost no control over new hires or teacher transfers thus we only have voluntary transitions. This variable too is not significant.

Third, we presents estimates of student characteristics. The results obtained are very

[28] See for example OECD (2012). OECD PISA data show that in 16 OECD countries and 13 partner countries and economies, students in privately managed schools tend to perform better than students in publicly managed schools, but Italy is one of the very few exceptions. See also Brunello and Rocco (2008).

[29] Due to this mechanism, Italian students sorting is also significantly driven by the family background. See Brunello and Checchi (2007). Recent evidence shows that the learning divide due to family background originates in the early stages of the schooling process, in particular in lower secondary schools. See De Simone (2013).

[30] In this case, the school type is identified at class rather than school level.

[31] These schools include combinations of different types of curricula and are usually located in rural areas.

much consistent with the literature. First, our analysis confirms past results on international and Italian data with the presence of a significant "gender gap" in terms of boys outperforming girls in mathematics.[32] Second, the index of socioeconomic background is positively strongly associated with student achievement. Further, estimates confirm that children of immigrants face important gaps in math, with first-generation immigrants representing the most disadvantaged group. We also include a dummy that captures if a student is repeating the grade or if she/he is a retained student and find, as expected, a negative and significant coefficient.

Finally, in terms of area characteristics, we control for whether a school is located in densely populated, intermediate density or sparsely populated areas, and for the wealth level of the school catchment area. We also include regional dummies. In fact, previous studies show that geographical location is an important determinant of Italian student test scores (Cipollone, Montanaro, and Sestito, 2010; Bratti, Checchi, and Filippin, 2007). Moreover, regional dummies are important also because they enable us to control for possible local differences in principals' allocation rules. As said above, since actual assignments are made by each Regional School Authority, both formal and informal assignment rules differ across the different areas and our dummies should capture at least part of this sorting.

8 Robustness and Extensions

8.1 Endogeneity and Sorting

In this section we perform a set of robustness checks of the basic results discussed above. We first address selection issues. If SP managerial skills are not orthogonal to the unobserved component of student test scores, OLS results are clearly biased. Ex-ante, the direction of the bias could go either way. On one side, better SPs might be able to choose schools with better students. This would imply an upward bias of the OLS regressions. On the other, the informal pressure exerted by the RSA on good SPs to take up "difficult" schools may had an impact also on those SPs who declared in our survey to have chosen the school. In this case, OLS estimates would suffer from a downward bias.

To control for the possible endogeneity of managerial practices we exploit the new national competition introduced in 2006 to become SP. The competition was based on an exam that included testing specifically for managerial skills: that is, the SPs appointed

[32] Guiso et al. (2008) show that mathematics gender gaps results in favor of boys are rather unstable and rapidly evolving, with countries with a more gender-equal culture . For example, OECD PISA results report some male advantage in all rounds although not in all countries. See Eurydice (2010).

from 2006 onward might have had better managerial abilities than those selected before the reform. We therefore use a dummy equal to 1 for SP appointed after 2006 as an instrument for management. The rank condition requires that such reform had an impact on the managerial capabilities of SPs selected after the reform was implemented. The exclusion restriction requires that being selected after the reform does not systematically correlate with the student outcomes of the schools such SPs are assigned to (again, conditional on the rich set of observables, as well as on SPs characteristics, such as age, clearly correlated with the instruments). We deem both assumptions as reasonable.

Using this model specification, results on the intent to treat regression and first-stage regression are as expected. The former results show a significant and positive relationship between post-2006 SPs and student outcomes. The idea is that this relationship reflects the effect of the new selection rules on management and first stage results support this hypothesis, with the positive and significant sign of our instrumental variable also on the first stage. To save on space we do not include these results here but only show in Table 3 the second stage results when we alternatively include our six different proxies of managerial practices.

The point estimates of the IV regression reported in Table 5 suggest that, if anything, the OLS largely underestimate the importance of managerial practices. The coefficients on Management, Monitoring and Targets are all significant and larger than the OLS estimates. This is exactly what would result from a process of nonrandom allocation between good SPs and schools with low performing students as previously described. However, we stress that these results need to be interpreted with caution. In fact, the Kleibergen-Paap F statistic signals weak instruments problems. In general, as a rule of thumb, to confirm that our instruments are relevant we should find a first-stage F-statistic larger than ten, but in all columns we find significantly lower values. We therefore create confidence intervals robust to weak instruments that we include among results. For most management areas the Anderson-Rubin AR confidence intervals are conclusive in suggesting a positive effect of management on student results, even if we are still unable to predict the magnitude of the effect, as the interval is very large. Two exceptions are Leadership and Operations that show a disjoint pair of confidence intervals. We take these results as corroborating the OLS ones of a positive effect of managerial practices on student outcomes, but, given the weak instrument problem, acknowledge that the point estimate have to be taken with a grain of salt.[33]

[33]In a previous draft of this paper we have used a different set of instruments exploiting also another reform of the Italian schooling system, reaching similar conclusions. See Di Liberto, Schivardi, and Sulis

Our instrumental variable approach is mainly aimed at solving endogeneity issues arising from SPs self-selection. As pointed out in Section 6, a second concern is that the best students might choose to attend schools with better managerial practices. A simple descriptive evidence shows that the average quality of managerial practices adopted in schools attended by disadvantaged students is almost identical to that observed in high socioeconomic background schools (respectively, 2.01 and 2.09). Second, the characteristics of the Italian upper secondary school system should help to mitigate this problem. Unlike primary or lower secondary schools, in Italy the main students sorting mechanism is across the different types of schools and curricula rather than different schools characteristics, including SPs. We control for the type of school in our regressions, and thus we still have to control for the within type of school students sorting. Since this sorting process arises mostly in densely populated areas, where families can choose between different schools offering the same curriculum, we have run separate regressions for students attending schools that do and do not have other schools in competition within 30 minutes driving distance. The rationale for this exercise is that, for the isolated schools sample, families have no choice and, if endogenous sorting were driving our results, we would expect to find that managerial practices matter more for the sub-sample of schools in competition. Results in Panel A of Table 6 show the opposite: the effect of good managerial practices is larger (a coefficient of 4.97 vs 2.67) in the isolated schools sample. In sum, this evidence rules out this alternative interpretation about the positive relation between test scores and SP, and support the causal interpretation of our OLS estimates.

8.2 Other Robustness Checks

As a first robustness check we focus on the important issue of the role of principals managerial practices when serving different types of students. As stressed in Branch, Hanushek, and Rivkin (2012), case studies and anecdotal accounts find that the effect of a good SP seems most apparent when serving disadvantaged populations, but other studies assume that some specific school policies affect the elite students more than the disadvantaged ones.[34] We therefore run separate regressions on two sub-samples defined on the basis of our school average socio-economic background index: since this indicator has normalized zero mean and unit standard deviation, we split the low and high socio-economic background samples according to this index being below or above zero. When we test the possibility of effect heterogeneity, results in Panel B of Table 6 indicate a larger effect of managerial practices

(2013).

[34]For example, this is the case of discipline in Lazear (2001).

for the disadvantaged group.

This result appears also to contradict another possible explanation for a positive correlation between managerial practices and student outcomes, namely, complementary parental investments. In this case, the positive effect of managerial practices may arise if parents invest more in children trained in schools managed by more able SPs. However, it is plausible that only parents from a high socioeconomic background alter investments in their children in response to changes in schooling inputs. If this is the case, our results seem rather to suggest that SPs practices and parental investments might be substitute more than complements (Datar and Mason, 2008).[35]

A potential mechanism influencing our results is also the possibility that SPs adopt different strategies of cheating, thus inflating test scores to attract better students.[36] For example, SPs may dissuade bad students to show up on the day of the test, they may suggest the correct answers to weak students, or they may directly manipulate students tests to inflate results. Although such behavior may well take place, there are various reasons for ruling out such mechanism in this context. First, there is no built in relation in the way our measures of managerial practices are constructed and possible cheating behavior of SPs. The latter may emerge in a context of high levels of competition between schools to attract better students, or simply in a context of low accountability where moral hazard problems for teachers are possible. The survey questions detailed in Table A10 in the online Appendix C suggest that there is no obvious relation between such behaviour and the quality of managerial practices. Second, our dataset enable also to identify a representative and random sample of monitored classrooms where external inspectors invigilate students during the test and also help to both compute results and prepare the documentation relative to the test. Existing evidence shows that students in the non-monitored classroom received a more "benevolent" supervision, allowing student cheating behaviour more easily (Lucifora and Tonello, 2012). Panel C of Table 6 shows the results when we introduce in our basic specification a dummy for the presence of an inspector in the class plus an interaction term between the latter and our measures of managerial quality. Results do not suggest that SPs that adopt better managerial practices tend also to cheat in test scores: the interaction term is never statistically significant, while results for managerial practices are virtually unaltered with respect to those reported in Table 3.

[35] We also split the sample according to parents education with no significant differences for different groups.

[36] See Angrist, Battistin, and Vuri (2014) and Lucifora and Tonello (2012) on cheating in test scores in Italy.

In Table 7 we return to the baseline OLS model and interact our measure of managerial practices with tenure in school, while keeping the remaining controls as in Table 3. School tenure should capture accumulation of skills that are specific to the school environment in which the SP is operating, ranging from organization and motivation of teachers and administrative staff to a general knowledge of the school operations or particular issues related to the school specificity. Results, reported in Table 7, are mixed. First, the effect of managerial skills increases and becomes more significant with respect to the baseline results. Second, the effect of tenure is positive, although it is marginally significant only for the dimensions Leadership and Monitoring. The interaction term between managerial practices and school tenure is negative and marginally significant in the same two cases as above.[37] One possible explanation is that newly appointed SPs do rely relatively more on their managerial skills when they are assigned to a new school, implementing formal managerial practices and methods to gain leadership and accumulate knowledge about the school, while they rely more on informal and less managerially oriented methods when they have accumulated tenure in the school.[38]

As we have discussed above, an important source of variability in terms of managerial practices is given by cohort effects. In fact, for SPs appointed before the reform of year 2000 there was little emphasis on managerial capabilities and more on instructional leadership. These SPs might therefore tend to manage the school relying less on the formal managerial procedures investigated in our survey. For this sample, therefore, this measure might be less telling than for the sample of SPs first appointed after 2000. Moreover, this cohort might be characterized by lower managerial ability but, possibly, be assigned to better schools, as seniority plays a role in the assignment process. In what follows, we further investigate this issue by estimating our baseline OLS model excluding SPs appointed up to year 2000. We expect that for this restricted sample of SPs good managerial practices should matter more for student outcomes. Results reported in Table 8, obtained on a sample of 14060 observations, confirm our a priori expectations: for the sample of SPs that were firstly appointed after 2000, we find a large and statistically significant positive effect for all dimensions of managerial abilities on student outcomes. The coefficient for the overall managerial index reported in column 1 of Table 8 indicates that an increase in the index by one point increases student scores by more than 6 points with an increase of about

[37] To capture non linear effects, we also tried a different specification including a quadratic term for school tenure and interacting managerial practices accordingly, results are virtually unchanged.

[38] Indeed, results in the literature on the effect of tenure of SPs on student outcomes are mixed. See Branch, Hanushek, and Rivkin (2012), Coelli and Green (2012) and Dhuey and Smith (2012) among others.

12%, an effect that is almost 3 times as large as the one found in Table 3 on the sample of SPs that chose the school and including all cohorts of SPs. The effect of management on student outcomes turns positive and statistically significant when we also use other measures of managerial skills. As stressed above, while the effect in the sections Leadership and Monitoring are confirmed, it is interesting to note that an increase by one point in managerial abilities related to People and Targets increases student scores by more than 5 points, suggesting that for this cohort of SPs all dimensions of managerial abilities matter for the good performance of students. Finally we report a positive effect for Operations practices, even if the magnitude of the coefficient is the smallest across dimensions (2.15).

In Section 4, we have stressed the importance of institutional constraints as a factor that may help to explain the observed gap in managerial abilities between Italian SPs and those operating in other countries and we have also classified each single question according to the relevance of such institutional constraints. When institutional constraints are binding, as it is the case in Italy for the the human resource management area, we expect that management scores are more similar across SPs and that SPs activities may be less effective. Conversely, we should expect that managerial practices can make more of a difference in areas where the institutional constraints are less binding. In what follows, we exploit this classification to group the individual questions into three different measures of managerial practices classified as low, medium and high level of institutionally constraint activities, with each managerial practices variables obtained as the mean of the corresponding questions (see the online Appendix B and Table A9 for details about the classification of each question). Results reported in Table 9 indicate a consistent pattern: managerial practices have a positive and statistically significant effect (about 2.5 points) on student outcomes only when we consider the index constructed using areas of managerial practices that have low institutional constraints (see column 1), while we observe a positive but not statistically significant effect when we consider the other two indexes in columns 2 and 3.[39]

We finally investigate the issue of what kind of students (in terms of school test results) benefit more from good managerial practices.[40] To this aim, we run quantile regression models for the 25, median and 75 percentiles. The results, which are reported in Table 10, first reveal that our management coefficients are significant for the leadership and monitor-

[39] We run such regressions on the sample of SPs that chose the school used in Table 3 to estimate our baseline model.

[40] In a previous version of this paper, we have also run probit regressions in which the dependent variable is equal to one if the student is lagging behind with respect to the age/grade ladder and zero otherwise finding that good managerial practices reduce the probability of student retention by 3 percentage points, with an effect that is almost constant across different dimensions of management. See Di Liberto, Schivardi, and Sulis (2013).

ing areas and for the overall management index, as for the OLS regressions. Second, for the latter index the evidence suggests that the effect of managerial practices is significant only for the median and 75 percentile, while the two areas of management seem to affect students in a more similar way. Overall, this analysis suggests that,if anything, the effects of managerial practices are lower for the left tail of student outcomes. Due to the high correlation existing between student performance and socio-economic background, these results seem at odds with those found in Table 6 for the two sub-samples of low and high socio-economic background schools. We therefore further investigate this issue and run quantile regressions also on these two separate sub-samples.[41] In this case, we find that the coefficients on management is positive and significant only for the sub-sample of low socio-economic background school, while coefficients are non significant for the remaining schools. That is, the above significant results were mainly driven by the disadvantaged schools sample. Further, for this sub-sample we also consistently find that the effect is stronger in the upper part of the test score results distribution: that is, the weakest students are the more difficult to recover, but the good students of disadvantaged schools are those benefiting more from well organized schools.

9 Comparing WMS with OECD PISA

Since 2006, the OECD includes a section in the PISA (Programme for International Student Assessment) survey that investigates SPs managerial practices.[42] This data collection system is based on a self-reported measures and it is clearly much simpler (and cheaper) than the one of the WMS. It is therefore interesting to investigate if the two surveys supply comparable information. If this were the case, one could resort to the simpler, cheaper collection method to obtain reliable indicators of managerial practices.

We use the most recent PISA release of 2009 where we identify three indices of managerial practices. The main one is the index of principal leadership (LDRSHP), which captures if principals are active in improving both schools teaching practices and the working environment. The PISA dataset constructs this index as the average of the answers that the SP, or a designate, reports in Q26 of the School Questionnaire on the frequency with which fourteen areas of activities are performed during the school year. The indicator ranges from 1 (never) to 4 (very often). There is therefore a clear difference between what the PISA and the WMS measure. In fact, PISA is a measure of self-reported effort, while in the WMS the

[41] To save on space we do not report these results, but they are available upon request.
[42] An analysis on the 2006 data can be found in Pont et al. (2008).

interviewer directly assesses the quality of the implemented practices. In addition to the overall indicator, the fourteen specific items can be classified in terms of three of the WMS key areas of management, namely, Leadership, Operations and Monitoring.[43] This enables us to construct three additional sub-indices to be used in our comparison.

A second PISA variable that can be used as a measure of school management activities is the index of school responsibility for resource allocation (RESPRES). In this case the PISA survey asks to report whether (and/or) 1) principals, 2) teachers, 3) school governing board, 4) regional or local education authority or 5) national education authority has/have responsibility for both human resource management and in deciding on budget allocations within the school. Finally, a third index, calculated as the previous one, refers to the school responsibility for student curriculum and assessment (RESPCURR). Positive values on these two final indices imply relatively more responsibility for schools than local, regional or national education authorities. The three indices, LDRSHP, RESPRES and RESPCURR are all standardized with (OECD) mean zero and a standard deviation one.

We start by comparing the two main indices, that is, LDRSHP for PISA and Management for WMS for the six countries in our WMS sample. To make the indicators comparable, we use the average values of Q26 items and rescale both the WMS and PISA data into a range from 0 to 1. Figure 3 reports the average values at the country level. First, note that the PISA indicator is in all cases substantially larger than the WMS one. This is evidence that self reported values might inflate the measure with respect to that based on the assessment of an external reviewer. Second, the two surveys display significant differences even in terms of country ranking: while the UK is the best performer in both datasets, Sweden performs relatively worse in the PISA ranking while the opposite is true for Italy.

In Figure 4 we report the distribution of both our WMS management measure (in bright-yellow) and the PISA main index (in darker-green). This figure clearly shows that, unlike the WMS data, the PISA index distribution is, for all countries, concentrated on high values. In particular, for the UK and Italy (the best and the worst performers according to the WMS) the two distribution seldom overlap. We interpret this as a further indication that the PISA data suffer self-assessment bias and/or that the frequency with which certain activities are performed has little to tell about the quality of the managerial practices.[44]

[43]Leadership includes answers to sub-questions (a), (k) and (l); Operations items (b), (c), (e), (h), (j) and (m); Monitoring (d), (f), (g) and (i). The last question/item (n), explores if the SP takes over lessons from absent teachers, an issue which is not explicitly covered by our WMS survey. Note that we also have replicated the analysis using a new variable of PISA-Management that exclude this item, and results are almost identical.

[44]For example, the LDRSHP is based on the frequency with which the principal is involved in a series of organizational activities: about 75% of the 3026 principals/schools interviewed in the PISA six country

We have also used an alternative WMS management indicator that takes into account only the three areas of management (Leadership, Operations and Monitoring) included in the LDRSHP index provided by PISA but this hardly changes the picture.[45] We next move on to correlate student achievements with the indicators of managerial practices, following the same scheme as before. The dependent variable is the individual student PISA test scores in mathematics, while the set of regressors is chosen in order to replicate our baseline specification of Table 3.[46] The only exceptions are the SPs demographic variables, not included in the PISA dataset.

Table 11 shows the OLS results on the Italian sample. With more than 1000 schools, this is one of the largest country sample in PISA 2009. The first row reports the estimated coefficients of the six measures of PISA managerial practices: LDRSHP (overall leadership measure), RESPCURR (responsibility over curricula) and RESPRES (responsibility over expenditure) in columns from 1 to 3, and the 3 sub-indices of Leadership, Operations and Monitoring in columns from 4 to 6. Robust standard errors, clustered at principal/school level, are reported in parenthesis. We find that the measures of managerial practices are always statistically insignificant.[47] The coefficients of the additional controls are very similar with that found in Table 3, suggesting that the difference is in variables that measure managerial practices rather than in other features of the survey.

We also extend the same analysis to all countries, exploiting the cross-country dimension of the PISA dataset by including the 33 available OECD countries in the regressions. Results, reported in Table A8 in the online Appendix, lead to very different conclusions from the ones based on the WMS: in the PISA sample the effect of managerial practices is always negative and, in two cases (LDRSHP and Operations), statistically significant. This is a puzzling evidence: although one might argue against the importance of SPs managerial practices for student performance, it seems hard to justify a *negative* effect. One possibility

sample rates her/his overall activity between 3 and 4 (recall that the range is 1-4), the remaining 25% between 2 and 3, while not a one reports a value between 1 and 2.

[45] See Figure A1 reported in the online Appendix.

[46] Among controls we include: a set of dummies for the number of school in competition (Q5), the number of students in the school (SCHSIZE), a dummy for private schools (SCHTYPE), the share of part-time teachers with respect to the total number of teachers, student immigration status (IMMIG), their gender, their socio-economic background (ESCS), and an indicator for students repeating grades (Q7). We also control for area characteristics, including dummies for population density and alternatively regional-type of school. When we focus on the sample of 33 OECD countries we include countries fixed effects. Data for France are not available.

[47] We have replicated the same analysis with an alternative measure of management that excludes the last item of Q26, that refers to a managerial activity not explored by the WMS survey, and we obtain the same results.

is that less capable SPs are also more indulgent in self-assessing their performance.[48]

Overall, this analysis does not support the assertion that the WMS and the PISA surveys supply comparable indicators of management practices. Both the direct comparison and the regression analysis reveal marked differences between the two surveys. These disparities are likely to be due to both methodological differences in data collection and possible mis-measurement of the self-reported PISA managerial indices. They suggest that self assessment cannot substitute the direct assessment of SPs managerial capabilities by a third party.

10 Conclusions and Policy Implications

This study investigates if the quality of managerial practices in schools affects student achievements. To this aim, we measure managerial practices using the World Management Survey, an innovative methodology that enables us to construct robust measures of management quality comparable across countries and overcome standard mis-measurement problems that are typical of self-assessed measures of managerial ability that characterized previous studies in the field. Moreover, this methodology identifies the specific areas of managerial activity that matter most for successfully leading a school.

We first compare data on six industrialized countries (Canada, Italy, Germany, Sweden, the UK, the US) and find substantial heterogeneity in managerial practices across these countries. Countries that obtain the lowest scores on schools managerial quality (Italy and Germany) are also characterized by more centralized systems, with less autonomy at school level in many areas of management. Second, our analysis focuses on the Italian case for which we are able to match our indexes of managerial practices at the school level with the student outcomes in the maths standardized test. In particular, in order to investigate the relationship between SPs managerial practices and student outcomes we estimate a standard education production function where student test performance in mathematics is modeled as a function of SP managerial skills and of rich set of student, SPs and schools additional controls. We argue that our set of covariates allows us to control for the most likely sources of endogeneity, in particular, self selection of best principals into best schools.

The OLS estimates imply that increasing managerial practices by one unit, student test results would increase by about 4.6%. In other words, if Italian SPs implemented managerial practices of the same average quality as the the UK SPs (the best performer

[48] An alternative explanation is selection: better SPs are assigned to worse schools. To assess this possibility, one would need to consider the SPs appointment process in each country, something clearly beyond the scope of this paper.

in terms of management), Italian students would close the gap in the math OECD PISA test with respect to the OECD average. In terms of specific areas of managerial practices, we find that the coefficients are positive in all categories, although statistically significant only for leadership and monitoring activities of school processes. The results survive a large series of checks, including the most likely endogeneity and selection concerns. All our exercises confirm that management quality is an important input of our estimated education production function.

Overall, the policy implications of these results are clear cut. A well-run school improves student achievements. Governments should therefore enact policies that take into account principal selection and training in terms of managerial capabilities. Policies that explicitly target SPs are also likely to be cost effective with respect to those only targeting teachers. In fact, although teachers are the key players in any school system, improving their teaching skills directly will generally require a substantial amount of resources, given the large number of subjects involved. Of course, policies directed at SPs and teachers are complementary and ideally should be pursued simultaneously. In a period of tight public finances, however, it might be more feasible to start from SPs.

Another important issue relates to the "partial" versus the "general equilibrium" effects of improving the managerial capabilities of SPs. If the effects of managerial practices work mostly through allowing teachers to work in a better organized environment, and therefore teach more effectively, an improvement in the average managerial capabilities of SPs will immediately improve student performance. On the contrary, if the effects of better management work mostly through attracting and retaining good teachers, a general increase in the managerial capabilities of SPs will not affect the average student achievement, *given the population of teachers*. Even in this case, however, we should expect long run effects of better managerial practices in schools, as a well managed school system will increase the appeal of the teaching profession. We have no direct evidence on the relative importance of the two channels. An indirect indication can be obtained by considering the subcategories of managerial practices. The two categories that tend to be significantly related to student performance across specifications are leadership and monitoring. These areas are more likely to affect the behaviour of the current population of teachers rather than their mix. Instead, the category related to human resource management is never significantly correlated with students performance, possibly because of the high degree of institutional constraints that SPs face along this dimension. This suggests that at least part of our results come from the better performance of the given teachers pool, and that improving the average managerial

practices of all SPs would immediately improve student achievements.

Finally, as argued by Hanushek, Link, and Woessmann (2013), a growing body of evidence suggests that the combination of school autonomy and accountability is the most effective policy to improve student performance in well developed countries. As a consequence, many countries are injecting more autonomy in the school system. This will imply that managerial competencies of school principals will become even more important as this process develops, pointing to the risk of giving autonomy without improving their managerial capabilities.

References

Angrist, Joshua D, Erich Battistin, and Daniela Vuri. 2014. "In a Small Moment: Class Size and Moral Hazard in the Mezzogiorno." *NBER Working Paper* 20173.

Barbieri, Gianna, Claudio Rossetti, and Paolo Sestito. 2011. "The determinants of teacher mobility: Evidence using Italian teachers transfer applications." *Economics of Education Review* 30 (6):1430–1444.

Barro, Robert J and Jong-Wha Lee. 1994. "Sources of economic growth." In *Carnegie-Rochester conference series on public policy*, vol. 40. 1–46.

Bertrand, Marianne and Antoinette Schoar. 2003. "Managing with style: The effect of managers on firm policies." *The Quarterly Journal of Economics* 118 (4):1169–1208.

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, Raffaela Sadun, and John Van Reenen. 2014. "Does management matter in schools?" *NBER Working Paper* 20667.

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.

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

———. 2010b. "Why do management practices differ across firms and countries?" *The Journal of Economic Perspectives* 24 (1):203–224.

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

Böhlmark, Anders, Erik Grönqvist, and Jonas Vlachos. 2012. "The headmaster ritual: the importance of management for school outcomes." *IFAU Working Paper* 201216.

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." *NBER Working Paper* 17803.

Bratti, Massimiliano, Daniele Checchi, and Antonio Filippin. 2007. "Geographical Differences in Italian Students' Mathematical Competencies: Evidence from PISA 2003." *Giornale degli Economisti e Annali di Economia* :299–333.

Brewer, Dominic J. 1993. "Principals and student outcomes: Evidence from U.S. high schools." *Economics of Education Review* 12 (4):281–292.

Brunello, Giorgio and Daniele Checchi. 2007. "Does school tracking affect equality of opportunity? New international evidence." *Economic Policy* 22:781–861.

Brunello, Giorgio and Lorenzo Rocco. 2008. "Educational Standards in Private and Public Schools." *The Economic Journal* 118 (533):1866–1887.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2011. "The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood." *NBER Working Paper* 17699.

Cipollone, Piero, Pasqualino Montanaro, and Paolo Sestito. 2010. "Value-Added Measures in Italian High Schools: Problems and Findings." *Giornale degli Economisti* 69 (2):81–114.

Clark, Damon, Paco Martorell, and Jonah Rockoff. 2009. "School Principals and School Performance. Working Paper 38." *National Center for Analysis of longitudinal data in Education research* .

Coelli, Michael and David A Green. 2012. "Leadership effects: School principals and student outcomes." *Economics of Education Review* 31 (1):92–109.

Cullen, Julie Berry and Michael J. Mazzeo. 2008. "Implicit Performance Awards: An Empirical Analysis Of The Labor Market For Public School Administrators." *mimeo.*

Datar, Ashlesha and Bryce Mason. 2008. "Do reductions in class size crowd out parental investment in education?" *Economics of Education Review* 27 (6):712–723.

De Simone, Gianfranco. 2013. "Render unto Primary the Things which are Primary's. Inherited and Fresh Learning Divides in Italian Lower Secondary Education." *Economics of Education Review* 35 (C):12–23.

Dhuey, Elizabeth and Justin Smith. 2012. "How School Principals Influence Student Learning." *Society for Research on Educational Effectiveness* .

Di Liberto, Adriana, Fabiano Schivardi, and Giovanni Sulis. 2013. "Managerial practices and students performance." *FGA Working Paper* 49.

Eurydice. 2010. "Gender Differences in Educational Outcomes." Tech. rep., European Commission.

Faini, R., G. Galli, and F. Rossi. 1996. "Mobilità e disoccupazione in Italia: un'analisi dell'offerta di lavoro." In *La mobilità della società italiana*, edited by G. Galli.

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

Guiso, Luigi, Monte Ferdinando, Paola Sapienza, and Luigi Zingales. 2008. "Culture, Gender, and Math." *Science* 320:1164–1165.

Hallinger, Philip and Ronald H Heck. 1998. "Exploring the Principal's Contribution to School Effectiveness: 1980-1995." *School effectiveness and school improvement* 9 (2):157–191.

Hanushek, Eric A, Susanne Link, and Ludger Woessmann. 2013. "Does school autonomy make sense everywhere? Panel estimates from PISA." *Journal of Development Economics* 104:212–232.

Hanushek, Eric A and Steven G Rivkin. 2006. "Teacher quality." *Handbook of the Economics of Education* 2:1051–1078.

———. 2012. "The distribution of teacher quality and implications for policy." *Annu. Rev. Econ.* 4 (1):131–157.

Hanushek, Eric A. and Ludger Woessmann. 2011. "How much do educational outcomes matter in OECD countries?" *Economic Policy* 26 (67):427–491.

INVALSI. 2011. "Rapporto tecnico sulle caratteristiche delle prove INVALSI 2011." Tech. rep., INVALSI.

Jacob, Brian A. 2011. "Do principals fire the worst teachers?" *Educational Evaluation and Policy Analysis* 33 (4):403–434.

Lavy, Victor. 2008. "Does Raising the Principal's Wage Improve the School's Outcomes? Quasi-experimental Evidence from an Unusual Policy Experiment in Israel." *The Scandinavian Journal of Economics* 110 (4):639–662.

Lazear, Edward P. 2001. "Educational Production." *Quarterly Journal of Economics* 116 (3):777–803.

Li, Danielle. 2012. "School Accountability and Principal Mobility: How No Child Left Behind Affects the Allocation of School Leaders." *mimeo, MIT*.

Loeb, Susanna, Demetra Kalogrides, and Tara Béteille. 2012. "Effective schools: Teacher hiring, assignment, development, and retention." *Education Finance and Policy* 7 (3):269–304.

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

Lucifora, Claudio and Marco Tonello. 2012. "Students' Cheating as a Social Interaction: Evidence from a Randomized Experiment in a National Evaluation Program." *IZA Working Paper* 6967.

OECD. 2012. *Public and Private Schools: How Management and Funding Relate to their Socio-economic Profile*. Paris, OECD Publishing.

Pont, Beatriz, Deborah Nusche, Hunter Moorman, and Hopkins David. 2008. *Improving school leadership, Volume 1: Policy and practice*. Paris, OECD Publishing.

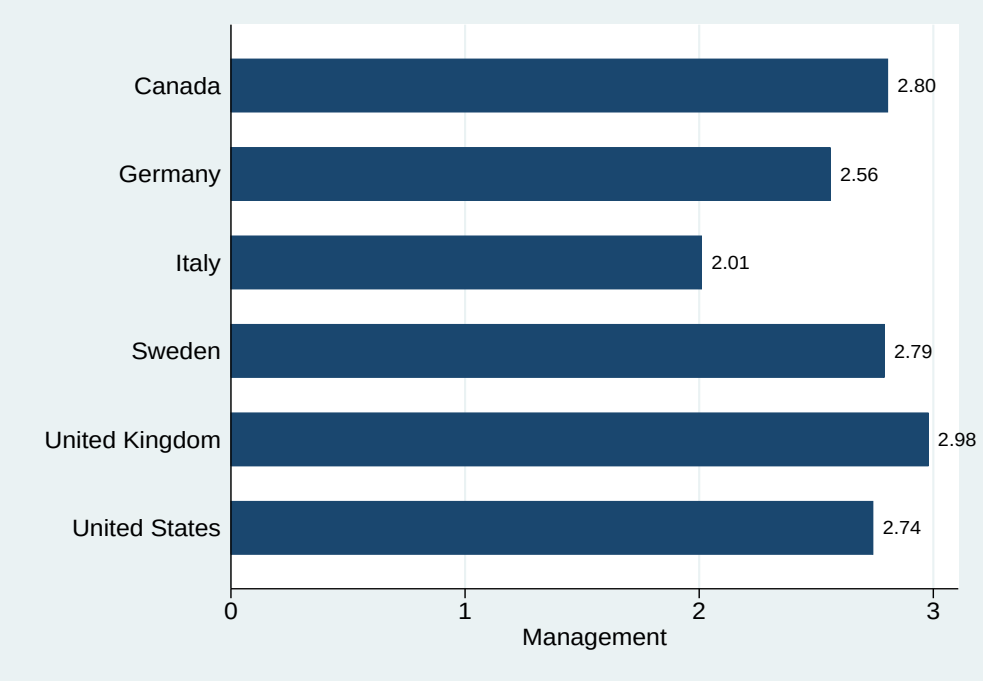
Rivkin, Steven G, Eric A Hanushek, and John F Kain. 2005. "Teachers, schools, and academic achievement." *Econometrica* 73 (2):417–458.

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* 44:635674.

Rockoff, Jonah E. 2004. "The impact of individual teachers on student achievement: Evidence from panel data." *The American Economic Review* 94 (2):247–252.

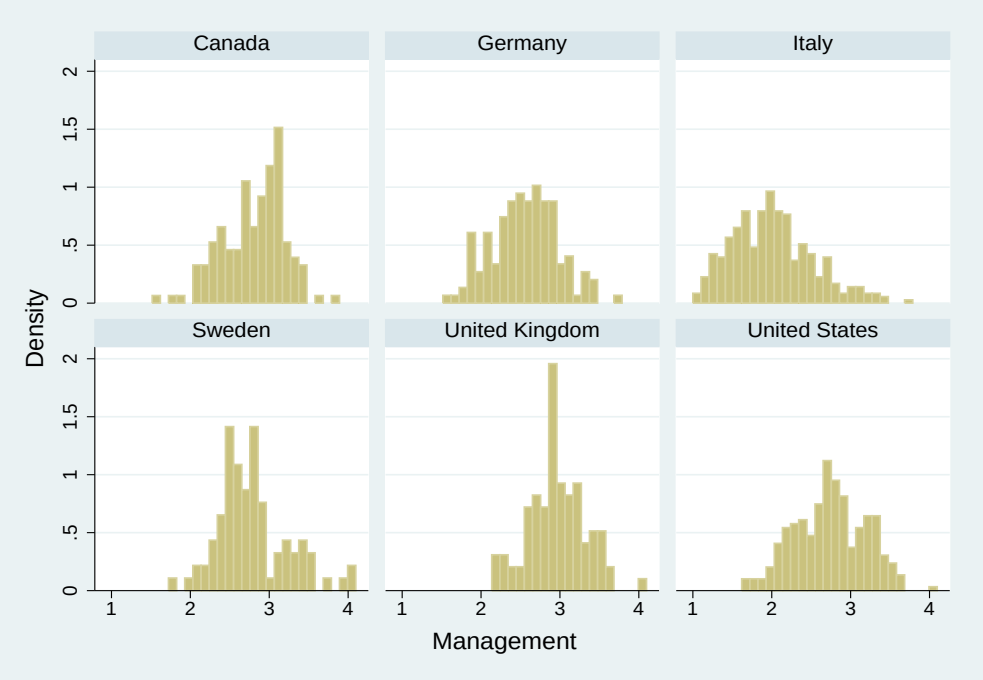
Figures and Tables

Figure 1: Average Managerial Practices across Countries



Notes: For WMS cross country data, excluding Italy, see Bloom et al. (2012).

Figure 2: Distribution of Managerial Practices within Countries



Notes: For WMS cross country data, excluding Italy, see Bloom et al. (2012).

Figure 3: PISA and WMS School Management Indices Rescaled

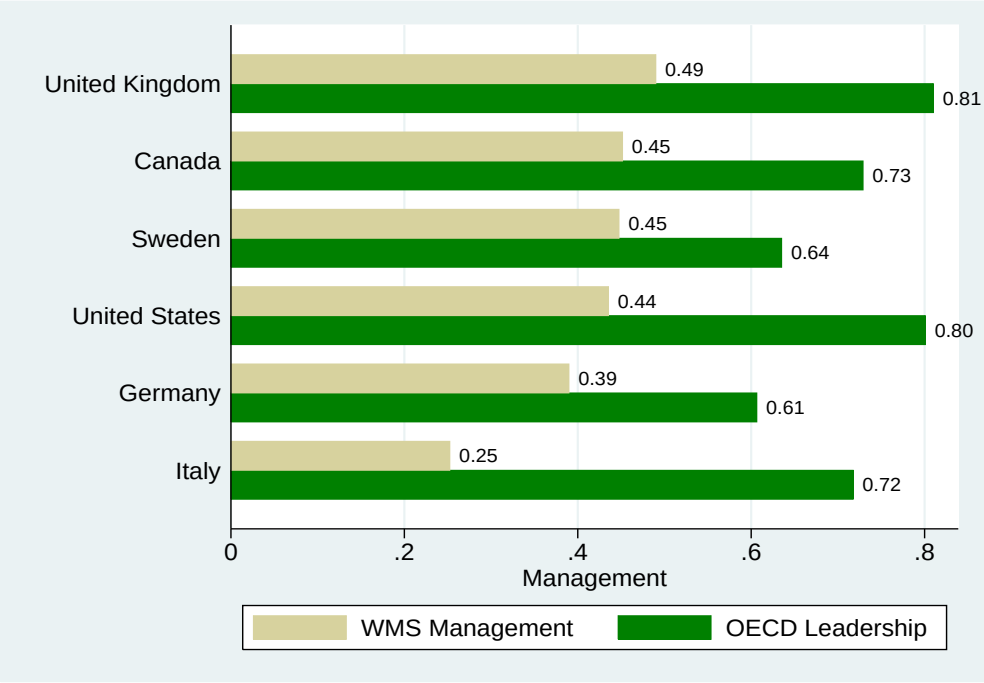


Figure 4: Distribution of PISA and WMS School Management Indices Rescaled

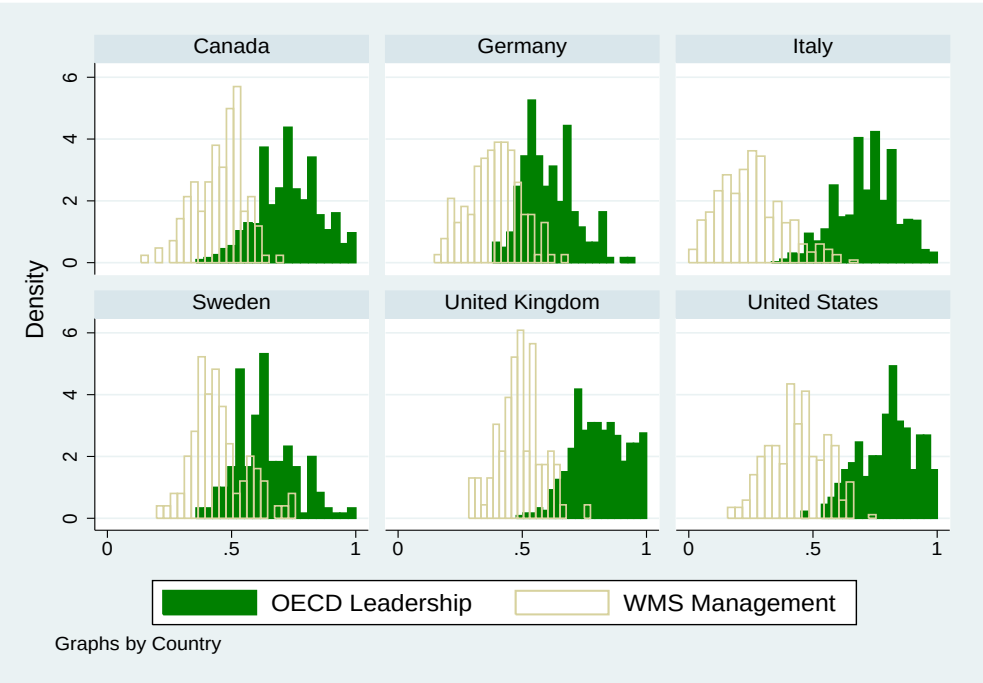


Table 1: Descriptive Statistics: Cross Country Comparison

	Canada	Germany	Italy	Sweden	United Kingdom	United States
percentiles						
10th	2.26	1.97	1.37	2.32	2.53	2.15
50th	2.87	2.55	1.99	2.74	2.95	2.75
90th	3.26	3.10	2.71	3.40	3.44	3.32
mean	2.81	2.56	2.01	2.80	2.96	2.74
st. dev.	0.39	0.41	0.51	0.44	0.36	0.45
percentile ratios						
90/10	1.44	1.57	1.98	1.47	1.36	1.54
50/10	1.27	1.29	1.45	1.18	1.17	1.28
90/50	1.14	1.22	1.36	1.24	1.17	1.21
observations	147	143	341	89	94	285

Notes: For more on cross country data see Bloom et al. (2012, 2014). Our overall measure of Management includes all sections, i.e., Leadership, Operations, Monitoring, Targets and People.

Table 2: Descriptive statistics

Variable	Mean	Std. Dev.	Min.	Max.	Obs.
Panel A: School/Principal Characteristics					
management	2.01	0.5	1	3.74	260
people	1.82	0.47	1	3.5	260
targets	1.78	0.54	1	3.6	260
operations	2.25	0.68	1	4.5	260
leader	2.19	0.58	1	4	260
monitoring	2.16	0.68	1	4.60	260
principal's age	58.18	6.73	32	80	257
gender of principal, female=1	0.37	0.48	0	1	260
tenure in school	6.55	5.86	1	40	259
competition	2.34	2.34	0	11	259
school size	6.27	0.8	3.93	7.5	260
dummy private school	0.17	0.37	0	1	260
teacher turnover	4.78	6.13	0	50	257
dummy=1 post-2006 reform	0.32	0.47	0	1	260
Panel B: Student Characteristics					
test score math normalized	49.04	17.95	0	100	29198
Lyceum	0.48	0.5	0	1	29649
Vocational	0.2	0.4	0	1	29649
Technical	0.32	0.47	0	1	29649
dummy "istituto superiore"	0.37	0.48	0	1	32728
immigrant 1st generation	0.06	0.24	0	1	32366
immigrant 2nd generation	0.02	0.15	0	1	32366
gender of student, female=1	0.5	0.5	0	1	32361
escs, socio-economic background	0.03	0.98	-3.66	2.07	28828
student behind	0.22	0.41	0	1	32358
dummy rural area	0.12	0.32	0	1	32728
dummy non rural area	0.35	0.48	0	1	32728
dummy urban area	0.54	0.5	0	1	32728
ln of bank deposits per capita	4.27	1.2	0	5.38	32728

Notes: This descriptives are based on the sub-sample of schools/SPs used during the regression analysis of SPs. In panel A data are aggregated at the principal level, in Panel B at student level. *principal's age* and *tenure* are measured in years, *competition* is the number of schools with similar curricula within 30 minutes drive from school, *dummy post-2006 reform* is equal to 1 for principals with 4 years (or less) of experience, *school size* is the natural logarithm of the number of students in school, *teacher turnover* is the share of teachers that left the school in previous year, *immigrant 1st generation* are students born abroad of foreign-born parents, *immigrant 2nd generation* are native-born students of foreign-born parents, *student behind* is a dummy for students older than regular ones.

Table 3: Baseline model

Dependent Variable Test results (Math)	(1) Management	(2) Leadership	(3) Operations	(4) Monitoring	(5) Targets	(6) People
managerial practices	2.24*	2.72**	1.20	2.36**	0.98	0.73
	(1.31)	(1.13)	(0.86)	(0.96)	(1.24)	(1.44)
age principal	0.17	0.17	0.16	0.19*	0.14	0.13
	(0.12)	(0.12)	(0.12)	(0.11)	(0.12)	(0.12)
gender principal (female=1)	0.44	0.28	0.53	0.63	0.60	0.69
	(1.31)	(1.27)	(1.31)	(1.25)	(1.35)	(1.32)
tenure in school	-0.03	-0.03	-0.01	-0.03	-0.02	-0.02
	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
competition	0.25	0.29	0.22	0.27	0.22	0.21
	(0.27)	(0.27)	(0.28)	(0.27)	(0.28)	(0.28)
school size	-0.99	-1.18	-0.83	-0.92	-0.87	-0.80
	(1.01)	(1.00)	(1.03)	(1.01)	(1.03)	(1.01)
private school	-5.67**	-5.69***	-5.35**	-5.00**	-5.49**	-6.00**
	(2.22)	(2.18)	(2.23)	(2.20)	(2.23)	(2.40)
teacher turnover	0.11	0.12	0.11	0.11	0.11	0.11
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
lyceum school	5.20***	5.22***	5.09***	5.19***	5.11***	5.18***
	(1.33)	(1.32)	(1.35)	(1.28)	(1.35)	(1.36)
vocational school	-9.25***	-9.27***	-9.27***	-9.23***	-9.41***	-9.40***
	(1.32)	(1.33)	(1.35)	(1.30)	(1.35)	(1.34)
"istituto superiore"	-2.85**	-2.93**	-2.72**	-2.79**	-2.76**	-2.62**
	(1.21)	(1.20)	(1.19)	(1.17)	(1.25)	(1.21)
immigrant 1st generation	-1.98***	-1.87***	-1.94***	-2.08***	-1.99***	-1.96***
	(0.60)	(0.60)	(0.60)	(0.60)	(0.60)	(0.60)
immigrant 2nd generation	-2.26***	-2.21***	-2.23***	-2.30***	-2.31***	-2.29***
	(0.63)	(0.63)	(0.63)	(0.62)	(0.63)	(0.63)
gender_student (female=1)	-6.60***	-6.59***	-6.62***	-6.59***	-6.66***	-6.66***
	(0.53)	(0.53)	(0.54)	(0.52)	(0.54)	(0.54)
socio_economic background	1.07***	1.06***	1.08***	1.09***	1.07***	1.06***
	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)
student behind	-5.22***	-5.18***	-5.25***	-5.22***	-5.27***	-5.28***
	(0.40)	(0.40)	(0.41)	(0.40)	(0.40)	(0.40)
dummy rural areas	-0.26	-0.50	-0.29	0.10	-0.42	-0.45
	(2.04)	(1.97)	(2.03)	(2.05)	(2.04)	(1.99)
dummy non rural areas	0.03	-0.15	-0.10	0.33	-0.10	-0.19
	(1.20)	(1.17)	(1.18)	(1.24)	(1.18)	(1.16)
ln deposits per capita	0.80**	0.83**	0.77**	0.81**	0.79**	0.77**
	(0.38)	(0.38)	(0.39)	(0.39)	(0.38)	(0.38)
Constant	42.25***	42.58***	42.96***	39.65***	45.72***	46.96***
	(11.73)	(11.32)	(12.08)	(11.35)	(11.55)	(11.41)
Observations	27775	27775	27775	27775	27775	27775
R-squared	0.250	0.252	0.248	0.253	0.247	0.247

Notes: Standard errors clustered at school level in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variable is standardized test score in mathematics. All regressions include regional dummies. *Age of principal* and *tenure in school* are measured in years. *Competition* is the number of schools with similar curricula within 30 minutes drive. *School size* is the natural logarithm of the number of students in school. *Private* is a dummy equal to one for private schools. *istituto superiore* is a dummy for schools offering different curricula (See also Section 7.2). *Lyceum* and *vocational* are dummies equal to 1 for each school type (technical school is the omitted dummy). *Teacher turnover* is the share of teacher turnover in the previous school year. *Immigrant 1st generation* are students born abroad of foreign-born parents, *immigrant 2nd generation* are native-born children of foreign-born parents. *Student behind* is a dummy equal to 1 if the student is older than regular. *Ln of deposits* is the natural logarithm of bank deposits per capita at the municipal level. Dummies for urbanization are equal to 1 for rural areas and non rural areas (urban area is the omitted category).

Table 4: Model with other areas of management

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Test results (Math)	Management	Leadership	Operations	Monitoring	Targets	People
managerial practices	2.24*	3.45**	-0.72	3.55**	-2.86	-2.71
	(1.310)	(1.592)	(1.522)	(1.596)	(1.818)	(1.811)
other areas of man.		-1.04	3.41	-2.05	4.77**	3.89***
		(1.763)	(2.466)	(2.138)	(1.931)	(1.470)
principal characteristics	Yes	Yes	Yes	Yes	Yes	Yes
school characteristics	Yes	Yes	Yes	Yes	Yes	Yes
student characteristics	Yes	Yes	Yes	Yes	Yes	Yes
area characteristics	Yes	Yes	Yes	Yes	Yes	Yes
regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,775	27,775	27,775	27,775	27,775	27,775
R-squared	0.250	0.253	0.251	0.254	0.253	0.252

Notes: Standard errors clustered at school level in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The set of additional controls is identical to that included in Table 3: these results are not reported but available upon request. Other areas of management is calculated as the average of the other (excluded) areas.

Table 5: IV model

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Test results (Math)	Management	Leadership	Operations	Monitoring	Targets	People
managerial practices	19.32*	29.52	30.73	11.82**	16.71*	25.37
	(11.11)	(23.65)	(35.63)	(5.96)	(9.35)	(16.71)
principal characteristics	Yes	Yes	Yes	Yes	Yes	Yes
school characteristics	Yes	Yes	Yes	Yes	Yes	Yes
student characteristics	Yes	Yes	Yes	Yes	Yes	Yes
area characteristics	Yes	Yes	Yes	Yes	Yes	Yes
regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27775	27775	27775	27775	27775	27775
First-stage F statistic	4.29	1.59	0.77	6.09	5.44	3.28
AR (95%)	[3.93, 62.85]	[-63.20,-35.10] U [6.11, 122.24]	[-108.91,-18.64] U [6.75, 170.38]	[2.62, 35.17]	[3.01, 53.38]	[4.86, 90.85]
No. clusters	249	249	249	249	249	249

Notes: Standard errors clustered at school level in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The instrumental variable for managerial practices is a dummy=1 if the principal has experience in post less or equal to 4 years (or dummy post-2006 reform). *AR (95%)* reports the Anderson-Rubin 95% confidence intervals (on managerial practices) robust to weak instruments. More on this in Section 8. For all these specifications, additional controls are identical to those included in Table 3, results not reported but available upon request. See also Notes to Table 3 for additional details regarding control variables.

Table 6: Sorting, heterogeneity and cheating

Dependent Variable Test results (Math)	(1) Management	(2) Leadership	(3) Operations	(4) Monitoring	(5) Targets	(6) People
Panel A: Sorting of students						
Isolated schools						
managerial practices	4.97*	7.37***	1.73	3.29*	3.07	4.75*
	(2.623)	(2.235)	(2.128)	(1.708)	(2.400)	(2.501)
Observations	5,954	5,954	5,954	5,954	5,954	5,954
Other schools in competition						
managerial practices	2.67*	2.69**	1.40	2.72**	1.49	1.17
	(1.416)	(1.175)	(0.954)	(1.093)	(1.343)	(1.604)
Observations	21,821	21,821	21,821	21,821	21,821	21,821
Panel B: Effect heterogeneity						
Low socio-economic background						
managerial practices	4.17**	3.62**	1.98*	4.04***	1.67	3.70
	(1.938)	(1.704)	(1.133)	(1.314)	(1.920)	(2.374)
Observations	14,099	14,099	14,099	14,099	14,099	14,099
High socio-economic background						
managerial practices	1.52	2.31*	0.68	1.71	1.14	-0.72
	(1.649)	(1.261)	(1.278)	(1.093)	(1.608)	(1.917)
Observations	13,676	13,676	13,676	13,676	13,676	13,676
Panel C: Cheating						
managerial practices	2.29*	2.70**	1.13	2.40**	1.15	0.84
	(1.356)	(1.167)	(0.890)	(0.999)	(1.272)	(1.480)
managerial X inspector	-0.60	0.22	0.89	-0.48	-1.67	-1.50
	(1.800)	(1.863)	(1.087)	(1.500)	(1.604)	(2.096)
inspector	1.21	-0.59	-2.06	1.09	3.08	2.67
	(3.709)	(4.303)	(2.636)	(3.394)	(3.014)	(3.712)
Observations	27,775	27,775	27,775	27,775	27,775	27,775

Notes: Standard errors clustered at school level in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Panel A: results on separate regressions for the samples of students attending schools that do and do not have other schools in competition within 30 minutes driving distance. Panel B: results on separate regressions for the samples of low socio-economic background students (average school level *escs* index below zero) and high socio-economic background students (*escs* above zero). Panel C: nspector is a dummy equal to one if there is an external inspector randomly assigned by Invalsi to invigilate students during the test and help to compute results and documentation relative to the test. For all these specifications, the set of additional controls is identical to that included in Table 3, results not reported but available upon request. See also Notes to Table 3 for additional details regarding control variables.

Table 7: Interaction with tenure

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Test results (Math)	Management	Leadership	Operations	Monitoring	Targets	People
managerial practices	4.41**	4.87***	2.77**	3.89***	2.77	2.00
	(1.83)	(1.56)	(1.25)	(1.32)	(1.71)	(2.05)
managerial X tenure	-0.36	-0.32*	-0.27	-0.26*	-0.31	-0.19
	(0.22)	(0.17)	(0.18)	(0.14)	(0.20)	(0.22)
tenure in school	0.69	0.70*	0.58	0.54*	0.52	0.32
	(0.42)	(0.37)	(0.37)	(0.31)	(0.33)	(0.38)
principal characteristics	Yes	Yes	Yes	Yes	Yes	Yes
school characteristics	Yes	Yes	Yes	Yes	Yes	Yes
student characteristics	Yes	Yes	Yes	Yes	Yes	Yes
area characteristics	Yes	Yes	Yes	Yes	Yes	Yes
regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27775	27775	27775	27775	27775	27775
R-squared	0.252	0.255	0.251	0.255	0.249	0.248

Notes: Standard errors clustered at school level in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Tenure is measured in years.The set of additional controls is identical to that included in Table 3: these results are not reported but available upon request.

Table 8: Different sample: excluding pre-2000 reform principals

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Test results (Math)	Management	Leadership	Operations	Monitoring	Targets	People
managerial practices	6.11***	5.24***	2.15**	4.80***	5.88***	5.23***
	(1.73)	(1.65)	(1.04)	(1.18)	(1.68)	(1.94)
principal characteristics	Yes	Yes	Yes	Yes	Yes	Yes
school characteristics	Yes	Yes	Yes	Yes	Yes	Yes
student characteristics	Yes	Yes	Yes	Yes	Yes	Yes
area characteristics	Yes	Yes	Yes	Yes	Yes	Yes
regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14060	14060	14060	14060	14060	14060
R-squared	0.318	0.315	0.305	0.322	0.317	0.311

Notes: Standard errors clustered at school level in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The sample is composed of principals with 9 (or less) years of experience in post (those appointed post-2000 reform). The set of additional controls is identical to that included in Table 3: these results are not reported but available upon request.

Table 9: Different institutional constraints

Dependent variable Test results (Math)	(1) Low constraints	(2) Medium constraints	(3) High constraints
managerial practices	2.48**	1.41	0.59
	(1.16)	(1.08)	(1.56)
principal characteristics	Yes	Yes	Yes
school characteristics	Yes	Yes	Yes
student characteristics	Yes	Yes	Yes
area characteristics	Yes	Yes	Yes
regional dummies	Yes	Yes	Yes
Observations	27775	27775	27775
R-squared	0.251	0.248	0.247

Notes: Standard errors clustered at school level in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. A full discussion on how each WMS survey question is classified based on how much SPs activity is likely to be constrained by the institutional setting can be found in the online Appendix B and Table A9 for details. The set of additional controls is identical to that included in Table 3: results not reported but available upon request.

Table 10: Quantile regressions

	(1) Management	(2) Leadership	(3) Operations	(4) Monitoring	(5) Targets	(6) People
quantile .25	1.3325	2.2495*	0.8607	1.6433*	0.2573	-0.2345
	(1.200)	(1.186)	(0.861)	(0.902)	(1.102)	(1.231)
quantile .50	2.4152*	3.1826***	1.3880	2.3875**	0.9405	0.5750
	(1.450)	(1.171)	(1.017)	(1.015)	(1.417)	(1.605)
quantile .75	2.7904*	2.8813**	1.4707	2.6978**	1.1570	1.4886
	(1.554)	(1.286)	(1.012)	(1.118)	(1.488)	(1.766)
principal characteristics	Yes	Yes	Yes	Yes	Yes	Yes
school characteristics	Yes	Yes	Yes	Yes	Yes	Yes
student characteristics	Yes	Yes	Yes	Yes	Yes	Yes
area characteristics	Yes	Yes	Yes	Yes	Yes	Yes
regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,775	27,775	27,775	27,775	27,775	27,775

Notes: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The set of additional controls is identical to that included in Table 3: results not reported but available upon request. See Table 2 for additional details regarding control variables.

Table 11: Baseline model. PISA Data, Italy

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Test results (Math)	LDRSHP	Curricula	Resources	Leadership	Operations	Monitoring
managerial practices	-0.45	-0.48	-4.35	4.70	-4.30	1.31
	(1.841)	(2.056)	(4.304)	(3.645)	(4.269)	(3.484)
competition (two or more schools)	-3.32	-3.30	-2.27	-3.85	-3.37	-3.52
	(6.648)	(6.541)	(4.053)	(6.627)	(6.686)	(6.638)
competition (one school)	-11.77	-11.68	-6.12	-11.79	-11.88	-11.87
	(8.263)	(8.118)	(5.027)	(8.228)	(8.319)	(8.250)
school size	9.22***	9.19***	10.59***	9.17***	9.20***	9.27***
	(3.397)	(3.427)	(2.668)	(3.386)	(3.380)	(3.407)
private school	-40.71***	-40.90***	-15.22	-40.77***	-40.56***	-40.89***
	(13.022)	(13.068)	(9.730)	(13.114)	(12.930)	(13.071)
share part-time teachers	6.84	7.05	13.21	6.96	6.78	7.21
	(19.335)	(19.361)	(11.501)	(19.422)	(19.250)	(19.419)
socio_economic background	7.24***	7.24***	8.33***	7.22***	7.25***	7.23***
	(0.913)	(0.914)	(0.723)	(0.912)	(0.912)	(0.914)
immigrant student, 1st gen.	-39.96***	-39.96***	-36.48***	-39.97***	-40.03***	-39.91***
	(3.324)	(3.318)	(2.632)	(3.325)	(3.320)	(3.330)
immigrant student, 2nd gen.	-23.73***	-23.78***	-15.75***	-23.98***	-23.66***	-23.84***
	(5.862)	(5.884)	(4.106)	(5.869)	(5.852)	(5.872)
gender student (female=1)	-31.94***	-31.92***	-31.43***	-31.87***	-32.00***	-31.88***
	(1.868)	(1.881)	(1.426)	(1.873)	(1.868)	(1.865)
student behind	-42.79***	-42.80***	-48.44***	-42.80***	-42.80***	-42.78***
	(2.073)	(2.070)	(1.549)	(2.068)	(2.072)	(2.070)
dummy village or rural area	12.29	12.31	11.72	11.48	13.10	11.71
	(10.989)	(11.039)	(10.432)	(10.919)	(11.049)	(10.859)
dummy small town	5.51	5.39	4.14	5.24	5.38	5.13
	(9.358)	(9.295)	(8.948)	(9.310)	(9.304)	(9.375)
dummy town	2.53	2.38	1.28	2.29	2.73	2.13
	(7.831)	(7.763)	(8.284)	(7.748)	(7.785)	(7.840)
dummy city	-6.01	-6.03	-2.87	-5.82	-6.20	-6.12
	(8.672)	(8.654)	(8.814)	(8.631)	(8.692)	(8.676)
Constant	384.95***	345.05***	447.48***	367.92***	400.05***	380.43***
	(22.094)	(21.071)	(19.937)	(26.087)	(26.679)	(25.200)
Observations	24,882	24,915	24,915	24,882	24,882	24,882
R-squared	0.434	0.434	0.405	0.434	0.434	0.434

Notes: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variable is standardized test score in mathematics. Managerial practices are as follows: in cols.1-3 we use the variables LDRSHP, RESPCUR, RESPRES as in the OECD PISA dataset, in col.4 we use the average of items (a), (k) and (l) of Q26 to construct the Leadership indicator, in col. 5 items (b), (c), (e), (h), (j) and (m) for Operations indicator, in col.6 (d), (f), (g) and (i) for Monitoring indicator. Competition is the number of schools in competition (no competition is the omitted dummy). School size is the natural logarithm of the number of students in school. Private is a dummy equal to one for private schools. Immigrant student 1st generation are students born abroad of foreign-born parents, immigrant student 2nd generation are native-born children of foreign-born parents. Student behind is a dummy equal to 1 if the student has repeated a grade. Dummies for urbanization are defined in Table (large city is the omitted category). Regional and school type dummies included.