PROGRAMMA EDUCATION
FGA WORKING PAPER

N. **49** (07/2013)

Managerial Practices and Students' Performance

Adriana Di Liberto*, Fabiano Schivardi,
Giovanni Sulis***

***** Università di Cagliari**
****Università LUISS**

Fondazione
Giovanni Agnelli

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

Managerial Practices and Students' Performance*

Adriana Di Liberto
University of Cagliari
and
CRENoS

Fabiano Schivardi
LUISS, EIEF and CEPR

Giovanni Sulis
University of Cagliari
and
CRENoS

July 18, 2013

FIRST DRAFT

Abstract

We study the effects of managerial practices in schools on students' outcomes. We measure managerial practices using the World Management Survey, a methodology that enables 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 students' outcomes in a math standardized test. We find that managerial practices are positively related to students' 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 self selection of best principals into best schools 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 students' 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, M1, O32.

*Contacts: Adriana Di Liberto, Department of Economics and Business, University of Cagliari, Viale S. Ignazio da Laconi 78, Cagliari, Italy, diliberto@unica.it; Fabiano Schivardi, Department of Economics and Finance, LUISS University, Viale Romania 32, 00197 Roma, fschivardi@luiss.it; Giovanni Sulis, Department of Economics and Business, University of Cagliari, Viale S. Ignazio da Laconi 78, Cagliari, Italy, gsulis@unica.it. We thank Daniele Checchi, Gianfranco De Simone, Andrea Gavosto, Erik Hanushek, Victor Lavy, Angelo Paletta, Paolo Sestito, Marco Sideri and John Van Reenen for useful comments and discussions. We thank Renata Lemos and Raffaella Sadun for their support in the World Management Survey data collection process. We also thank INVALSI and in particular Patrizia Falzetti for providing the data on students outcomes. We are solely responsible for all the remaining errors. Sulis acknowledges hospitality at the Center for Labor Economics at Berkeley University. Financial support from the Regione Sardegna, Legge 7/2007 - grant CUP F71J09000300002, and from the Fondazione Giovanni Agnelli is gratefully acknowledged.

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 students' 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 students' achievement only in well-developed school systems (Hanushek, Link, and Woessmann, 2011). In fact, only when schools are well managed students can benefits from decentralization, while giving autonomy to badly run institutions can indeed worsen students' outcomes.

Indeed, there is a growing attention regarding the role of school principals (SPs in what follows) as managers in charge or running the school. However, we still know relatively little on this issue, because assessing the role of managerial practices on students' 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's outcome, only few recent studies have attempted to quantify the role played by the SPs on students' outcomes. This paper addresses this question. We collect data on school managerial practices through extensive phone interviews of around 400 Italian SPs of upper secondary schools. The interviews are based on the World Management Survey (Bloom and Van Reenen, 2010b; Bloom et al., 2012) 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. 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, 2007, 2010a).

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. 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 SPs 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 students' results in a standardized math test administrated by the INVALSI, the Italian institute in charge of evaluating schools' performance. The Italian case is an interesting one to study the effects of SPs' managerial practices on students' 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 details 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 characteristics we can control for. Finally, we use two reforms of the Italian school system to construct an IV regression.

Our baseline model is an OLS regression of students' 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 students' characteristics. Data on students' 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 2.05 and s.d. of .5) would improve the students' average test score results by 4.6%, approximately the distance of Italian students from the OECD average in the standardized OECD-PISA tests. While a unit increase in managerial practices is clearly substantial, it

is also approximately the distance from the average value in Italy and 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 2000, schools have enjoyed greater organizational autonomy that has increased the emphasis on SPs' managerial duties. Moreover, in 2006 a new national competition was introduced that also explicitly assesses managerial skills. Both events 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 construct two dummies for those who became SP after 1999 and 2005. To allow for potential local differences in the implementation of the reform, we interact them with macro-area dummies and use them as exogenous shifter of managerial practices. The IV results confirm the OLS ones, with the effect becoming substantially larger. We also experiment with different specifications, outcome variables, and subsamples. All the exercises confirm that management quality is an important input of our estimated education production function and that the OLS results are a lower bound of the effect of managerial practices on students' outcome. We finally investigate if managerial practices are more relevant for the achievements of some type of students, such as those lagging behind. Unlike other studies, we find that the increase in performance is fairly similar across the distribution of students (Branch, Hanushek, and Rivkin, 2012). Consistently, effects of managerial practices on the dispersion in students' outcomes is not statistically significant (although the point estimate is always negative).

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 Section 4 compares the survey results for the six countries. We describe the students' data and the additional controls in Section 5 and the identification and the empirical design in the following one. Section 7 discusses the main results and Section 8 the extensions and robustness checks. We conclude in Section 9.

2 Literature review

The role of the SPs on students' learning is increasingly identified as crucial. The literature recognizes an important but often indirect role of the SPs through their influence

on teachers, and there is no consensus about what attributes a SP should have in order to successfully lead a school. Three main channels are identified as the most important affecting students outcomes and they relate to specific managerial practices. The first one, also called by education scholars *instructional leadership*, focuses on the role of SPs in promoting the introduction of organizational innovations enabling teachers to work more effectively. In general, these include all activities that facilitate and improve the quality of teaching and learning.[1] 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."[2] Finally, the literature stresses the importance of the processes of selection of teachers through the recruitment of new staff and dismissal of the existing one which can be decided in part or in whole by the SP depending on the institutional context, since collective bargaining agreements may prevent principals from engaging in firing/hiring low/high performing teachers.[3] The change in the quality of teachers is considered as one of the primary channels through which SPs may affect educational outcomes.[4]

Most existing analysis 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.[5] 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.[6] Most quantitative studies use the value added approach, a methodology already introduced to estimate the effect of individual teachers on student performance but employed also outside the education framework to identify the role of CEOs in firms productivity.[7]

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 achieve-

[1] See Robinson, Lloyd, and Rowe (2008) and Grissom and Loeb (2011).

[2] See Robinson, Lloyd, and Rowe (2008), p. 639.

[3] 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).

[4] Grissom, Loeb, and Nakashima (2013) use data from the Miami districts involuntary transfer policy finding that principals in the lowest-performing schools tend to transfer/dismiss low-performing teachers for transfer who would have been unlikely to leave on their own. Moreover, sending schools were able to hire teachers who achieved higher student's test scores.

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

[6] 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.

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

ment using the semi-parametric approach provided by Bertrand and Schoar (2003) in their study of corporate management styles.[8] They find significant variation in principal quality and identify a large effect, similar to that found for teachers, of principal leadership on student outcomes.[9] 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 students test results have also been found in Coelli and Green (2012) for Canada and Böhlmark, Grönqvist, and Vlachos (2012) for Sweden. Together with students' 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 students' test results.[10]

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 for equity. 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 students' performance targets.[11] 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.[12] 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 produce exactly the opposite result.

[8]They calculate different measures of SPs efficiency. A first estimate is obtained by introducing effects-by-school principal in a regression model that uses as dependent variable the results on cognitive tests (in both math and reading for students from three to eight grade), while a second specification is obtained by adding to the previous regression model school fixed effects.

[9]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.

[10]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.

[11]Performance targets, called Adequate Yearly Progress, are set dividing students into 9 demographic subgroups. They require that students in each subgroup reach a particular threshold for reading and math scores. If only one subgroup fails to make this target, the entire school is declared failing. Data shows that for disadvantaged schools it was difficult to improve test scores among every low-performing demographic subgroup.

[12]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.

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 students' academic achievements.[13]

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 SP's assistants to the same questionnaire as well as alternative indicators of teachers and parents satisfaction with school quality.[14]

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.[15] Second, value added measures do not control for the possibility that SPs quality change over time with tenure and experience.[16] Third, they produce an overall measure of the SPs impact but they do not tell much about what SPs actually do to influence student

[13]On this see also Cullen and Mazzeo (2008) and Brewer (1993).

[14]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 students academic performance.

[15]In this case, such problems conflate the true SP effect with other factors for test results change. However, Li argues that compared to teachers value-added measures "...these concerns are less of a problem in the context of studying principals. While principals have substantial knowledge about the test scores and other characteristics of students in their own school and may use this information in assigning teachers to classrooms, they have less information about the test score gains of students at other schools and are thus less likely to use this information in their own mobility decisions." Li (2012), p.17.

[16]On this see Li (2012).

learning.[17] As seen above, the survey approach overcome the latter criticism. However, the 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.[18]

3 The World Management Survey

As highlighted in the previous section, the main difficulty encountered when analyzing the effect of SPs on school outcomes is to provide a reliable quantitative measure of SPs' abilities in terms of leadership capacity and organizational skills. While there is an established literature that suggests that such components are important determinants of 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. As a matter of fact, previous studies dealing with the role of SPs and based on the "survey approach" suffer from severe limitations, mostly related to mis-measurement of managerial abilities and self-assessment bias.

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 in order to obtain such measures. 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, managers' responses 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, at 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 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

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

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

evaluated using a scoring grid (Bloom and Van Reenen, 2011, 2010a; Bloom et al., 2012). 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 production units, as schools or firms, and focuses on the solutions adopted by principals/managers to solve specific problems.[19]

The questionnaire comprises five sections that consider different key areas of management practices.[20] The first 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 second 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 third section is Targets (five questions) and has the objective to assess the managerial capacity of SPs to identify quantitative and qualitative targets, their interconnection and their temporal cascade. The fourth section 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. Finally, the fifth dimension is Leadership (three questions) and measures the leadership capacity of the SP jointly with a clear definition of roles and responsibilities within the school. [21]

The remaining part of the questionnaire is dedicated to obtain data on the main principal and school characteristics. We collected information on demographic characteristics of the SP such as gender, age, tenure in school, experience in post, marital status, teaching field of specialization and place of birth. We also obtained data on the number of students in the school, number of teachers and administrative staff, number of schools in competition, type

[19] Although the questionnaire is very similar across different sectors with the focus on key management areas as monitoring, targets and incentives, there are some differences in terms of specific sections and questions included for each different sector. We will discuss such issues in more detail in next sections.

[20] 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.

[21] The WMS international survey on education is based on 23 questions. We include two more question in the Italian version of the questionnaire. The first measures the ability of the SP in raising additional funds with respect to those received by the national Government. The second is related to ICT adoption in schools, as the use of a website to provide information regarding the school or the use of the email for communication with teachers and admin staff. However, to ease comparability with cross-country data, we exclude the latter questions from the calculation of our main index of managerial practices.

of school, religious orientation and ownership (private versus public), presence of possible selection criteria for students, type of administrative procedures needed for hiring teachers and to increase the number of students.

We collected our data on managerial practices during the period from February to May 2011, with a team of five analysts and two managers that were adequately trained and monitored by the international WMS team. During the training period the team was first prepared to the principles and techniques of the WMS project, then a series of pilot interviews were conducted to familiarize with interview techniques, the questionnaire and the scoring grid. Most importantly, a very large fraction (about half) of the interviews subsequently conducted by the analysts were double scored by the managers or by another analyst. In order to reduce difference in scoring across analysts, some of the interviews were jointly scored by the whole team. The sample constitutes a representative sample of the population of Italian secondary schools.

4 Cross country comparisons

In addition to Italy, the same survey has been run in Canada, Germany, Sweden, the UK and the US. In this section we compare the measures of managerial practices across such countries. Given the standardized data collection process, the indicators are in fact fully comparable across countries. Given that the econometric analysis of the effects of managerial practices on students' outcomes will be performed for Italy, the only country for which we have students' outcomes data, in what follows we benchmark the discussion on this country.

Cross country differences in the quality of managerial practices can stem from two main sources. First, 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 USA 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. In fact, the survey by design measures actual practices rather than the intrinsic ability of the SP. 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 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 overall average of 2.01, the managerial skills of Italian school leaders are significantly lower than those of other countries.[22] 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, 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 a few SPs with values less than 2, especially in the case of Sweden and Great Britain. Table 10 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 biding 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

[22]Note that the average value of the overall Management index reported in Figure 1 slightly differs from the one reported in Table 1, as the former is calculated on the original sample of 341 schools, while the latter is obtained using the schools matched with students' outcomes.

performance, we should find that this is mostly concentrated in the areas in which such constraints are more binding.

To save on space, here we only report the main results, referring the interested reader to the Appendix and Table 11 for all the details. 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. 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 four 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 a 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 Students' performance measures and additional individual-level data

Our second source of 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 students 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.[23]

The 2010-11 was the first school year that these evaluation tests were performed by upper secondary school students.[24] 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.[25] 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's background and family characteristics.[26] 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.[27] The latter is calculated based on the parents' 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 students 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.[28]

Our WMS survey dataset on principals and schools is therefore matched to the INVALSI dataset through an anonymous school identifier. Table 1 sums up the major characteristics of these additional variables for our overall sample. We find that our sample is not significantly different from the INVALSI 10th grade students census one (data in brackets)

[23]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 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.

[24]Even if the Italian high school system is based on three main differentiated curricula (Vocational, Technical and Lyceum) both the Language and Mathematics tests and their administration procedures were identical for all schools and students. In fact, 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.

[25]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.

[26]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 students general information part compiled from school administrative staff.

[27]More precisely, 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 if the student is older than "regular" students, that is, if, at the end of 2011, he/she is older than 16 years old.

[28]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.

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 48.59 (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 and the percentage of retained students is 22% (22%).[29] Thus, only the presence of first generation immigrant students is somewhat oversampled. Finally, both the mean value and the standard deviation of the ESCS index confirm that, even in terms of students 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).[30] 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 1 show that most upper secondary schools are located in densely populated areas, with only 11% located in rural areas. As explained in more detail in the following section, we will also focus on a specific sample that will be used in our regression analysis. In this case, as reported in Table 2 for each variable both sample averages and standard deviations are almost identical to those reported in Table 1.

6 Empirical framework and identification

We study the effects of managerial practices on students' 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 students' 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, it could be that more capable SPs self select into schools with better students. In this case, interpreting β in a causal sense would clearly be incorrect. We take particular care in addressing the endogeneity concerns.

[29] Note that this retained students dummy includes also non-native students that are allocated to a lower grade on the basis of their language skills and not on the basis of a simple age rule.

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

First, our data include an extremely rich set of students, 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 students outcomes to identify the effects of managerial capabilities on students' 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 students 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 follows. In practice, RSA try to accommodate SPs requests, but have to fill in the positions for the schools that where 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 students' outcomes in standardized tests. Formal self-evaluation procedures are just starting to be introduced in the Italian system and students achievements 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 the residence.[31]

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 requested.

Although we argue that the selection mechanism is unlikely to give rise to serious endogeneity constraints, we cannot completely rule out this possibility. We will also perform a set of IV regressions. We focus on two institutional changes that should determine exogenous shifts in SPs managerial abilities. First, since 2000, schools have enjoyed greater organizational autonomy in compliance with art. 21 of the Law 59/1997. Before that, the main role of SPs was that of "*instructional leader*" rather than of manager. Moreover, 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. Following the 2000 reform, this constraint has been eliminated. Moreover, SPs became not only instructional leaders but also school managers with full responsibility over the school budget. Since 2000, therefore, managerial skills and specific training in human resource management were required to become SP. This implies that the SPs appointed from 2000 onward might have had better managerial abilities than those selected before the reform. Second, in 2006 a new national competition was introduced to become SP. The competition was based on an exam, that included testing specifically for managerial skills. Although the competition was established at the national level, it was implemented at the regional level. We derive two instruments from these institutional changes: a dummy equal to 1 for SP appointed after 1999 and another dummy equal to 1 for SP appointed after 2005. To allow for potential local differences in the implementation of the reforms, we interact these dummies with macro-area dummies. The rank condition requires that such reform had an impact on the managerial capabilities of SPs selected after the reforms were implemented. The exclusion

[31]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 teachers' mobility can be found in Barbieri, Rossetti, and Sestito (2011).

restriction requires that being selected after the reform does not systematically correlate with the students' 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.

7 Main results

To investigate the relationship between SP's managerial skills and students 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), students 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.[32] Robust standard errors, clustered at principal/school level, are reported in parenthesis. We start in column 1 with our main SP's 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 students 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 students 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

[32] We report specific descriptive statistics for the sample used in regressions in Table 2. Note that averages and standard deviations are very similar to those reported in Table 1 for the overall sample, that also includes SPs that have not chosen the school.

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, students's test scores would increase by around 4.6%.

In columns 2 to 6 we investigate the effects of the various dimensions of managerial skills. Column 2 focuses on Leadership, an area of management often identified by the education literature as an important channel through which SPs may influence teachers behaviour and, thus, students outcomes. In general, leadership skills, also called *transformational leadership*, 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 (Grissom and Loeb, 2011; Robinson, Lloyd, and Rowe, 2008). 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. However, they are also aimed to capture all activities that SPs perform to informally stimulate/incentivize teachers work.[33] 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 students outcome is monitoring of school processes. The five questions on Monitoring included in our survey focus on the tracking of school performance, reviewing performance with teachers and staff, and consequence management (e.g., making sure that, if a problem is identified, the appropriate sanctions and rewards are in place). Column 4 suggests that 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. 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 (Robinson, Lloyd, and Rowe, 2008). Column 5 includes results on Targets. This section of the survey examines mainly the type of targets set by the school in terms of students outcomes, their realism and whether

[33]Within the WMS project, Leadership is investigated only in the Educational (3 questions) and the Health care (1 questions) sectors. As stressed by Bloom et al. (2012), leadership is an important part of any business success but they consider this skill harder to quantify than other types of management and they do not examine its role in their analysis on manufacturing and retail sectors. Unlike the former two sectors, the latter are mainly privately owned organizations, where formal incentives promotions can be fully implemented.

they are given consistently throughout the school organization. As also stressed by Branch, Hanushek, and Rivkin (2012) public sector CEO 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. 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). Finally, column 6 includes People, that is, the incentives section that includes 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.

7.2 Other controls

We now turn our attention to examining the results obtained 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 in regards to the relationship between school performance and principal characteristics.[34] Principal's 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 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 students educational outcomes: both principal's age and gender show a positive and non significant coefficient, while the coefficient on tenure in school is statistically insignificant.

Results found in Table 3 for the remaining indicators are largely consistent with the

[34] In particular, studies that use the value added approach often find that the vast majority of variation in educator (both principals and teachers) quality cannot be explained by their observed characteristics. On this see Li (2012).

literature. One important exception is the private schools dummy that shows a negative and significant coefficient, implying that private school students' performance on standardized tests is worse than that of public Italian school students. This should come as no surprise since, unlike most industrialized countries, private schooling in Italy is associated with poorer rather than better outcomes. As suggested by Brunello and Rocco (2008), one of the main reasons is that the large part of Italian private schools focus more on the recovery of less able students than on across the board high quality education. That is, Italian families often choose private, including religious, schools, to obtain degrees from less demanding private institutions.[35]

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 students 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.[36] 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 between these two curricula there is the curriculum offered by Technical schools. Moreover, only Vocational schools can last for 3 rather than 5 years, even if graduates from all three school types, after five years, may continue to tertiary education. 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.

Among additional type of school controls, we introduce another dummy variable, "Istituto Superiore", that identifies if a school offers different types of curricula.[37] 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 students outcomes.[38] Further, we also find that both school size and school competition do not significantly affect students' test results.[39] We finally include a measure of teacher's

[35]See for example OECD (2012). 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.

[36]Due to this mechanism, in addition to their cognitive abilities (test scores, grade repetition, previous marks), 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).

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

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

[39]Schools (and teachers) facing a direct competition may exert more influence and put more efforts to

turnover. Italian principals have almost no control over new hires or teachers transfers thus we only have voluntary transitions. This variable too is not significant.

Third, we presents estimates of students' characteristics. The results obtained are very much consistent with the literature. In general, cross country analysis shows that mathematics gender gaps results in favor of boys are rather unstable and rapidly evolving, with countries with a more gender-equal culture often showing a nil or even reversed gender gap.[40] Our analysis confirms past results on Italian data with the presence of a significant "gender gap" in terms of boys outperforming girls in mathematics.

As expected, the index of socioeconomic background is positively strongly associated with student achievement. Although it does not include any specific data on family income, this variable is created on the basis of the occupational and educational level of the students parents, home educational and cultural resources. Further, estimates confirm that children of immigrants face important gaps in math, with first-generation immigrants representing the most disadvantaged group. Note that significant immigration flows in Italy only started in the Nineties but, from then on, have risen sharply over the last decade and almost 8% of students in our sample are first or second generation immigrants. We also include a dummy that captures if a student is repeating the grade or if she/he is a retained students and find, as expected, a negative and significant coefficient.

In terms of catchment area characteristics, we construct three dummies for whether a school is located in densely populated, intermediate density or sparsely populated areas. None of this measures significantly affect math test scores. Conversely, the log of per capita bank deposit, a proxy for the wealth level of the school catchment area, is always positive and significant. We also include regional dummies. In fact, previous studies show that geographical location is an important determinants of Italian students test scores, with students in the North-East area usually outperforming the others, and those from the South being substantially behind (Cipollone, Montanaro, and Sestito, 2010; Bratti, Checchi, and Filippin, 2007). Results (available upon request) confirm this pattern. Moreover, regional dummy are important also because they enable us to control for possible local differences in principals allocation rules. In fact, as said above, since SPs' 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.

improve students outcomes, while Böhlmark, Grönqvist, and Vlachos (2012) suggest that principals in smaller schools have a larger influence on students' academic results.

[40] Guiso et al. (2008). For example, PISA results report some male advantage in all rounds although not in all countries. See Eurydice (2010).

8 Robustness and Extensions

In this section we perform a set of our robustness checks of the basic results discussed above. First concerns are selection issues. If SPs' managerial skills are not orthogonal to the unobserved component of students' 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 two institutional changes that occurred in the past that modified the legal requirements for Italian school principal selection as described above. As in Table 3, columns from 1 to 6 show the second stage results when we alternatively include our six different proxies of managerial practices. In Table 4 we also report the p-value of the Hansen J test of over-identifying restrictions to test the validity of our excluded instruments and check for the relevance of our instruments.

The point estimates of the IV regression reported in column 1 of Table 4 suggest that, if anything, the OLS underestimate the importance of managerial skills. The coefficient on Management implies that a one point increase in managerial scores is associated with a 8.2 points increase in student test results. This coefficient appears to be precisely estimated and it is statistically significant at the five percent confidence level. When we focus on the different areas of management in columns 2 to 6 we find that with the exception of Targets, all IV estimates are 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, while the outcome of the Hansen J-test for overidentifying restriction tells us that altogether our instruments are valid, when we test for the strength of the instruments, the Kleibergen-Paap F statistic casts concern over their strength. 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 column 1 we find a significantly lower value, 2.8, and in the remaining models it is always below 4. We take this results as corroborating the OLS ones of a positive effect of managerial practices on students' outcomes, but, given the weak instrument problem, acknowledge that the point estimate have to be taken with a grain of salt.

In Table 5 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 5, are mixed. First, the effect of managerial skills increases and become 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.[41] 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.[42]

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 1999. 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 6, 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 students' outcomes. The coefficient for the overall managerial index reported in column 1 of Table 6 indicates that an increase in the index by one point increases student scores by more than 6 points with an increase of about 12%, an effect that is almost 3 times as large as the one found in Table 3 on the sample

[41] 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.

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

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).

Interestingly, inspection of Table 6 reveals some differences for our set of controls with respect to results reported in Table 3. While results in terms of direction and statistical significance for most school, students and regional area characteristics are in line with those found in our baseline regression, we find differences on principals characteristics coefficients. In fact, for this restricted sample of SPs we do find a positive effect of age and school tenure on students' outcomes, while the gender dummy turns out to be negative and sometimes statistically significant at 10 percent levels. Although effects are marginally significant and do not always hold across columns, such results suggest that the legislative change we exploited for our IV strategy had a significant impact on the characteristics of the SPs hired after the year 2000.

In section 4, we have stressed the importance of institutional constraints as a factor that may help 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 Table 11 and the Appendix for details about the classification of each question). Results reported in Table 7 indicate a consistent pattern: managerial practices have a positive and statistically significant effect (about 2.5 points) on students' 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.[43]

The identification strategy of the effect of managerial practices adopted in our paper exploits the variability of students' outcomes across students, but the independent variable of interest only varies across schools. To account for this, we have always clustered the standard errors at the school level. As a further check, we supplement our main findings with regressions based on school averages in test scores as the dependent variable, so that we have one observation by school. Other regressors are also obtained as school level averages. We report our results in the first row of Table 8 for the overall managerial index and separately for all different sections of managerial practices.[44] Overall, the results confirm our previous findings: while we do find a positive but not statistically significant effect for our main Management index, and for Operations and Targets, the positive effect for Leadership and Monitoring is confirmed to be statistically significant, with an increase of student test scores that is equal to about 2 points for an increase of 1 point in the respective index of managerial ability.

One important issue is what kind of students benefit more from good managerial practices: is it good students or the more disadvantaged ones? To investigate this issue, we replicate our previous analysis using as dependent variable the level of student test scores at the 10th, 50th and 90th percentile. In fact, managerial practices can have different effects on the shape of the distribution, depending on the particular targets SPs want to reach. Results, which are reported in rows 2 to 4 of Table 8, indicate that good managerial practices in terms of Leadership and Monitoring do have a positive and statistically significant effect on student outcomes, and such effect is quite similar in size across different parts of the distribution. For example, good managerial practices in Monitoring increase test scores similarly for both the bottom and upper part, respectively low achieving and elite students, of the test scores distribution (with an effect equal to 1.48 and 1.86 respectively). In other words, good managerial practices shift the distribution of student scores to the right without basically changing its shape. Consistently, we also find no statistically significant effect of managerial practices on different measures of dispersion of student outcomes, reported at the bottom of Table 8, such as the 90/10 percentile difference or the standard deviation of average scores between classes in the same school.

We conclude our regression analysis by considering the effect of managerial practices on

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

[44]The number of observations for these regressions drops to 221 schools, while regressions reported in Tables 3 and 4 reported 27775 observations for 249 schools. The difference is due to those schools for which test scores for only one class of students were available.

a different students' outcome. In Table 9, we 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. Such variable was used in our previous regression as an additional control for student characteristics.[45] Our set of regressors is identical to those reported in Table 3 and subsequent ones, hence includes standard controls for SP demographics, and different school, student and area characteristics. Even using an alternative student outcome variable results show a positive role for SPs managerial practices. In fact, we find 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. Again, we do find a larger and statistically significant effect for Leadership. While most of our controls turn out not to be statistically significant, there are a few results that are worth mentioning. As expected, we find that students in private schools have higher probability of lagging behind with respect to the age/grade ladder, confirming our claim above that private schools in the Italian system are mostly focused on recovering less able students (Brunello and Rocco, 2008). Similarly, first and second generations immigrants have higher probabilities of lagging behind, while better family background reduces such probability.

9 Conclusions

This study investigates if the quality of managerial practices in schools affects students' 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 a 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 students' outcomes in the mathematics standardized test. In particular, in

[45] We already stressed that while a large fraction of such students are those with grade retention, there is also a fraction of students that are older than their classmates for different reasons, such as immigration. Hence, in the interpretation of results one should keep this caveat in mind.

order to investigate the relationship between SPs managerial skills and students outcomes we estimate a standard education production function where students' test performance in mathematics is modeled as a function of SP managerial skills and of rich set of students, 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, students' test results would increase by about 4.6%. In other words, if Italy had the same managerial practices as the UK (the best performer in terms of management), it 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.

As a further check we also perform a set of IV regressions using two institutional reforms in SPs' recruitment and training processes as instruments in order to identify the effect of managerial practices. Although standard tests suggest that we cannot rule out weak instruments problems, our IV results confirm the OLS ones, with the effect becoming substantially larger, and statistically significant for all (but one) areas of management, i.e., operations, leadership, monitoring and people (but not targets). We also perform a number of further robustness checks. First, we replicate our analysis on different sub-samples, i.e. considering only the more recent cohort of SPs. Second, we consider how the effect of managerial practices varies depending on the importance of institutional constraints. Third, we analyze if managerial practices affect the distribution of test scores within the school. Finally we consider an alternative students' oucome 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 and that the OLS results are a lower bound of the effect of managerial practices on students outcome.

Overall, our results suggest that policies directed at improving students' cognitive achievements should take into account principals' selection and training in terms of managerial capabilities.

References

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." NBER Working Paper 17850.

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.

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.

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* .

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

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. SIPI.

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.

Grissom, Jason A., Susanna Loeb, and Nathaniel Nakashima. 2013. "Strategic Involuntary Teacher Transfers and Teacher Performance: Examining Equity and Efficiency." NBER Working Papers 19108.

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

Hanushek, Eric A., Susanne Link, and Ludger Woessmann. 2011. "Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA." NBER Working Paper 17591.

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.

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.

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.

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.

A Figures and Tables

Figure 1: Average Managerial Practices across Countries

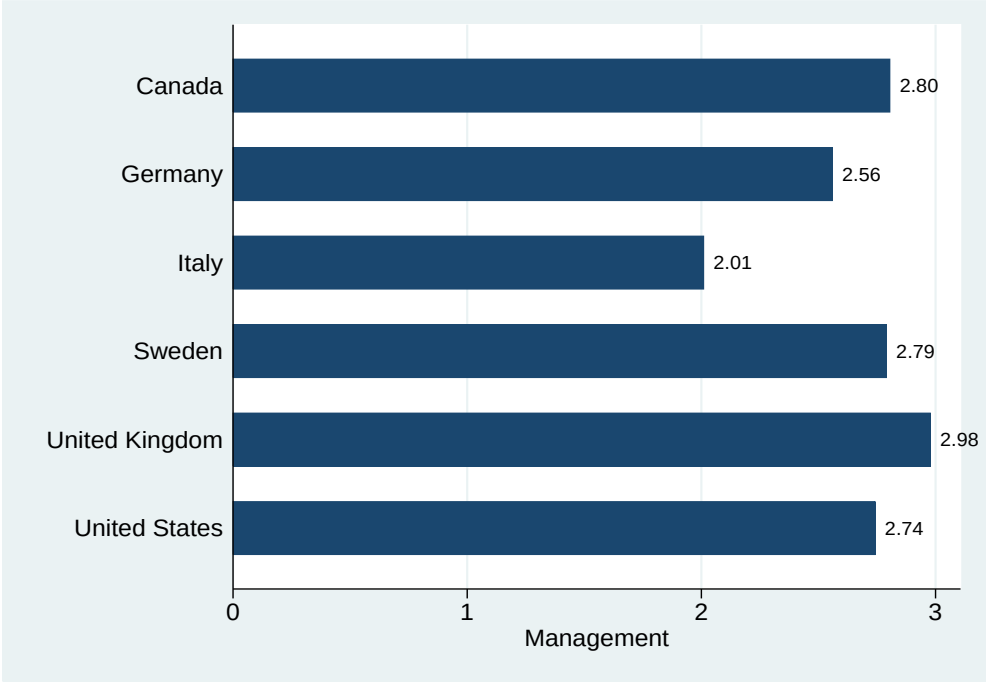

Figure 2: Distribution of Managerial Practices within Countries

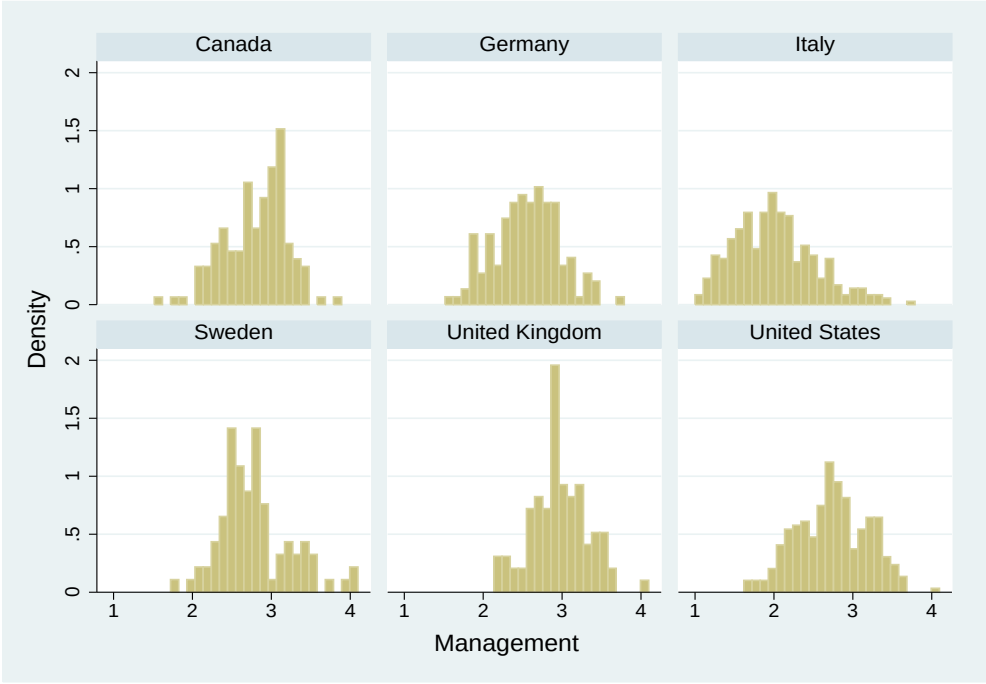

Table 1: Descriptive statistics. Overall sample

Variable	Mean	Std. Dev.	Min.	Max.	Obs.
test score math normalized	48.59	17.96	0	100	33727
management	2.04	0.5	1	3.74	37777
people	1.8	0.45	1	3.67	37777
targets	1.83	0.53	1	3.6	37777
operations	2.32	0.71	1	4.5	37777
leader	2.23	0.57	1	4	37777
monitoring	2.22	0.67	1	4.60	37777
principal's age	58.66	5.33	32	80	37632
gender of principal, female=1	0.38	0.48	0	1	37777
tenure in school	5.8	5.01	1	40	37595
competition	2.48	2.58	0	15	37598
ln of students	6.62	0.54	3.93	7.5	37777
dummy private school	0.05	0.21	0	1	37777
teacher turnover	4.79	5.78	0	50	37428
Lyceum	0.48	0.5	0	1	34244
Vocational	0.21	0.41	0	1	34244
Technical	0.31	0.46	0	1	34244
dummy "Istituto Superiore"	0.35	0.48	0	1	37777
immigrant student, 1st gen.	0.06	0.23	0	1	37362
immigrant student, 2nd gen.	0.02	0.16	0	1	37362
gender of student, female=1	0.5	0.5	0	1	37360
escs, socio-economic background	0.04	0.99	-3.66	2.07	33254
dummy=1 if student "retained"	0.22	0.42	0	1	37349
dummy rural area	0.11	0.31	0	1	37777
dummy non rural area	0.36	0.48	0	1	37777
dummy urban area	0.54	0.5	0	1	37777
ln of bank deposits per capita	4.29	1.18	0	5.38	37777
dummy=1 post-2000 reform	0.53	0.5	0	1	37777
dummy=1 post-2006 reform	0.33	0.47	0	1	37777
dummy=1 between 2000 and 2006 reform	0.2	0.4	0	1	37777

Notes: Age and tenure are measured in years, competition is the number of school with similar instructional programs within 30 minutes drive from school, teacher turnover is the share of teachers that left the school in previous year, immigrant student 1st generation are students born abroad of foreign-born parents, immigrant student 2nd generation are native-born children of foreign-born parents. Dummy post-2000 reform is equal to 1 for principals that have less (or equal) 9 years of experience in post, 0 otherwise. Dummy post-2006 is equal to 1 for those with less (or equal) 4 years of experience. Dummy between 2000 and 2006 is equal to 1 for those with experience in post between 9 and 4 years.

Table 2: Descriptive statistics. Sub-sample of principals leading a school they have chosen

Variable	Mean	Std. Dev.	Min.	Max.	Obs.
test score math normalized	49.04	17.95	0	100	29198
management	2.05	0.5	1	3.74	32728
people	1.79	0.43	1	3.5	32728
targets	1.83	0.53	1	3.6	32728
operations	2.32	0.71	1	4.5	32728
leader	2.25	0.57	1	4	32728
monitoring	2.22	0.68	1	4.60	32728
principal's age	58.69	5.27	32	80	32583
gender of principal, female=1	0.37	0.48	0	1	32728
tenure in school	6.01	5.04	1	40	32546
competition	2.37	2.32	0	11	32585
ln of student	6.65	0.52	3.93	7.5	32728
dummy private school	0.03	0.17	0	1	32728
teacher turnover	4.87	5.84	0	50	32547
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 student, 1st gen.	0.06	0.24	0	1	32366
immigrant student, 2nd gen.	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
dummy=1 if student "retained"	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
dummy=1 post-2000 reform	0.51	0.5	0	1	32728
dummy=1 post-2006 reform	0.33	0.47	0	1	32728
dummy=1 between 2000 and 2006 reform	0.18	0.39	0	1	32728

Notes: The sample is composed by observations for schools in which SPs have chosen the school. Age and tenure are measured in years, competition is the number of school with similar instructional programs within 30 minutes drive from school, teacher turnover is the share of teachers that left the school in previous year, immigrant student 1st generation are students born abroad of foreign-born parents, immigrant student 2nd generation are native-born children of foreign-born parents. Dummy post-2006 is equal to 1 for those with less (or equal) 4 years of experience. Dummy between 2000 and 2006 is equal to 1 for those with experience in post between 9 and 4 years.

Table 3: Baseline model

Dependent Variable Test results (Math)	(1) Management	(2) Leadership	(3) Operations	(4) Monitoring	(5) Targets	(6) People
managerial skills	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)
ln students	-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 student, 1st gen.	-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 student, 2nd gen.	-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: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variable is standardized test score in mathematics. Age of principal and tenure in school are measured in years. Competition is the number of school with similar instructional programs within 30 minutes drive. Ln students is the natural logarithm of the number of students in school. Private is equal to one for private schools. Dummy "istituto superiore" is described in text. Lyceum and vocational are dummies equal to 1 for each school type (technical school is the omitted dummy). Teacher turnover is the share of teachers that left the school in previous year. 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 been previously retained. 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, city dummy is the omitted category. All regressions include regional dummies.

Table 4: Instrumental variables model

Dependent Variable Test results (Math)	(1) Management	(2) Leadership	(3) Operations	(4) Monitoring	(5) Targets	(6) People
managerial skills	8.22**	10.19**	5.10*	5.85**	5.15	9.20*
	(3.811)	(4.764)	(2.975)	(2.565)	(3.637)	(4.823)
age principal	0.33**	0.32**	0.30**	0.30**	0.25*	0.28**
	(0.142)	(0.150)	(0.142)	(0.132)	(0.136)	(0.143)
gender principal (female=1)	-0.47	-1.10	-0.32	0.40	-0.20	-0.45
	(1.410)	(1.651)	(1.444)	(1.253)	(1.471)	(1.486)
tenure in school	-0.08	-0.07	-0.02	-0.07	-0.04	-0.15
	(0.131)	(0.141)	(0.124)	(0.126)	(0.120)	(0.153)
competition	0.36	0.54*	0.26	0.37	0.31	0.29
	(0.289)	(0.330)	(0.298)	(0.285)	(0.281)	(0.296)
ln students	-1.71	-2.44*	-1.21	-1.23	-1.51	-1.76
	(1.199)	(1.408)	(1.161)	(1.075)	(1.221)	(1.283)
private school	-5.84**	-5.94**	-4.53*	-4.10*	-5.01**	-10.59***
	(2.377)	(2.455)	(2.353)	(2.401)	(2.371)	(3.805)
teacher turnover	0.12	0.16	0.13	0.11	0.13	0.07
	(0.087)	(0.098)	(0.090)	(0.082)	(0.081)	(0.095)
lyceum school	5.46***	5.53***	5.07***	5.34***	5.15***	6.07***
	(1.327)	(1.404)	(1.372)	(1.228)	(1.318)	(1.526)
vocational school	-8.65***	-8.70***	-8.60***	-8.88***	-9.15***	-8.57***
	(1.445)	(1.488)	(1.549)	(1.333)	(1.383)	(1.512)
"istituto superiore"	-3.64***	-3.96***	-3.25**	-3.14**	-3.64**	-3.37**
	(1.308)	(1.416)	(1.262)	(1.226)	(1.462)	(1.331)
immigrant student, 1st gen.	-2.04***	-1.64**	-1.87***	-2.26***	-2.14***	-1.92***
	(0.613)	(0.665)	(0.626)	(0.610)	(0.616)	(0.620)
immigrant student, 2nd gen.	-2.14***	-1.93***	-2.00***	-2.28***	-2.33***	-2.13***
	(0.624)	(0.653)	(0.642)	(0.621)	(0.617)	(0.643)
gender_student (female=1)	-6.38***	-6.33***	-6.41***	-6.44***	-6.54***	-6.40***
	(0.530)	(0.533)	(0.569)	(0.506)	(0.539)	(0.559)
socio_economic background	1.08***	1.06***	1.11***	1.11***	1.07***	1.01***
	(0.194)	(0.193)	(0.190)	(0.190)	(0.192)	(0.205)
student behind	-5.02***	-4.86***	-5.08***	-5.11***	-5.15***	-5.04***
	(0.407)	(0.427)	(0.439)	(0.392)	(0.408)	(0.420)
dummy rural areas	0.30	-0.59	0.27	0.94	-0.24	-0.19
	(2.258)	(2.244)	(2.186)	(2.253)	(2.295)	(2.183)
dummy non rural areas	0.75	0.06	0.31	1.17	0.47	0.31
	(1.386)	(1.444)	(1.326)	(1.428)	(1.355)	(1.307)
ln deposits per capita	0.89**	1.01**	0.81*	0.87**	0.90**	0.80*
	(0.418)	(0.438)	(0.435)	(0.412)	(0.409)	(0.423)
Constant	19.31	21.34	21.85	21.72	30.43**	22.11
	(15.016)	(15.511)	(14.220)	(14.198)	(13.239)	(14.797)
Observations	27,775	27,775	27,775	27,775	27,775	27,775
R-squared	0.229	0.210	0.231	0.240	0.237	0.215
Hansen J p-value	0.621	0.637	0.573	0.579	0.457	0.661
F (excluded instruments)	2.76	2.13	1.72	3.32	3.48	2.02
No. clusters	249	249	249	249	249	249

Notes: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Instruments for managerial practices are a dummy=1 if the principal has experience in post less or equal to 4 years (dummy post-2006 reform), a dummy=1 if the principal has experience in post between 9 and 4 years (dummy=1 between 2000 and 2006 reform), and their interaction with geographical area dummies: North-West, North-East, Center, South. See Notes to Table 3 for additional details regarding control variables.

Table 5: Interaction with tenure

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Test results (Math)	Management	Leadership	Operations	Monitoring	Targets	People
managerial skills	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 char.	Yes	Yes	Yes	Yes	Yes	Yes
student char.	Yes	Yes	Yes	Yes	Yes	Yes
area char.	Yes	Yes	Yes	Yes	Yes	Yes
regional dum.	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: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. See Notes to Table 3 for additional details regarding control variables.

Table 6: Different sample: excluding pre-2000 reform principals

Dependent Variable Test results (Math)	(1) Management	(2) Leadership	(3) Operations	(4) Monitoring	(5) Targets	(6) People
managerial skills	6.11***	5.24***	2.15**	4.80***	5.88***	5.23***
	(1.73)	(1.65)	(1.04)	(1.18)	(1.68)	(1.94)
age principal	0.31**	0.22	0.20	0.31**	0.30**	0.19
	(0.14)	(0.15)	(0.17)	(0.13)	(0.15)	(0.15)
gender principal (female=1)	-2.75*	-2.49	-2.61	-2.10	-3.21*	-3.04*
	(1.59)	(1.59)	(1.69)	(1.53)	(1.63)	(1.67)
tenure in school	0.86*	0.70	0.68	0.85*	1.02*	0.72
	(0.50)	(0.49)	(0.52)	(0.49)	(0.51)	(0.51)
competition	-0.35	-0.31	-0.34	-0.25	-0.32	-0.41
	(0.41)	(0.40)	(0.43)	(0.42)	(0.37)	(0.37)
ln students	-4.36***	-3.76***	-3.64***	-4.35***	-4.83***	-4.18***
	(1.29)	(1.30)	(1.37)	(1.27)	(1.34)	(1.33)
private school	-10.72***	-10.14***	-10.64***	-9.62**	-10.73***	-14.02***
	(3.76)	(3.61)	(3.81)	(3.78)	(3.77)	(3.89)
teacher turnover	0.09	0.12	0.08	0.07	0.08	0.05
	(0.08)	(0.08)	(0.08)	(0.07)	(0.08)	(0.08)
lyceum school	7.21***	6.87***	6.96***	6.91***	7.10***	7.67***
	(1.71)	(1.77)	(1.86)	(1.68)	(1.73)	(1.74)
vocational school	-11.08***	-11.45***	-11.24***	-11.22***	-11.41***	-11.11***
	(1.44)	(1.47)	(1.49)	(1.38)	(1.45)	(1.48)
"istituto superiore"	-2.29	-2.32	-1.85	-1.78	-2.62	-2.00
	(1.62)	(1.66)	(1.62)	(1.58)	(1.68)	(1.59)
immigrant student, 1st gen.	-2.50***	-2.20***	-2.36***	-2.65***	-2.66***	-2.31***
	(0.64)	(0.65)	(0.65)	(0.63)	(0.63)	(0.63)
immigrant student, 2nd gen.	-1.79**	-1.68**	-1.83**	-1.93**	-2.01***	-1.74**
	(0.76)	(0.77)	(0.78)	(0.77)	(0.76)	(0.77)
gender_student (female=1)	-5.72***	-5.89***	-5.89***	-5.79***	-5.69***	-5.81***
	(0.78)	(0.79)	(0.85)	(0.77)	(0.79)	(0.80)
socio_economic background	1.21***	1.22***	1.24***	1.25***	1.17***	1.17***
	(0.24)	(0.24)	(0.24)	(0.25)	(0.24)	(0.23)
student behind	-4.50***	-4.52***	-4.67***	-4.55***	-4.57***	-4.55***
	(0.49)	(0.49)	(0.51)	(0.50)	(0.50)	(0.48)
dummy rural areas	2.90	1.78	1.97	3.42	2.65	2.72
	(2.38)	(2.39)	(2.47)	(2.35)	(2.55)	(2.40)
dummy non rural areas	0.83	0.30	0.26	1.11	0.36	0.34
	(1.62)	(1.59)	(1.62)	(1.58)	(1.65)	(1.63)
ln deposits per capita	1.78***	1.67***	1.62***	1.87***	1.90***	1.62***
	(0.56)	(0.59)	(0.55)	(0.52)	(0.57)	(0.57)
Constant	51.15***	52.65***	60.22***	51.65***	55.67***	60.18***
	(13.93)	(14.28)	(14.83)	(13.21)	(12.95)	(13.98)
Observations	14060	14060	14060	14060	14060	14060
R-squared	0.318	0.315	0.305	0.322	0.317	0.311

Notes: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The sample is composed of principals with less (or equal) than 9 years of experience in post (those appointed post-2000 reform). See Notes to Table 3 for additional details regarding control variables.

Table 7: Different institutional constraints

Dep. var.: Invalsi test	(1) Low constraints	(2) Medium constraints	(3) High constraints
OLS			
managerial skills	2.48**	1.41	0.59
	(1.16)	(1.08)	(1.56)

Notes: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Controls are as in Table 3; see Table 11 and the Appendix for additional details regarding the definition of constraints.

Table 8: School level regressions: dispersion measures and percentiles

	(1) Management	(2) Leadership	(3) Operations	(4) Monitoring	(5) Targets	(6) People
row 1 - avg test scores						
managerial skills	1.40	2.02**	0.47	1.93**	0.62	-0.25
	(1.162)	(1.002)	(0.837)	(0.835)	(1.105)	(1.302)
row 2 - 10 pct						
managerial skills	1.35	2.29**	0.27	1.48*	1.16	-0.16
	(1.209)	(1.040)	(0.871)	(0.874)	(1.147)	(1.354)
row 3 - 50 pct						
managerial skills	1.66	2.20*	0.60	2.24**	0.76	-0.16
	(1.293)	(1.116)	(0.932)	(0.928)	(1.230)	(1.450)
row 4 - 90 pct						
managerial skills	0.99	1.42	0.24	1.86**	-0.12	-0.38
	(1.259)	(1.090)	(0.905)	(0.905)	(1.195)	(1.407)
row 5 - 90 10 pct diff						
managerial skills	-0.36	-0.88	-0.03	0.38	-1.28	-0.22
	(0.969)	(0.839)	(0.696)	(0.703)	(0.914)	(1.081)
row 6 - st dev test scores						
managerial skills	-0.40	-0.75	-0.43	-0.03	-0.40	-0.09
	(0.572)	(0.495)	(0.410)	(0.416)	(0.543)	(0.640)
Observations	221	221	221	221	221	221

Notes: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variables are as follows: row 1: average test scores in math; rows 2 to 4: 10th, 50th and 90th percentile of the distribution of test scores; row 5: 90-10 percentile differential in test scores; row 6: standard deviation of the test scores across classes in the same school. Controls are as in Table 3; all variables are calculated at the school level mean. Schools with only one class are excluded. See Notes to Table 3 for additional details regarding control variables.

Table 9: Probit model. Marginal effects

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Retained student dummy	Management	Leadership	Operations	Monitoring	Targets	People
managerial skills	-0.03**	-0.03***	-0.02**	-0.02*	-0.02**	-0.03**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
age principal	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
gender principal (female=1)	0.00	0.00	0.00	-0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
tenure in school	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
competition	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ln students	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
private school	0.06**	0.06**	0.06**	0.06**	0.06**	0.08***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
teacher turnover	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
lyceum school	-0.09***	-0.09***	-0.09***	-0.09***	-0.09***	-0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
vocational school	0.17***	0.17***	0.17***	0.17***	0.17***	0.17***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
"istituto superiore"	0.00	0.00	0.00	-0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
immigrant student, 1st gen.	0.40***	0.40***	0.40***	0.40***	0.40***	0.40***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
immigrant student, 2nd gen.	0.13***	0.13***	0.13***	0.13***	0.13***	0.13***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
gender_student (female=1)	-0.04***	-0.04***	-0.04***	-0.04***	-0.04***	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
socio_economic background	-0.02***	-0.02***	-0.02***	-0.02***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
dummy rural areas	-0.03**	-0.03*	-0.03**	-0.03**	-0.03*	-0.03*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
dummy non rural areas	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
ln deposits per capita	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	28047	28047	28047	28047	28047	28047
No. clusters	249	249	249	249	249	249

Notes: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variable is a binary variable equal to one if the student is lagging behind. See Notes to Table 3 for additional details regarding control variables.

Table 10: 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: The overall measure of Management includes all sections, i.e., Leadership, Operations, Monitoring, Targets and People. See Notes to Table 11 for more details concerning the sections of the questionnaire.

Table 11: Index Numbers and Institutional Constraints

		Canada	Germany	Sweden	UK	US	Constraint
1	Leadership vision	129.72	112.02	103.81	126.23	118.54	1
2	Standardisation of Instructional Processes	131.48	125.56	107.22	127.38	130.44	2
3	Personalization of Instruction and Learning	146.19	121.93	166.00	132.75	133.81	2
4	Data-Driven Planning and Student Transitions	138.20	115.29	124.79	150.45	131.37	1
5	Adopting Educational Best Practices	154.66	141.61	133.62	148.49	147.80	1
6	Continuous Improvement	135.44	130.07	142.97	145.96	132.96	1
7	Performance Tracking	140.62	122.73	154.28	144.74	139.43	1
8	Performance Review	137.40	122.83	141.48	145.12	136.31	1
9	Performance Dialogue	150.44	139.86	164.88	156.66	143.97	1
10	Consequence Management	118.48	115.58	120.40	127.05	121.08	2
11	Target Balance	159.73	134.16	160.73	160.24	144.55	2
12	Target Inter-Connection	186.79	150.50	191.57	194.71	164.99	1
13	Time Horizon of Targets	157.41	140.01	132.33	161.58	146.32	1
14	Target Stretch	153.69	134.35	141.45	158.68	143.08	1
15	Clearly Defined Accountability for School Leaders	140.88	136.05	146.53	151.11	148.34	2
16	Clearly Defined Leadership and Teacher Roles	136.82	124.88	135.54	152.50	128.77	2
17	Clarity and Comparability of Targets	150.68	145.06	144.76	163.20	142.25	1
18	Rewarding High Performers	155.67	143.78	155.65	181.78	169.14	3
19	Removing Poor Performers	105.39	130.44	174.87	136.37	118.63	3
20	Promoting High Performers	109.01	111.52	132.52	142.54	140.34	3
21	Managing Talent	145.39	111.65	112.77	174.21	139.46	3
22	Retaining talent	138.26	131.90	122.16	140.51	128.84	3
23	Creating a Distinctive Employee Value Proposition	105.34	108.03	125.36	125.35	113.02	2
	Leadership	135.23	123.10	126.39	141.57	130.47	
	Operations	142.48	125.72	133.25	139.74	135.63	
	Monitoring	136.10	125.83	144.26	143.54	134.51	
	Targets	161.34	140.60	153.56	167.28	148.02	
	People	126.40	122.81	136.02	149.04	134.28	
	Management	139.80	127.36	139.30	147.26	136.32	

Notes: Each value in the Table is equal to (value country/value Italy)*100. Management is the average score calculated using all questions, Leadership is the average for questions 1, 15 and 16; Operations is the average for questions 2 to 5, Monitoring is the average for questions 6 to 10; Targets is the average for questions 11 to 14 and 17, People is the average for questions 18 to 23.

B Cross country comparisons for individual questions of the survey

In section 4 we discussed how different areas of management are influenced by institutional constraints and we claimed that the delay of Italy with respect to other countries is not completely attributable to such constraints. In this Appendix we explain in more detail how we classify both the single survey questions and the different areas of management based on how much SPs activity is likely to be constrained by the institutional setting. The idea is that, if the lower scores on managerial practices of Italian SPs were mainly due to its institutional setting, then we should expect that Italy's lowest relative scores are in those questions and sections where constraints are more binding while, conversely, in managerial practices where constraints are less binding, we should expect that Italian SPs scores are more similar to that observed in other countries. In other words, if the distance from other countries does not vary systematically with the importance of institutional constraints, then the evidence would suggest that the lower performance of Italian SPs cannot be fully explained by differences in institutional settings.

This classification can also be used to compare the performance of the Italian SPs with that of other countries. We start checking if the Italian data are closer to that of Germany the country where, according to existing evidence, both education and public sector institutional settings are relatively more similar to the Italian one, compared to the UK or the US (Pont et al., 2008). In general, Anglo-Saxon legal systems, based on common law, are less subject to institutional constraints and bureaucratic procedures, while our two remaining countries have more complex institutional settings. Existing evidence indicates a lack of autonomy of Canadian SPs during the process of personnel selection, while Swedish SPs seem able to affect more significantly the selection of the teaching staff. Overall, the reality experienced by Swedish SPs seems closer to the one of the Anglo-Saxons, in contrast to Canada where the institutional system has characteristics more similar to those of Germany and Italy.

Second, we focus on the different areas of management. In most countries schools are dominated by the public sector, with a strong union representation, and tend to be particularly poor in certain areas of human resource management such as promoting and rewarding high performing and/or firing badly performing teachers (Bloom et al., 2012; Pont et al., 2008). In fact, we expect People being the area of management where institutional factors are more binding in some countries than others. Therefore, if the presence of constraints were the main explanation for the Italian gap in managerial practices, People should be the area showing the largest gap between Italy and the other countries, particularly the Anglo-Saxon ones. On the other hand, areas as Monitoring and Targets, that capture other aspects of managerial ability, as the capacity of identifying and communicating targets, or monitoring performance and reviewing results, should be less subject to such institutional constraints.

Third, we focus on the single questions results. In particular, for each question of our survey, Table 11 shows the index numbers using Italy as reference (Italy = 100). All values are then easily interpretable in terms of percentage deviations. That is, for the first question on *Leadership vision* a value of 130 for Canada, implies that Canadian SPs managerial practices obtain a score 30% higher than the Italian SPs. First of all, this Table always shows values higher than 100. Thus, a first clear evidence is that, compared to the other five countries analyzed here, the Italian SPs managerial practices always obtain lower scores. In the last column of the Table we also classify each question on a scale from 1 (low constraints)

to 3 (highest institutional constraints). [46]

In detail, numbers indicate that the distance reported by Italian SPs in the incentives (People) section is smaller than that found the other areas of management. Remember that People is the managerial area where institutional constraints are more binding and where the score is lower in all countries. Conversely, Target is the area where Italian performance is relatively worse while institutional constraints are not significantly binding. In this case, our index numbers vary from a minimum of 140 (compared to Germany) to 167 (UK). In general, the analysis of the scores of the different sections does not seem to show any systematic relationship between the importance of institutional constraints and the relative performance of Italy. [47] When we focus on single questions scores we confirm the previous analysis and find additional interesting evidence. For example, in question 18 on "Rewarding High Performers" (People), which mainly concerns on how SPs identify and promote more talented teachers and where institutional constraints are considered more stringent, index numbers show that the Italian SPs scores are much lower than in Germany (almost 44%), while they are 81% lower than in UK schools. In turn, in question 17 on Clarity and Comparability of Targets, an area in which the institutional constraints are very low, if not absent, Italian SPs are equally distant from SPs in other countries, with a gap in scores that varies from about 40% (Germany, Sweden and the US) to about 60% (the UK).

As a final check, we use our classification of institutional constraint to calculate the correlation between the former and the index numbers reported in Table 11. Indeed, if the differences between Italy and the other countries were mainly due to institutional constraints we should find a positive correlation. Conversely, consistently with what we have seen before, we find a negative correlation (-0.22): as the institutional constraints become more binding, the distance of Italy from the other countries appears to be lower rather than higher. In other words, a greater freedom of action in favor of the SPs is related to a lower relative quality of the management, and this negative correlation is even stronger for Italian SPs when compared with those of other countries with similar institutional setting (-0.41 and -0.25 with Canada and Germany respectively).[48] Also in this case, if the key determinant of the performance of SPs was the Italian institutional framework we would expect the opposite result.

In sum, this evidence shows that institutional constraints cannot fully explain the observed cross countries heterogeneity in managerial SPs activities and the low scores of Italian SPs. Indeed, even when answering to questions where institutional constraints are hardly binding, the distance of the Italian results from those reported by other countries SPs remains significant.

[46] It is fair to say that this classification process involves an unavoidable degree of arbitrariness. It is based, however, on the experience accumulated during the data collection phase, in which we discussed with the SPs for a total of about 600 hours.

[47] In order to control for the goodness of our classification, we have also selected our Italian sample and we have calculated the (within country) standard deviation for each of our 23 questions. In this case, we expect that when institutional constraints are more stringent there is also less variability of performance across the Italian SPs. Our data confirm this hypothesis: they indicate a smaller heterogeneity (standard deviation equal to 0.49) for the section People, for Targets the standard deviation is equal to 0.54, while it has the highest variability for the Operations and Monitoring sections (about 0.68).

[48] The value of the correlation coefficient is reduced but remains negative even when we compare Italy than in countries where institutional constraints are less stringent (Sweden, United Kingdom and United States).