

Evaluating the impact of a principals' professional development program on school management practices: Evidence from Brazil

Bruna Borges^a Gabriel Leite^b Ricardo A. Madeira^c Luís Meloni^d

This version: July 17, 2023

Abstract

Several programs focus on the professional development of school principals by assuming that the quality of management affects the school's functioning and, therefore, students' learning. However, little is known about what skills we should develop to translate better school management into proficiency. In this paper, we propose a survey instrument to measure thirteen managerial practices specifically designed to capture variations between schools within a context of scarce resources and limited autonomy. Then, we assess the impact of a school management program (*Jovem de Futuro*) – with proven effects on students' proficiency – on these practices, exploiting the program's phase-in randomization strategy. We find that receiving the program increased the overall management practices index, especially improving the schools' evaluation processes. Analyzing each dimension separately, we observe that the program affects how principals pursue pre-established targets, use assessments to learn about student learning, identify school leaders, and attribute responsibilities to the school's management team.

Keywords: *Jovem de Futuro*, School management practices, Impact evaluation, Educational economics, Human capital.

JEL Classifications: I20, M10, J01, L38.

^aCorresponding author at: Getulio Vargas Foundation, São Paulo School of Economics, Rua Itapeva, 474, São Paulo CEP 01332-000, Brazil. E-mail: bruna.borges@fgv.edu.br

^bGetulio Vargas Foundation, São Paulo School of Economics, Rua Itapeva, 474, São Paulo CEP 01332-000, Brazil. E-mail: gabriel.leite@fgv.edu.br

^cUniversity of São Paulo, Avenida Professor Luciano Gualberto, 908, São Paulo CEP 05508-010, Brazil. E-mail: rmadeira@usp.br

^dUniversity of São Paulo, Avenida Professor Luciano Gualberto, 908, São Paulo CEP 05508-010, Brazil. E-mail: luis.meloni@usp.br

1 Introduction

Brazil has experienced an increase in public education expenditure at one of the fastest rates through the last decade.¹ However, when we look at international student assessment rankings, we note that the country still struggles to translate this rising educational investment into student learning.² Additionally, from a national point of view, the large disparities in learning within the country cannot be explained by differences in per-student spending (de Barros, de Carvalho, Franco, Garcia, Henriques, and Machado, 2019).

A possible interpretation for this pattern is that poor governance and management are important factors to explain the persistence of low school effectiveness in developing countries (Glewwe and Muralidharan, 2016) and that the leadership role of school principals has been understated over the last decades (Grissom, Egalite, and Lindsay, 2021), which may have diverted effective policies at the school level from this actor.

Based on those hypotheses, there has been a growing interest in the impact of school governance and management practices on students' academic achievement. See, for instance, the works by Fryer (2017) and Fryer (2014), showing that providing management training to principals of public schools increases student achievement in the short-run.³

Despite this robust correlation between management practices and student achievement, the overall understanding of which management practices are important and in which contexts they are more or less likely to be effective is still limited. Measuring management practices is challenging. Even though there are recognized instruments in the literature (Bloom, Lemos, Sadun, and Van Reenen, 2015), the set of measures that school managers can take varies significantly depending on the institutional setting.

This paper contributes to deepening the understanding of governance and management practices within the school. More specifically, the paper empirically investi-

¹Brazil currently spends approximately 6% of its gross domestic product (GDP) on basic education every year, and it has nearly doubled this share between 2008 and 2017 (Tesouro Nacional, 2018). More details are provided in Section 2.

²Brazil was among the lowest positions on PISA 2015 (OECD, 2016). The scenario became even worse in PISA 2018. According to OECD (2019), more than two-thirds of students are below basic knowledge in Mathematics and more than half in Portuguese Language and Sciences.

³Dobbie and Fryer (2013), and Lemos, Muralidharan, and Scur (2021) also document a positive association between management practices and school effectiveness in the United States and India, respectively. Another stream of the literature on school management estimates principal's value-added and finds significant principal effects on student achievement in the United States (Branch, Hanushek, and Rivkin, 2012; Dhuey and Smith, 2018; Miller, 2013; Walsh and Dotter, 2020), Canada (Coelli and Green, 2012), Sweden (Böhlmark, Grönqvist, and Vlachos, 2016) and Chile (Munoz and Prem, 2022).

gates the impact of the *Jovem de Futuro* program, a randomized school management program, with documented impacts on both learning and dropout rates⁴ on school management practices in Brazil.

The *Jovem de Futuro* program was created by *Instituto Unibanco* in 2007.⁵ Its main goal is to increase students' learning and graduation rates by improving the management of public high schools and by developing good management practices for school principals. Since its beginning, the program followed a phase-in strategy in each partner State, which randomized the schools to receive the program immediately (treatment schools) and schools to receive the program from the fourth year on (control schools).

Instituto Unibanco originally designed the experiment to allow the evaluation of the program's effect on student performance through a Randomized Controlled Trial (RCT) in a three-year window. We take advantage of this exogenous variation provided by the original experimental setup to estimate the causal impact of the program on management practices, applying our management survey instrument to the treated and control schools in the same evaluation window. To assess the program's impact on management practices, we interviewed 297 school principals in 2017, three years after the randomized implementation of the program in the Brazilian States of Pará and Espírito Santo.

Considering the context of Brazilian public schools, where principals lack autonomy, we employ a survey instrument specifically developed to capture variation between schools in this setting. The survey instrument investigates thirteen managerial practices regarding organization, planning and execution of school activities, innovation and support of the instructional process, definition and evaluation of school leaders, usage of available tools and information to assess student learning, and school climate and reputation. We have graded the practices on a pre-defined scale ranging from one to five, according to the complexity of each of these practices.

Our results show that *Jovem de Futuro* program increased the overall management practices index by 0.25 standard deviations. The most affected management practices relate to evaluation processes, which increased on average 34% of a standard deviation. Particularly, the program affects how principals use and pursue pre-established

⁴Previous work documented that the program had a positive impact on Mathematics and Language achievement, decreased the probability of dropout (Henriques, de Carvalho, and de Barros, 2020b; de Barros, de Carvalho, Franco, Garcia, Henriques, and Machado, 2019; Rosa, 2015) and also increased higher education access (Finamor, 2017). In subsection 6.2, we replicate the program's positive impacts on students' learning in our sample of schools.

⁵*Instituto Unibanco* is a non-profit organization founded in 1982. Its main focus is to improve public education through educational management. More details on *Instituto Unibanco's* website.

targets, how principals use student learning assessments to evaluate students' academic conditions, and how principals identify school leaders and attribute responsibilities among the management team and teachers. Most affected practices correspond to dimensions that had actions deliberately designed by the program.

Although we cannot identify all the mechanisms through which the *Jovem de Futuro* positively influences students' achievement, we observe for our sample of schools a positive impact of the program both on management practices and students' learning outcomes, indicating that part of the impact comes from managerial practices. Next, we document a positive correlation between average school performance and the management practices index, suggesting that improving school management practices play a relevant role. Considering this correlation between these two variables, the program's impact on management practices would be responsible for at least 50% of the effect of the program on students' achievement.

This paper contributes to two strands of the literature. Our first contribution relates to the literature that causally evaluates programs and interventions deliberately designed to improve school management practices in developing countries, which provide mixed results. For example, [Muralidharan and Singh \(2020\)](#) experimentally evaluated a large-scale program in India that featured comprehensive assessments, detailed school ratings, and customized school improvement plans. The authors find that the program did not affect school functioning or students' outcomes. Alternatively, [De Hoyos, Ganimian, and Holland \(2020\)](#) conducted a similar experiment in Argentina, where treated principals received training and access to an online dashboard to develop, implement and monitor school-improvement plans. The paper documents a reduction in repetition rates and an increase in passing rates, but the effects only emerge after two years of the intervention. In Brazil,⁶ [Tavares \(2015\)](#) evaluated a management program using a fuzzy regression discontinuity design and found significant positive impacts on 8th-graders' math scores after a year, especially for low-performing students. These effects were mainly channeled by management practices such as performance monitoring, target setting, managers' engagement in school planning, and increasing the sharing of performance indicators with teachers and parents.

The mixed results could relate to at least two major facts. First, school management programs and interventions require a massive mobilization of actors, and the quality of the implementation might decrease with its complexity. [Muralidharan and](#)

⁶In a correlational exercise, [Teodorovicz, Lazzarini, Cabral, and Nardi \(2022\)](#) construct a management index that reflects students' perceptions of management practices for Brazilian upper secondary schools. The authors find that public schools are less likely to use the most efficient management practices and that the management index positively correlates with student learning outcomes.

Singh (2020) argue that their null effects can be related to implementation problems, which could hamper even well-designed interventions. Second, those programs may need time to operate and agents to adapt, especially in interventions that systematically change the educational environment and demand new management tools. For instance, reported effects in De Hoyos et al. (2020) increased with time of exposure, which could relate to adaptation. Therefore, our paper also contributes to this debate by adding rigorous evidence of the impact of a well-designed and well-implemented management program (*Jovem de Futuro*) on the management practices of public schools' principals.

Our second contribution relates to understanding which managerial practices matter at the school level in different educational contexts. Recently, related studies are going further to understand the “black box” of principals' roles and investigate relevant school management practices.⁷ The most notable example is Bloom et al. (2015). The authors proposed an international management index (named World Management Survey) for schools allowing comparisons of management practices across countries in four areas: operations, monitoring, target setting, and human resources. They find that management practices are strongly associated with better educational outcomes, especially in autonomous schools.⁸

Although Bloom et al. (2015)'s index enables international comparisons, their survey instrument is not designed to investigate variations within contexts where principals lack autonomy. This fact could help to explain why Brazil performs poorly in their management index and why the correlation between those practices and student outcomes is weaker for Brazil than for other countries in their study. Di Liberto, Schivardi, and Sulis (2015), for example, find a weakly positive association between Math test scores and the management measures of Bloom et al. (2015) in Italy, a country that also performs poorly in this index and where the educational system is more centralized than in other countries surveyed in the World Management Survey (WMS).

To address within-country comparison issues, Lemos et al. (2021) adapted the WMS survey to allow more measurement granularity and capture variations in management practices in low-capacity systems such as India. However, they did not contextualize their survey instrument to Indian-specific characteristics to maintain comparability with WMS.⁹ In our paper, we follow a similar strategy, constructing a management

⁷In a recent publication, Grissom et al. (2021) synthesized quantitative and qualitative studies and suggested that four principal practices are linked to effective outcomes: high-leverage instructional activities, building a productive culture and climate, facilitating collaboration and learning communities, and the strategic management of personnel and resources.

⁸Examples of autonomous schools given by the authors are UK academies and US charters.

⁹Their study finds a positive correlation between measures of people management and teacher prac-

survey index adapted to the Brazilian public schools' setting. Specifically, we present a survey designed to measure managerial skills and investigate the effects of a randomized school management program on these skills.

The remainder of the article is organized as follows: Section 2 discusses the institutional background of the Brazilian Educational System and details the *Jovem de Futuro* program. In Section 3, we describe our data, discuss the management practices' instrument and present descriptive statistics. Section 4 describes the empirical strategy and presents suggestive evidence of its validity. Section 5 reports our main findings. Section 6 discusses the interpretation of our results. Lastly, Section 7 concludes.

2 Institutional Background

2.1 Brazilian context: underachievement and limited principal's autonomy

Despite rising educational investments, Brazil is still one of the most typical examples of student performance stagnation. If Brazil continues to evolve its educational indicators as in recent periods, by 2024, less than a third of high school students will achieve appropriate levels of Portuguese learning and less than 10% in Mathematics (de Barros et al., 2019).

The country spends approximately 6% of its GDP on public education every year,¹⁰ with real growth of 91% between 2008 and 2017 – 7.4% per year (Tesouro Nacional, 2018). However, Brazil performs far worse in the PISA ranking (OECD, 2019). According to de Barros et al. (2019), the country is 0.64 standard deviations below what we expect from countries with the same per-student spending. Brazil also has large differences in per-student spending within the country, but this factor alone cannot explain disparities in learning. Almost 90% of learning differences at the end of elementary education relate to differences between schools with the same per-student spending.

The inefficiency in educational investment raised concerns about how resources are being used at the school level and gave rise to policies aiming to improve school management quality at municipal and state levels.¹¹ Not surprisingly, the importance

tices.

¹⁰The share is higher than the OECD average spending of 5.5% and greater than 80% of the 141 countries with reported values on World Bank Open Data (Tesouro Nacional, 2018).

¹¹One educational policy example at the State level is the pro-EMI program (*Programa Ensino Médio*

of the principals' role became more salient, as well as their attributions and responsibilities, and the restrictions (legal or not) imposed on their activities.

Institutionally, the provision of public education in Brazil is shared by Federal, State, and Municipality authorities. In particular, the high school management is mainly under the State level responsibility, which is liable for developing educational policies, considering the national guidelines established at the Federal level.¹²

Although this decentralized system provided some degree of autonomy and led to different ways of implementing educational policies at the State management level, public schools still face some restrictions that interfere with school management, such as rigid bureaucratic rules, process rigidity, legal and administrative constraints, limited control of their financial resources, and political pressure (de Mendoca and de Barros, 1997; Teodorovicz et al., 2022).

Public institutions' managers have limited autonomy regarding their human resources management, as they cannot freely hire and dismiss employees, and the employees have weak incentives to increase performance because they have job stability, their wages are flat, and promotions are tenure-based. Moreover, principals cannot use school resources for payroll expenses (staff salary, gratifications, and other monetary incentives) (de Mendoca and de Barros, 1997; Teodorovicz et al., 2022).

Another limitation the public school management faces in Brazil is the different principal selection methods. Miranda (2015) and Akhtari, Moreira, and Trucco (2022) find that party turnover for mayoral positions leads to substantial changes in the composition of school's principals and that this turnover is negatively associated with school performance. These findings suggest that, in many cases, the school management position is politically motivated, which could produce management practices not aligned with school goals. The principal chosen by political criteria may not be considered legitimate by the school community or not have the necessary skills for the job position. Elected principals tend to be more qualified in terms of leadership and managerial abilities (Pereda, Lucchesi, Mendes, and Bresolin, 2019).¹³

Inovador), which started in 2009 to support innovative initiatives related to the high school curriculum. At the municipal level, the *Programa Nacional Escola de Gestores da Educação Básica* aims to improve the quality of managers in primary education. More details are provided on the website of the [Ministry of Education](#).

¹²In 2020, 7.6 million students were enrolled in high schools in the country. Of those enrollments, the State educational system comprises around 84%, and the private system accounts for 12% (INEP, 2021).

¹³As stated by Tavares (2015), it is worth noting that once a principal or educational coordinator is appointed, changing professionals is usually a slow and bureaucratic process. Muñoz, Pascual, and Sáez (2021) document that 39.5% of the principals under the State system are elected, whereas a public selection process chooses 12%, and 25% are appointed.

In a context where school managers have scarce resources, limited autonomy regarding investment and personnel decisions, and selection processes do not necessarily follow technical criteria, improvements in the management practices of principals are difficult to assess. Our study takes advantage of a program specifically designed to improve management quality¹⁴, considering both the possibilities and limits regarding school principals' actions. Moreover, we develop a survey management instrument that specifically captures variation in management practices for Brazilian public schools.

2.2 The *Jovem de Futuro* program

The *Jovem de Futuro* program was created by *Instituto Unibanco* and started as a pilot in 2007. After over a decade, it has been implemented in 12 Brazilian states and attended more than 4,700 schools, reaching over 4 million students (Henriques, de Carvalho, and Bittar, 2020a). The program originally focused on school management but has recently progressively expanded to influence managerial practices at other levels of the educational system. The program's main goal is to increase high school students learning and graduation rates by improving the management of public schools.

The program starts by implementing 'trigger actions' in five major areas (Henriques, de Carvalho, and Bento, 2021):

1. Governance: establishment of decision-making groups, composed of members of the Secretary of Education and *Instituto Unibanco*, to monitor the program's functioning by analyzing the implementation data gathered by a system provided by *Instituto Unibanco*.
2. Management training: (a) 168 classroom hours and 120 remote hours for technicians of the secretary department; (b) 72 classroom hours and 144 remote hours for managers of regional secretaries of education; (c) 72 classroom hours and 160 remote hours for school supervisors; and (d) 56-64 classroom hours and 120-160 remote hours for school managers (principals and educational coordinators).

¹⁴Teodorovicz et al. (2022) highlight the performance heterogeneity of Brazilian secondary schools within the same governance level and relate the observed differences to schools' management practices and resources. The authors argue that professional development interventions that improve the school's human capital could lead principals to adopt performance-enhancing practices. The paper cites the *Jovem de Futuro* program as an example of a large-scale program that focuses on human capital enhancement by providing a professional development program to school managers. The authors argue that the channel through which the program impacts students' learning is by enabling principals to select and employ better management practices.

The training focuses on the program's methodology, data system operation, and how to set performance targets and the school's action plan.

3. Technical support: (a) the implementation of a data system that unifies the existing and new information necessary to monitor the program's execution at the schools and to track students' performance, and (b) the provision of a local team of technicians (*Instituto Unibanco* staff) to administer the system and support data interpretation, analysis and the production of technical reports.
4. Mobilization: (a) yearly seminars on school management topics provided by specialists and local leaders, directed to the Education department technicians, principals, and educational coordinators; (b) immersion days in another Brazilian State to experience and share good practices with managers from a different reality; and (c) events that encourage high school students to participate in the management of their schools.
5. Knowledge Management: (a) research committees, composed of secretary members, data technicians, and researchers from local universities; (b) experimental impact evaluation to estimate the program's efficacy in improving students' learning; and (c) the revision of the program's theory of change to guide other evaluations.

After providing management training and equipping managers with protocols, information systems, and technical support, the program stimulates and promotes expertise through the management cycles methodology, inspired by the PDCA method (Plan, Do, Check, and Act) with adaptations to the Brazilian public education context (Henriques et al., 2021). The cycle is composed of the following phases: (i) Agreement of learning and achievement goals; (ii) Planning geared toward achieving those goals; (iii) Coordination of the planning execution; (iv) Monitoring the plan's implementation and assessment of the results; (v) Sharing successful management practices; and (vi) Identification of necessary adjustments, route changes, and redesign of actions (de Barros et al., 2019). The management cycles are held three times a year according to the school calendar.

The program's theory of change states that the repeated interaction between training, management tools, and problem-solving methodology empowers managers to identify and tackle the most pressing school issues.¹⁵ The program is almost exclusively targeted to educational managers. Therefore, at the school level, the program

¹⁵Diagrams of the program's theory of change presented in [Henriques et al. \(2021\)](#) are displayed in Online Appendix B.

does not prescribe any actions specifically oriented to pedagogical practices, such as teacher classroom activities and teaching methods, nor training to teachers or other school employees besides the principal and educational coordinator.¹⁶

Moreover, the program does not offer additional funds to schools. Although resource management is an important dimension, the program does not claim to change schools' performance by relaxing their budget constraints. Hence, even for participating schools, principals still have the same limitation in terms of financial resources.

The theory of change considers that *Jovem de Futuro* positively influences the students' proficiency through improvements in the overall school functioning, which would generate a more appropriate environment for teaching and learning activities, indirectly enhanced by better management practices (Horng, Klasik, and Loeb, 2010; Henriques, de Carvalho, and de Barros, 2020b). Therefore, we would expect the effects of the *Jovem de Futuro* program on student performance to derive from indirect channels triggered by an improvement of managerial practices.

While *Jovem de Futuro* does not directly influence pedagogical practices or schools' financial resources, the program's strategy (particularly its management cycle methodology) is strongly and intensively based on data-driven management decision-making. In particular, most management cycle steps require using available data to inform school priorities, set realistic goals, monitor schools' actions, evaluate results during the year, and assess student attendance, learning, and progression (Henriques et al., 2021).

Considering the program's design, we would expect ex-ante that *Jovem de Futuro* does not directly affect the dimensions strictly related to pedagogical practices (e.g., teaching, pedagogical planning, new classroom methods). In contrast, we would expect the program to influence management practices linked to data analysis and problem-solving methods (e.g., setting up targets, monitoring and assessing indicators of attendance and performance).

The program has three implementation stages. The first three years correspond to the evaluation impact window specifically designed for a rigorous impact evaluation, in which schools were randomly assigned to receive *Jovem do Futuro* program. In Pará and Espírito Santo, the evaluation phase occurred between 2015 and 2017, and it is exactly the window considered in this paper to evaluate the program's effect on the schools' management practices. We applied our survey instrument at the end of 2017.

¹⁶Recently, *Instituto Unibanco* has been developing specific actions to address pedagogical dimensions of school management to increase its effects on students' performance.

The following three years correspond to the expansion phase, devoted to including schools from the control group in the program and strengthening the partnership between *Instituto Unibanco* and the State. This phase occurred between 2018 and 2020 for Espírito Santo, but the program was terminated in Pará in 2018, just after the evaluation window and before the program expansion.¹⁷ Lastly, the program enters the sustainability phase, which aims to transfer all processes and methodology to the partner State to sustain the program's changes after the partnership ends. In Espírito Santo, this phase comprised the period between 2021 and 2022 (Henriques et al., 2020a).

3 Data and management survey instrument

Measuring and evaluating school management quality is challenging for distinct reasons. First, management usually depends on the interaction between the manager's skills, the school, and the institutional context. Therefore, different educational contexts may impose specific resource-related constraints that we must consider to capture the relevant variation for the contexts examined. Second, management is usually a broad definition. Since managerial skills are latent traits, our survey instrument focuses on measuring management quality from what is observable in the field – the management practices – rather than trying to measure the principal's skills.

Bloom et al. (2015) develop the international World Management Survey school management index that comprises 20 management practices, evaluated in a 1 to 5 scoring grid. Despite enabling important international comparisons, the WMS tool is not specifically designed to capture variation in less well-managed institutions, such as the public sector of developing countries, where we observe left-skewed scores' distributions, bunching at the lowest scores (Lemos and Scur, 2017; Lemos et al., 2021). Lemos et al. (2021) propose adapting the WMS instrument for developing countries, constructing a more granular index to analyze school management in India that captures wider variation in a low-performing context. Similarly, we take the recognized WMS instrument as a reference and construct an instrument that adapts the management practices to the context of Brazilian public schools. Our instrument aims to capture the specificities of the national educational system, which will greatly vary according to each country's institutional context.

We orient the construction of our survey instrument in a diverse conceptual ground-

¹⁷The termination was unrelated to the program's implementation as explained in Henriques et al. (2020a). Despite observing considerable adhesion between school principals and the State Educational Office, the budgetary constraints hampered the possibility of continuing the partnership.

ing, exploiting theoretical, normative, and empirical references. For the United States, [Grissom et al. \(2021\)](#) surveys the previous literature and summarizes successful school management practices. Our instrument is based on the previous literature, encompassing relevant actions mentioned by the report, such as the establishment of data-driven decisions; the creation of learning communities; the maintenance of a cooperative school environment; the recruitment, retention, and development of effective employees; the promotion of school community engagement, etc.

We also rely on official documents for normative references. For England, Ofsted provides a school inspection handbook that guides good practices in four dimensions: students' broader development, education quality, behavioral aspects, and management and leadership. Regarding management practices, the handbook displays dimensions related to staff professional development and support, pedagogical content (such as the school curriculum), school environment, a culture of high expectations, parental and community engagement, financial management, and students' equality of opportunity. In Canada, an important reference is The Ontario Leadership Framework (OLF), which uses extensive research to build a practical guide that describes the effective practices of educational leaders. They identify five main capacities essential in leaders to advance educational outcomes: goal setting; resources (human, capital, financial, pedagogical) management; promotion of collaborative environments; data-driven decision making; and introduction of innovative practices through dialogue and feedback.

Finally, our instrument adapts the management practices to the context of Brazilian public schools. Brazil recently approved a normative document describing a school principal's basic skills and assignments. As our main goal when constructing the instrument was to adapt it to this setting, we note that our instrument covers several managerial dimensions, practices, and skills mentioned in this official document. The document defines general and specific abilities matrixes, grouped into four dimensions: political-institutional, pedagogical, administrative-financial, and personal and relational. In addition to the normative and theoretical references already mentioned, when adapting the instrument to the reality of Brazilian public schools, we held several rounds of conversation with public education managers, intending to build a measure focused on their real context, representing the school management actions, an empirical measure rather than a normative framework.

Our management survey instrument comprises thirteen major managerial practices, graded in a five-categories rubric system. The higher the score, the greater the quality and complexity of a task the school principal perform in a particular man-

agerial dimension. The topics addressed refer mainly to organization, planning, and execution of school activities, innovation and support of the instructional process, definition and evaluation of school leaders, use of available tools and information to assess student learning, and school identity. Figure 1 sums up our instrument and we provide details of each management practice in Online Appendix C.

We collected the primary data through phone interviews with principals between September and November of 2017, in partnership with the State Secretaries of Education, almost three years after the *Jovem de Futuro* program was implemented in the States of Espírito Santo and Pará. The interviews were based on a predefined script of open questions and conducted as a guided and anonymous conversation. Interviews were recorded, and responses were graded twice by different graders to reduce measurement error. The first interviewer graded the management practices in real-time, and then a second listener graded them after the conversation.

3.1 Survey Instrument Validity

We have taken several precautions to validate the survey instrument and the interview method. First, we applied the survey instrument as a pilot to test if we captured enough variation within the States, investigated if the script was response-inducing, and checked if we were not evaluating redundant management practices. Second, we conducted robustness checks regarding the agreement of grades to verify if the captured variation in the sample was not merely random.¹⁸ Third, we rigorously trained all interviewers based on real situations and did not mention the *Jovem de Futuro* program during the training. Finally, we employ a ‘double-blind’ interview method, where neither school principals nor interviewers knew that the interviews gathered data for an impact evaluation. To assure that there was no mention of *Jovem de Futuro* program or the impact evaluation, we presented this research as being conducted by *Universidade de São Paulo* (USP) in partnership with the State Secretary of Education. We told them that the research aimed to gather evidence to improve public policies related to educational management. We provide the official research statement sent by the Secretary to the schools in Appendix D.¹⁹

¹⁸The agreement rates varied from 43% to 65% in the 13 managerial practices. In Table A.1, we present the kappa-statistic measure of the interagreement rate, considering that we have two raters for each management practice, but their identities vary. Generally, we observe reasonable levels of agreement rates in all the analyzed dimensions. Therefore, we can reject the hypothesis that the raters randomly score schools for all thirteen managerial practices at a 5% significance level.

¹⁹Regarding the interview protocol, we instructed the interviewers to introduce themselves as OPE-Sociais researchers conducting a survey on management quality. When listening to the phone interviews, we noticed that the research presentation depends on each interviewer in practice. Therefore,

We present a series of exhibits to provide evidence that our management measure captures relevant variation and is associated with other (independent) measures at the school level. In Figure 2, we first show substantial variation in the overall management index, with schools scoring from 2.5 to almost 5. It also shows that, with a few exceptions, our survey instrument generates significant variation within practices, with grades ranging from 1 to 5 in most of them.

Next, in Figure 3, we present correlations between our management index and average school performance (average scores from the 2017 state evaluations) in Portuguese Language and Mathematics. Overall, we find strong correlations between our measures of school management practices and students' achievement.

Finally, in Figure 4, we report correlations between our management survey data and a self-reported survey instrument that asked principals and teachers about several aspects of the school's environment, including those related to management quality. Specifically, we correlate our management practices index with data collected by this independent instrument, considering only the questions related to school management.²⁰ More details about this survey instrument can be found in [Vinha, Morais, and Moro \(2017\)](#). We find a strong and positive correlation between our management index and the measure captured by this external instrument.

3.2 Sampling and Descriptive Statistics

The data on the original experiment assignment was provided by *Instituto Unibanco*. Table 1 illustrates the experiment design and sample composition. In the Brazilian state of Espírito Santo, the program evaluation design comprehended a sample of 221 schools with at least 120 high school students in 2015: 151 were assigned to the treatment group and 70 to the control group. The program design was similar in Pará, where 45 schools were allocated to the treatment group, and 42 schools remained in the control group. Therefore, the experiment sample comprises 308 schools, considering both states.

The program randomization occurred within strata based on the predicted Mathematics and Portuguese performances between 2014 and 2016 and the schools' socioe-

to deal with potential differences derived from the protocol used by each interviewer, we included a robustness analysis in Table A.5 where we control for fixed effects of the researcher who conducted the phone interview at that school. Importantly, as the interviewers did not know about the *Jovem de Futuro* program or its impact evaluation, they did not mention these elements in the interview.

²⁰We list the questions used in this particular exercise in Appendix E.

conomic index.²¹ In Espírito Santo, there were 59 strata with three schools, where two schools were allocated to the treatment group, and 11 strata with four schools, three of them assigned to the treatment group. In Pará, there were 39 strata with two schools for which the experiment randomized one school to the treatment group and the other school to the control group, and 3 strata with three schools, two schools were randomly assigned to the treatment group, and one was assigned to the control group. In Appendix F, we summarize *Instituto Unibanco's* technical report (*OPE-Sociais*), which describes the randomization process for schools in Espírito Santo and Pará.

From the total of 308 schools, we could not interview principals in eleven schools, so our final sample comprises 297 schools, 193 treated, and 104 controls. In Table A.2, we present some relevant school characteristics, comparing the interviewed schools with schools we did not interview. We note that they are similar, suggesting that our sample is representative of the schools that participated in the *Jovem de Futuro* program.

We present descriptive statistics about our management practices' survey instrument in Table 2. We observe that the practices' average scores by dimension in our sample vary from 2.70 to 4.66. Principals perform worse in practices related to human resources (workers evaluation and performance management and retention) and school targets, and perform better in practices related to using data to analyze students' flow and in assessments to evaluate students' achievement. When we split our sample by treatment status, we observe that the average scores across treatment units are systematically higher than those of control units in almost all thirteen analyzed practices.

4 Empirical Model

To evaluate the impact of *Jovem de Futuro* on school management practices, we took advantage of the random assignment of the program, an experiment originally designed by *Instituto Unibanco* to evaluate the impact of *Jovem de Futuro* on students' performance. The assignment was based on a phase-in strategy, which randomized schools within strata to receive the program immediately (treatment schools) or to receive the program from the fourth year on (control schools).

²¹This index is called School Socioeconomic Level Indicator (*Indicador de Nível Socioeconômico – INSE*). It comprises vulnerability variables to allow comparisons between schools. It was created by *Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira* (INEP), an agency related to the Brazilian Ministry of Education, responsible for the national educational evaluation system.

Considering the grouping procedure described, the assignment mechanism allows the program’s impact to be causally measured through a simple mean difference between treated and control schools in the three-year evaluation window. We evaluate the differences in management practices after almost three years of the program’s implementation in 2017.

We estimate the impact by the following equation:

$$Y_{is} = \beta_0 + \beta_1 P J F_{is} + \gamma_s + \epsilon_{is} \quad (1)$$

where Y_{is} is a measure of school management quality of school i in the stratum s , $P J F_{is}$ is a binary variable equal to one if school i in the stratum s received the program and zero otherwise, and γ_s is the stratum fixed effect.²² We also cluster the standard errors at the strata level, both to account for the correlation of the treatment assignment at this level (Abadie, Athey, Imbens, and Wooldridge, 2022) and to correct the degrees of freedom given the strata fixed effects in the main specification (Cameron and Miller, 2015; de Chaisemartin and Ramirez-Cuellar, 2020).

Our estimates will be unbiased (i) if there are no systematic differences in pre-determined covariates that might correlate with treatment status and the outcomes of interest. The randomization strategy should deal with this first issue; (ii) if attrition is independent of the potential outcomes. In Table A.2, we show evidence that attrition in our experiment does not correlate with the observable characteristics or the treatment status, suggesting the validity of this assumption.

4.1 Experiment validity

To assess the validity of the experimental design, we provide evidence in Table 3 that several relevant pre-determined characteristics are similar between schools in the treatment and control groups. In columns (1), (2), and (3), we present the averages and standard deviations of school characteristics considering the total sample, treatment, and control groups, respectively. Finally, the fourth column presents the estimated difference between the two groups conditional on strata and state fixed effects. We find no significant differences between the groups in any pre-determined characteristic at a 10% significance level. We test if the estimated coefficients are jointly significant as an additional exercise and fail to reject the null hypothesis, which provides additional

²²We have included specific dummies for each stratum in the states of Espírito Santo and Pará. The stratum fixed effect will absorb the state fixed effect in our specification. Therefore, the stratum dummies control both for stratum and state fixed effects.

evidence that treatment status does not correlate with the sample composition.

Also, we compare the mean score for the Management Complexity Index constructed by INEP, designed to capture the potential difficulties of running a school as a function of educational inputs. The index uses data from the Brazilian School Census to measure the complexity of school management. The Management Complexity Index is based on the Item Response Theory (IRT), considering a single latent trait. The following school characteristics compose the index: school size in terms of enrollment, number of school shifts, and quantity and complexity of the grades/teaching modalities offered. The higher the school size and the number of grades/modalities and shifts offered by the school, the higher the complexity index. Schools offering grades for older students (such as high school) have a higher score in this index. From the score obtained, INEP divides schools into six categories according to their similarity in the index (1-6 scale).²³ As our sample is composed of secondary schools, they are considered complex by the INEP index, with an average of 4.9. The Management Complexity Index could reveal unobservable school features regarding baseline characteristics. We verify that treated and control schools present similar values for this index before the program's implementation. This evidence suggests that our estimates are not capturing differences in the initial easiness of improving school management quality.

5 Results

This section presents the paper's main results, the effect of the *Jovem de Futuro* Program on management practices. We report the main results in Table 4. We pool the schools from the States of Espírito Santo and Pará, and control for strata-state fixed effects in all estimates. We begin by showing that the program had a positive and significant effect on the average management score of 0.130 points, corresponding to an increase of 0.25 of the control group's standard deviation.

In the first panel, we proceed to the program results in our set of *Pedagogical Practices*. Even though most of the estimated coefficients are positive, we can only reject the null hypothesis for the dimension related to *School leaders definitions and tasks*. For this particular dimension, we find that the program increases its score on average by 0.15 points, corresponding to 0.25 of the control group standard deviation.²⁴

²³Technical details of the Management Complexity Index provided by INEP (2014).

²⁴In most dimensions, we observe positive point estimates, although we fail to reject the null hypothesis of a zero effect. To better understand the concerns related to power, we present the minimum

In the second panel, we report the effects for dimensions related to *Human Resources & School Identity*. Likewise, although all the point estimates are positive, we fail to reject the null hypothesis for the three dimensions. For the dimension related to *Workers evaluation*, the point estimate is much larger than the ones we found in other dimensions but is not statistically significant at a 10% level.

Finally, in the last panel, we document the program's impacts on the school's *Evaluation processes*. In this dimension, we find a robust impact of the program. We document positive and significant coefficients for all three practices: *External learning assessment* (0.247 points, equivalent to 0.24 s.d. of the control group); *Internal learning indicators* (0.217 points, equivalent to 0.28 s.d. of the control group); *School targets* (0.359 points, equivalent to 0.29 s.d. of the control group).

Analyzing the program's effects on each of the thirteen dimensions of management practices can raise concerns related to multiple hypothesis testing. Therefore, in addition to the positive and significant results for the overall index of management – which already play a role in alleviating such concerns – we group the dimensions in each panel to investigate the overall impact of *Jovem de Futuro* program on *Pedagogical Practices*, *Human Resources & School Identity* and *Evaluation processes*. We verify that the program positively affects the management practices related to *Evaluation Processes* (0.274 points, equivalent to 0.34 s.d. of the control group), as we would expect, as this set concentrates most of the previous results. The estimates for *Pedagogical Practices*, which includes the dimension *School leaders definitions and tasks*, and *Human Resources & School Identity* are both positive but insignificant at standard levels.

As an additional exercise, we investigate whether our effects vary in the two different states the program was implemented. Since our sample size is not large, this exercise comes with the cost of losing precision in our estimates. The results are presented in Table A.4. Overall, the point estimates of most dimensions and the average effect index are positive. However, some lose statistical significance due to reduced statistical power when analyzing the states separately.

Finally, we run a robustness check for each of the thirteen management practices, including interviewers' fixed effects in our regressions. As Section 3 mentioned, two different interviewers rated each school. First, the researcher that conducted the phone interview rated the school. Then, a second researcher listened to the interview and rated it. In our main analysis, we use the average score for each school, and the unit of analysis is the school. In this exercise, each observation is the score given by each

detectable effect sizes in Table A.3, which suggests that we may not be able to capture small effects (in terms of magnitude) in our sample.

rater (2 observations per school), and we expand the model presented in equation (1) to include interviewer-order (first or second rater) fixed effects. Table A.5 reports the results for this alternative specification. Like in our main specification, we observe positive coefficients for most management practices (11 out of 13), but we only observe statistically significant impacts of the *Jovem de Futuro* program on the three management practices related to the evaluation processes: *External learning assessment*, *Internal learning indicators*, and *School Targets*. Besides, we note a marginally significant (10% significance level) positive impact on the practice *Teaching and learning customization*.

6 Discussion

This section discusses the interpretation of our results. First, we provide a qualitative interpretation of our effect sizes regarding modifications in school managerial practices. Second, we reproduce the program's impact on students' learning in our sample. Lastly, we look at the correlation of our management practices index and present some back-of-the-envelope calculations to assess how much of the observed effects of *Jovem de Futuro* on student proficiency are due to changes in management practices.

6.1 Qualitative interpretation of our main findings

Comparing the effect sizes in Table 4 to the existing evidence in the literature is not trivial. Related studies mainly investigate the impacts of management programs on students' performance or correlate variations in management practices indexes with school characteristics. However, they do not intend to evaluate the causal impact of a program on principals' management practices.

To our knowledge, few papers do that, especially in an experimental context. The closest paper in that matter is [Muralidharan and Singh \(2020\)](#). Although the authors experimentally evaluate a program in India similar to *Jovem de Futuro*, they focus on the impacts on students' test scores and measures of school functioning rather than on school management practices. Their results show that the Indian program is not effective, and – consistently with the null effects – they find no impact on monitoring and inspection by officials, as well as no changes in pedagogical practices (such as usage of textbooks and workbooks), instructional time, or teacher/student absence patterns. The authors attribute the failure of the Indian program to implementation

²⁴We interact the researcher identifier with the order he rated the scores, during the phone interview (first) or after the interview (second).

problems, especially regarding the support and monitoring of the schools' goals after the first year, which may have contributed to characterizing the program as a 'soft' intervention. Unlike them, we evaluate a well-implemented management program with documented impacts on student performance. In this sense, our setting allowed us to provide evidence of management practices as a potential mechanism to improve students' learning.

To better interpret our results, we will shed light on each one of the school management practices for which we find a statistically significant effect of the program. First, the dimension *School leaders definitions and tasks* relates to how principals identify school leaders and attribute responsibilities to the management team and teachers. Our survey instrument evaluates how principals define staff members' roles and how principals address changes in attributions from time to time. It also evaluates the principals' tasks (bureaucratic vs. pedagogical). Considering the average score of the control group, the estimated effect in this dimension is consistent with part of the treated principals adding a practice of developing and revising attributions with school employees' whenever required.

A second dimension we verify a positive and significant effect is *External learning assessment*. It relates to how principals use national and state-level external evaluations to analyze students' learning conditions. In particular, our survey instrument evaluates whether principals know about state-level standardized tests²⁵ and how they assess the school performance. It also evaluates the pedagogical actions triggered in the face of the school performance and how the results are shared with the school community. Considering the average score of the control group, the program's effect in this dimension is consistent with part of the treated schools adding a reflexive and diagnostic process about those results involving all school community (parents, students, teachers, and staff members) to elaborate a school plan.

The third dimension for which we document an impact is *Internal learning indicators*, measuring the presence of internal tools to assess student learning and its frequency and evaluating if principals contrast these indicators with external assessments, the actions taken in the face of students' performance, and how these results are shared with the school community. Considering the average score of the control group, the observed effect in this dimension relates to part of the treated schools adding the practice of comparing internal and external evaluations to elaborate the schools' action plan.

²⁵The national system of evaluation is called *Sistema de Avaliação da Educação Básica* (SAEB). Each state also has its evaluation program, such as *Programa de Avaliação da Educação Básica do Espírito Santo* (PAEBES) in Espírito Santo and *Sistema Paraense de Avaliação Educacional* (SisPAE) in Pará.

Lastly, *School targets* evaluates the relevance of establishing targets for student learning. Our survey instrument evaluates if principals define internal targets for their schools and use the externally established targets (by the Regional or State Educational Office, for example). Most importantly, we ask who is involved in elaborating those targets, which actions (if any) are taken to pursue and achieve them, and the frequency (if any) of target monitoring and revision. Considering the average score of the control group, the *Jovem de Futuro* impact in this dimension is equivalent to part of the treated schools adding practices related to the diversity of targets, monitoring those targets, and involving the school community in discussions about them.

6.2 Impact of *Jovem de Futuro* on students' learning

Our analysis finds that the program causally improved the management practices at the school level, suggesting that we can interpret the impacts on school management as a mechanism to improve students' performance.

Previous works document positive and significant effects of the program on student performance. Considering all States attended by the program²⁶, [de Barros et al. \(2019\)](#) and [Henriques et al. \(2020b\)](#) show that schools that received *Jovem de Futuro* increased Portuguese Language and Mathematics scores by 4.4 and 4.8 points on the national exam's scale, which is equivalent to 9% and 12% of a standard deviation.²⁷

However, to directly interpret the impacts on management practices as a potential channel, we investigate the program's impacts on students' performance in our sample of schools. Therefore, in this subsection, we employ the same empirical strategy (regression model (1)), including school performance in Language (Portuguese) and Mathematics as response variables. We use data from the state evaluations on students' learning for the same year (2017) and states (Espírito Santo and Pará) from our main analysis. The data comprises, at the school level, information on students' average proficiency in Math and Portuguese in state evaluations (PAEBES and SisPAE).

We present these estimates in Table 5. Importantly, we verify positive impacts of the *Jovem de Futuro* program on students' learning outcomes after three years of the program's implementation. Reassuringly, we verify impacts of similar magnitude on

²⁶Considering the first four states that implemented *Jovem de Futuro*, [Rosa \(2015\)](#) finds a similar impact, ranging from 16% to 18% of a standard deviation, and [Finamor \(2017\)](#) finds positive and sizable mid-term effects of the program on public college attendance.

²⁷According to the authors, the effects in Pará and Espírito Santo range between 2.5 and 4.6 points in Portuguese and between 3.5 and 10.1 in Mathematics.

the management index for the subsamples of schools.²⁸

To deal with potential attrition biases, in their impact evaluation, *Instituto Unibanco* dropped all schools from strata where they lost at least one school. Therefore, their sample comprises a subsample of our school data, containing 183 schools in Pará and Espírito Santo.²⁹ We report the exercise restricting the data to the same sample in Panel *Original impact evaluation*. *Jovem de Futuro* positively affects learning outcomes, increasing by around 5 points both Portuguese and Math average test scores, corresponding to an increase by 0.16 and 0.18 of the control group standard deviation, respectively. These results align with the program’s most recent performance evaluation ([de Barros et al., 2019](#)).

In Panel *All Schools*, we display the results for all schools we have information on average proficiency in Portuguese and Math in state evaluations. The program increased the average Mathematics and Portuguese test scores by 3 points, corresponding to 10% of the control group standard deviation. For Math test scores, we observe a large positive coefficient. Still, the impact is not statistically significant at a 10% level.

6.3 Correlation between school management practices and students’ performance

We have provided causal evidence that *Jovem de Futuro* improves management practices, particularly those related to evaluation processes. Moreover, our results in subsection 6.2 indicate the program also positively affects student proficiency, suggesting that management practices mediate the impacts on students’ learning. To better interpret our results, we investigate how management practices are related to the student’s proficiency.

Specifically, we investigate how management practices, as measured by our survey instrument, correlate with the standardized grades in Portuguese Language and Mathematics, measured by the school’s average score in the state evaluation exams.³⁰ Figure 3 visually reports this correlation, showing that an increase of 1 point in our aggregate measure of management practices is associated with an average increase in both Language and Mathematics standardized scores of 19.9 and 23.7 points, respec-

²⁸In Table A.6, we present the impact of *Jovem de Futuro* on the 13 management practices for these subsamples of schools. Results remain similar, with large impacts observed in *Evaluation processes*.

²⁹In our study, as we see that balanced treated and control schools in terms of pre-treatment characteristics (Table 3), we opted to keep all remaining schools to increase statistical power.

³⁰Since the program affects both management practices and student proficiency, we compute this correlation only for schools in the control group.

tively.

If we assume that the correlation between our survey instrument and test scores for the control group captures the effect in the absence of the program, a back-of-the-envelope calculation, given by the product of the causal effect of the program on management practices and the coefficient of the correlation between our survey instrument and school performance in the control group, would give the *Jovem de Futuro* impact on school test scores if all the effect was mediated by the school's management practices.

Considering the aggregate effect of the program for the subsample of schools presented in Panel *Original impact evaluation* in Table 5, as the impact on the management score is 0.129, the effect on student proficiency through an increase in management practices would be equivalent to 2.6-3.1 points in the state evaluation exams, corresponding to 51% of the impact on Portuguese test scores and 58% of effect on Math test scores.

Alternatively, if we consider the effect of 0.123 points on the overall management index and the impact on learning outcomes presented in Panel *All Schools* in Table 5, the effect on student Language test scores through an increase in management practices would be equivalent to 2.4 points in the state evaluation, which corresponds to about 82% of the total impact.

7 Conclusion

In this paper, we exploit a school-level randomized management program (*Jovem de Futuro*) to evaluate the causal impact on school management practices in Brazil. Furthermore, given the documented impacts of the program on student performance, we investigated if one of the intended mechanisms stated by its theory of change can be driving those results.

We took advantage of the phase-in randomization strategy and interviewed school principals of the treated and control groups three years after the program implementation, using a survey instrument specifically developed to capture variations between schools within a context in which school managers have scarce resources and little autonomy regarding investment decisions.

We find that the *Jovem de Futuro* program increased the overall management practices index. The program affects how principals use and pursue pre-established targets, use student learning assessments to examine students' academic conditions, and

identify school leaders and attribute responsibilities among the management team and teachers. The program's effect is strongly concentrated in the set of dimensions that we called *Evaluation Processes*.

Although we cannot directly identify all the mechanisms through which the *Jovem de Futuro* program positively influences students' performance, our results suggest that improving school management practices plays an important role. First, we show that the program positively impacted students' performance and management practices in our sample of interviewed schools. Then, we document a positive correlation between school performance and our aggregate measure of management practices for control schools. Finally, a back-of-the-envelope calculation suggests that improving management practices would account for a large portion of the program's effect on student performance, at least 50% of the impacts on learning outcomes.

Our paper rigorously evaluates the ability of a specific management program to generate improvements in practices at the school level and adds evidence to the literature on which of them matters the most in this particular context. Nevertheless, more research is needed to disentangle different possible mechanisms and understand what other interventions could trigger improvements in management practices.

Acknowledgements

The authors thank the co-editor and the two anonymous reviewers for the thoughtful comments provided to them during the reviewing process. We thank *Instituto Unibanco* for access to data and Beatriz Garcia, Fabiana Bento, Leonardo Rosa, Samuel Franco, and Vitor Menezes for helping us with the data collection. Also, we thank Cesar Nunes, David Plank, Maria Julia Azevedo, Martin Carnoy, Mirela de Carvalho, Ricardo Paes de Barros, Sergio Firpo, Telma Vinha, Tiago Borba, Ricardo Henriques, and the seminar audiences at *Instituto Unibanco* and Stanford University for helpful comments and discussions. We also thank Bruna Guidetti and Cristiano Martins for their excellent research assistance. The views expressed here are those of the authors alone. All errors and omissions are our own.

Declaration of interest

None.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- A. Abadie, S. Athey, G. W. Imbens, and J. M. Wooldridge. When Should You Adjust Standard Errors for Clustering? *The Quarterly Journal of Economics*, 138(1):1–35, 10 2022. doi:[10.1093/qje/qjac038](https://doi.org/10.1093/qje/qjac038).
- M. Akhtari, D. Moreira, and L. Trucco. Political turnover, bureaucratic turnover, and the quality of public services. *American Economic Review*, 112(2):442–493, 2022. doi:[10.1257/aer.20171867](https://doi.org/10.1257/aer.20171867).
- N. Bloom, R. Lemos, R. Sadun, and J. Van Reenen. Does management matter in schools? *The Economic Journal*, 125(584):647–674, 2015. doi:[10.1111/eoj.12267](https://doi.org/10.1111/eoj.12267).
- A. Böhlmark, E. Grönqvist, and J. Vlachos. The headmaster ritual: The importance of management for school outcomes. *The Scandinavian Journal of Economics*, 118(4): 912–940, 2016. doi:[10.1111/sjoe.12149](https://doi.org/10.1111/sjoe.12149).
- G. F. Branch, E. A. Hanushek, and S. G. Rivkin. Estimating the effect of leaders on public sector productivity: The case of school principals. Working Paper 17803, National Bureau of Economic Research, 2012. URL <http://www.nber.org/papers/w17803>.
- A. C. Cameron and D. L. Miller. A practitioner’s guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372, 2015. doi:[10.3368/jhr.50.2.317](https://doi.org/10.3368/jhr.50.2.317).
- M. Coelli and D. A. Green. Leadership effects: School principals and student outcomes. *Economics of Education Review*, 31(1):92–109, 2012. doi:[10.1016/j.econedurev.2011.09.001](https://doi.org/10.1016/j.econedurev.2011.09.001).
- R. P. de Barros, M. de Carvalho, S. Franco, B. Garcia, R. Henriques, and L. Machado. Assessment of the impact of the *Jovem de Futuro* program on learning. *The World Bank*, 2019. URL <https://documents1.worldbank.org/curated/es/825101561723584640/pdf/Assessment-of-the-Impact-of-the-Jovem-de-Futuro-Program-on-Learning.pdf>.
- C. de Chaisemartin and J. Ramirez-Cuellar. At what level should one cluster standard errors in paired experiments, and in stratified experiments with small strata? Working Paper 27609, National Bureau of Economic Research, 2020. URL <http://www.nber.org/papers/w27609>.

- R. De Hoyos, A. J. Ganimian, and P. A. Holland. Great things come to those who wait: Experimental evidence on performance-management tools and training in public schools in Argentina. 2020. URL <https://static1.squarespace.com/static/5990cfd52994ca797742fae9/t/5f4feecabb6d0e29e6cb2e53/1599073999482/dehoyosetal2020.pdf>.
- R. S. de Mendoca and R. P. de Barros. O Impacto de Gestão sobre o Desempenho Educacional (Brasil). Research Department Publications 3000, Inter-American Development Bank, Research Department, 1997. URL <https://ideas.repec.org/p/idb/wpaper/3000.html>.
- E. Dhuey and J. Smith. How school principals influence student learning. *Empirical Economics*, 54(2):851–882, 2018. doi:[10.1007/s00181-017-1259-9](https://doi.org/10.1007/s00181-017-1259-9).
- A. Di Liberto, F. Schivardi, and G. Sulis. Managerial practices and student performance. *Economic Policy*, 30(84):683–728, 2015. doi:[10.1093/epolic/eiv015](https://doi.org/10.1093/epolic/eiv015).
- W. Dobbie and R. G. Fryer. Getting beneath the veil of effective schools: Evidence from New York City. *American Economic Journal: Applied Economics*, 5(4):28–60, 2013. doi:[10.1257/app.5.4.28](https://doi.org/10.1257/app.5.4.28).
- L. Finamor. Quality of education and human capital decisions: experimental evidence from brazil. Master’s thesis, Fundação Getulio Vargas, 2017. URL <https://bibliotecadigital.fgv.br/dspace/handle/10438/18194>.
- R. G. Fryer. Injecting charter school best practices into traditional public schools: Evidence from field experiments. *The Quarterly Journal of Economics*, 129(3):1355–1407, 2014. doi:[10.1093/qje/qju011](https://doi.org/10.1093/qje/qju011).
- R. G. Fryer. Management and student achievement: Evidence from a randomized field experiment. Working Paper 23437, National Bureau of Economic Research, 2017. URL <http://www.nber.org/papers/w23437>.
- P. Glewwe and K. Muralidharan. Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the Economics of Education*, volume 5, pages 653–743. Elsevier, 2016. doi:[10.1016/B978-0-444-63459-7.00010-5](https://doi.org/10.1016/B978-0-444-63459-7.00010-5).
- J. A. Grissom, A. J. Egalite, and C. A. Lindsay. How principals affect students and schools. Technical report, Wallace Foundation, 2021. URL <https://www.wallacefoundation.org/knowledge-center/Documents/How-Principals-Affect-Students-and-Schools.pdf>.

- R. Henriques, M. de Carvalho, and M. Bittar. *Gestão na educação em larga escala: Jovem de Futuro – de projeto piloto em escolas para uma política de rede pública*. Instituto Unibanco, 2020a. URL <https://www.institutounibanco.org.br/wp-content/uploads/2020/09/Gestao-na-educacao-em-larga-escala.pdf>.
- R. Henriques, M. de Carvalho, and R. P. de Barros. *Avaliação de impacto em educação: a experiência exitosa do programa Jovem de Futuro em parceria com o poder público*. Instituto Unibanco, 2020b. URL https://www.institutounibanco.org.br/wp-content/uploads/2020/05/IU_avaliacao_impacto_educacao-26052020_baixa.pdf.
- R. Henriques, M. de Carvalho, and F. Bento. *Gestão e avanço contínuo em educação: A Teoria da Mudança no Programa Jovem de Futuro*. Instituto Unibanco, 2021. URL <https://observatoriodeeducacao.institutounibanco.org.br/api/assets/716a7066-0be1-427b-8763-ee37dc2eaeaa/>.
- E. L. Horng, D. Klasik, and S. Loeb. Principal’s time use and school effectiveness. *American Journal of Education*, 116(4):491–523, 2010. doi:10.1086/653625.
- INEP. Indicador para mensurar a complexidade da gestão nas escolas a partir dos dados do Censo Escolar da Educação Básica. Technical report, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Brasília, DF, 2014. URL https://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2014/escola_complexidade_gestao/nota_tecnica_indicador_escola_complexidade_gestao.pdf.
- INEP. Censo da Educação Básica 2020 – resumo técnico. Technical report, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Brasília, DF, 2021. URL https://download.inep.gov.br/publicacoes/institucionais/estatisticas_e_indicadores/resumo_tecnico_censo_escolar_2020.pdf.
- Instituto Unibanco. Sobre nós. <https://www.institutounibanco.org.br/o-instituto/sobre-nos/>. Last accessed: 2023-03-30.
- R. Lemos and D. Scur. Developing management: An expanded evaluation tool for developing countries. 2017.
- R. Lemos, K. Muralidharan, and D. Scur. Personnel management and school productivity: Evidence from India. Working Paper 28336, National Bureau of Economic Research, 2021. URL <http://www.nber.org/papers/w28336>.

- A. Miller. Principal turnover and student achievement. *Economics of Education Review*, 36:60–72, 2013. doi:10.1016/j.econedurev.2013.05.004.
- Ministry of Education. Apoio à gestão educacional. <http://portal.mec.gov.br/conaes-comissao-nacional-de-avaliacao-da-educacao-superior/195-secretarias-112877938/seb-educacao-basica-2007048997/18841-apoio-a-gestao-educacional>. Last accessed: 2023-04-12.
- J. G. Miranda. Descontinuidade política, rotatividade de diretores e desempenho dos alunos: efeitos adversos de novos prefeitos na educação municipal. Master’s thesis, Universidade de São Paulo, 2015. URL <https://www.teses.usp.br/teses/disponiveis/96/96131/tde-15092015-105243/en.php>.
- P. Munoz and M. Prem. Managers’ productivity and recruitment in the public sector. 2022. URL <https://ssrn.com/abstract=3480718>.
- K. Muralidharan and A. Singh. Improving public sector management at scale? Experimental evidence on school governance in India. Working Paper 28129, National Bureau of Economic Research, 2020. URL <http://www.nber.org/papers/w28129>.
- G. Muñoz, J. Pascual, and P. Sáez. Seleção de diretores escolares: desafios e possibilidades. Technical report, Instituto Unibanco and Universidad Diego Portales, 2021. URL <https://observatoriodeeducacao.institutounibanco.org.br/cedoc/detalhe/selecao-de-diretores-escolares-desafios-e-possibilidades,03b13c78-1219-45b0-8ae8-3842259278db>.
- OECD. Programme for International Student Assessment – PISA 2015. Technical report, The Organisation for Economic Co-operation and Development, Paris, 2016.
- OECD. Programme for International Student Assessment – PISA 2018. Technical report, The Organisation for Economic Co-operation and Development, Paris, 2019.
- OPE-Sociais. Processo de aleatorização do programa jovem de futuro: Piauí, Pará e Espírito Santo. Technical report, Instituto Unibanco.
- P. Pereda, A. Lucchesi, K. Mendes, and A. Bresolin. Evaluating the impact of the selection process of principal in brazilian public schools. *Nova Economia*, 29:591–621, 2019. doi:10.1590/0103-6351/4242.

- L. S. N. Rosa. Avaliação de impacto do Programa Jovem de Futuro. Master's thesis, Universidade de São Paulo, 2015. URL <https://www.teses.usp.br/teses/disponiveis/12/12138/tde-21082015-112704/en.php>.
- P. A. Tavares. The impact of school management practices on educational performance: Evidence from public schools in São Paulo. *Economics of Education Review*, 48:1–15, 2015. doi:[10.1016/j.econedurev.2015.05.002](https://doi.org/10.1016/j.econedurev.2015.05.002).
- T. Teodorovicz, S. Lazzarini, S. Cabral, and L. Nardi. Can public organizations perform like private firms? The role of heterogeneous resources and practices. *Organization Science*, 2022. doi:[10.1287/orsc.2022.1634](https://doi.org/10.1287/orsc.2022.1634).
- Tesouro Nacional. Aspectos fiscais da educação no Brasil. Technical report, Tesouro Nacional, Brasília, DF, 2018. URL https://sisweb.tesouro.gov.br/apex/f?p=2501:9:::::9:P9_ID_PUBLICACAO:28264.
- T. P. Vinha, A. d. Moraes, and A. Moro. Manual de orientação para a aplicação dos questionários que avaliam o clima escolar. Technical report, Campinas: FE/Unicamp, 2017. URL <http://www.bibliotecadigital.unicamp.br/document/?code=79559&opt=1>.
- E. Walsh and D. Dotter. The impact on student achievement of replacing principals in District of Columbia Public Schools. *Education Finance and Policy*, 15(3):518–542, 2020. doi:[10.1162/edfp_a_00279](https://doi.org/10.1162/edfp_a_00279).

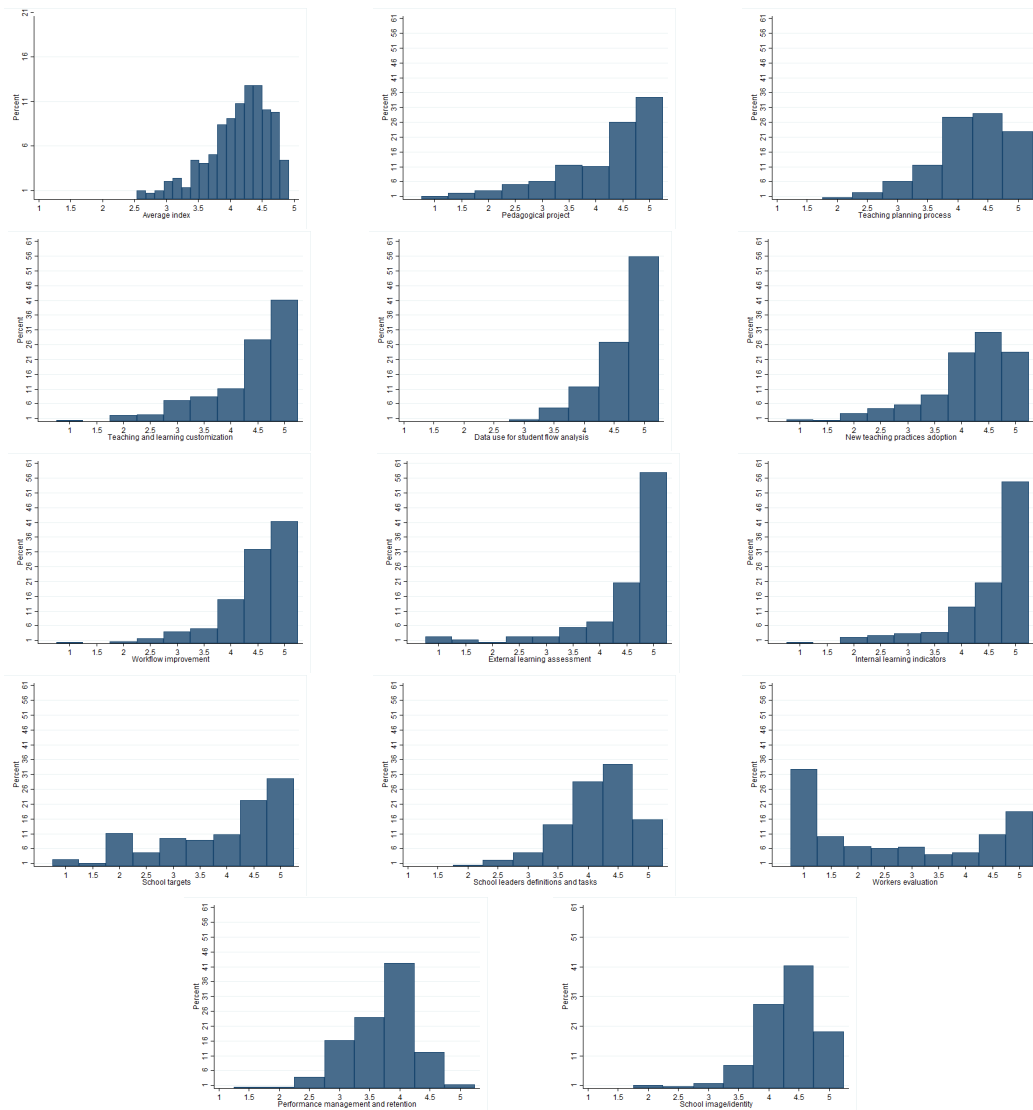
Tables and Figures

Figures

Figure 1: Survey Instrument - Brief Description

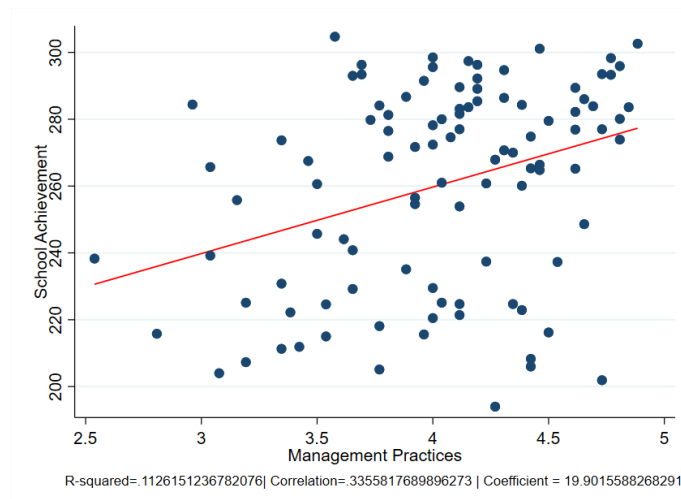
Groups	Practices	Description
Pedagogical Practices	Pedagogical project	Evaluates school elaboration of the pedagogical project and how (and if) it is used to guide decisions.
	Teaching planning process	Evaluates the quality of the pedagogical planning process.
	Teaching and learning customization	Evaluates how principals identify pedagogical strategies to work with different levels of students' learning.
	Data use for student flow analysis	Evaluates how principals deal with absence, repetition, and dropout.
	New teaching practices adoption	Evaluates if principals encourage the improvement of teaching practices and the search for innovative learning strategies.
	Workflow improvement	Evaluates the process of problem-solving adopted by principals.
	School leaders definitions and tasks	Evaluates how principals identify school leaders and assign responsibilities to the management team and teachers.
Human Resources & School Identity	Workers evaluation	Identifies and qualifies how principals evaluate the performance of school professionals.
	Performance management and retention	Analyzes how (and if) principals deal with both great and poor staff performances.
	School image/identity	Assesses whether there is a concern of the manager in creating an identity for the school.
Evaluation Processes	External learning assessment	Evaluates how principals use national and state-level external evaluations to analyze students' learning conditions.
	Internal learning indicators	Evaluates how the school internally evaluates student performance.
	School targets	Evaluates whether there is a management focused on goals for student learning.

Figure 2: Management Survey Instrument: Grading Distribution

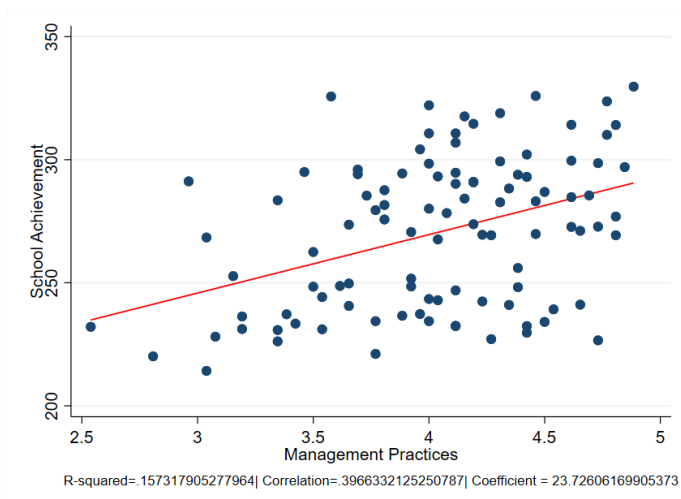


Notes: This figure presents histograms for the average management index and for the thirteen management dimensions.

Figure 3: Management practices and School's achievement



(A) Language



(B) Mathematics

Notes: This figure displays graphics reporting the correlations between the average management index and student's average performance in the state evaluations in 2017.

Figure 4: Management practices and School's environment



(A) Director's Management Survey



(B) Teacher's Management Survey

Notes: This figure displays two main graphics reporting the correlations between the average management index and a measure of school's environment (Vinha et al., 2017). Subfigure A presents the correlation using principal's answers related specifically to management aspects in the school's environment instrument. Subfigure B presents teachers' answers related specifically to management aspects in the school's environment instrument.

Tables

Table 1: Experiment and Sample Composition

	Experiment Design			Sample composition	
	Strata	Details	Schools	Strata	Schools
Pará	42	39 (1T, 1C) 3 (2T, 1C)	78 9	42	80
Espírito Santo	70	11 (3T, 1C) 59 (2T, 1C)	44 177	70	217
Total			308		297

Notes: This table reports the experiment design to evaluate the impact of *Jovem de Futuro* program on learning outcomes in the Brazilian states of Espírito Santo and Pará, and the sample composition of interviewed schools in our survey instrument. Columns ‘Strata’ presents the number of strata in each state. Column ‘Details’ present the composition of each stratum, composed of pairs, trios, or quartets of schools, assigned to the treatment (T) or control (C) groups. Finally, columns ‘Schools’ display the number of schools in each state.

Table 2: Descriptive Statistics

	Total sample		Treatment		Control	
	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.
Management Practices						
Average Index	4.13	0.497	4.19	0.476	4.03	0.520
Pedagogical Practices						
Pedagogical Project	4.14	0.966	4.16	0.974	4.10	0.956
Teaching planning process	4.21	0.655	4.22	0.651	4.19	0.665
Teaching and learning customization	4.35	0.784	4.42	0.685	4.23	0.932
Data use for student flow analysis	4.66	0.453	4.68	0.419	4.63	0.511
New teaching practices adoption	4.14	0.814	4.18	0.818	4.06	0.804
Workflow improvement	4.46	0.658	4.46	0.691	4.46	0.594
School leaders definitions and tasks	4.18	0.597	4.24	0.593	4.06	0.590
Human Resources & School Identity						
Workers evaluation	2.70	1.604	2.83	1.606	2.45	1.579
Performance management and retention	3.72	0.554	3.76	0.564	3.64	0.529
School image/identity	4.31	0.539	4.33	0.529	4.27	0.557
Evaluation Processes						
External learning assessment	4.48	0.893	4.59	0.801	4.30	1.020
Internal learning indicators	4.51	0.747	4.60	0.716	4.35	0.779
School targets	3.88	1.133	4.02	1.042	3.63	1.252
Observations	297		193		104	

Notes: This table presents Management Practices descriptive statistics. This first two columns report the average and standard deviation for the entire sample. Columns three and four report the same statistics for treatment units. Columns five and six report them for control units.

Table 3: Balance Test of Predetermined school characteristics

	(1)	(2)	(3)	(4)
	Full sample	Treatment	Control	Difference
Management Complexity Index	4.882 (1.038)	4.917 (1.062)	4.817 (0.993)	0.098 (0.128)
Elementary school	0.734 (0.443)	0.736 (0.442)	0.731 (0.446)	0.001 (0.059)
High school	0.997 (0.058)	1.000 (0.000)	0.990 (0.098)	0.008 (0.009)
Professional school	0.007 (0.082)	0.005 (0.072)	0.010 (0.098)	-0.003 (0.014)
Number of students (per 1,000)	0.873 (0.394)	0.865 (0.369)	0.888 (0.437)	0.003 (0.052)
Computer lab	0.926 (0.262)	0.933 (0.251)	0.913 (0.283)	0.007 (0.035)
Science lab	0.650 (0.478)	0.668 (0.472)	0.615 (0.489)	0.035 (0.066)
Sports court	0.838 (0.369)	0.824 (0.382)	0.865 (0.343)	-0.046 (0.052)
Internet	0.990 (0.100)	0.984 (0.124)	1.000 (0.000)	-0.017 (0.011)
Rural	0.027 (0.162)	0.031 (0.174)	0.019 (0.138)	0.008 (0.021)
Backup activities	0.010 (0.100)	0.010 (0.102)	0.010 (0.098)	0.005 (0.013)
Share of non-white teachers	0.479 (0.194)	0.477 (0.187)	0.485 (0.206)	0.025 (0.020)
Share of female teachers	0.607 (0.134)	0.611 (0.134)	0.600 (0.135)	-0.009 (0.015)
Observations	297	193	104	297
F-test of joint significance (F-stat)				1.307
F-test, number of observations				297

Notes: This table presents a balance exercise of pre-determined characteristics at the school level between units in control and treatment groups. Column (1) reports the average and standard deviation for the entire sample. Column (2) reports the average and standard deviation for the control group. Column (3) reports the average and standard deviation for treatment group. Column (4) reports the estimate difference between the two groups conditional on strata-state fixed effects. We also report F-test of joint significance in the last rows of the Table. Standard errors clustered at the strata-state level are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table 4: Effect of the *Jovem de Futuro* Program on management practices

Variable	(1) Control group average	(2) Program effect
Average Index	4.028 (0.520)	0.130** (0.063)
Pedagogical Practices	4.245 (0.487)	0.066 (0.058)
Pedagogical Project	4.096 (0.956)	0.076 (0.121)
Teaching planning process	4.188 (0.665)	-0.027 (0.075)
Teaching and learning customization	4.226 (0.932)	0.150 (0.097)
Data use for student flow analysis	4.625 (0.511)	0.041 (0.059)
New teaching practices adoption	4.058 (0.804)	0.081 (0.093)
Workflow improvement	4.462 (0.594)	-0.010 (0.083)
School leaders definitions and tasks	4.063 (0.590)	0.150** (0.070)
Human Resources & School Identity	3.457 (0.701)	0.137 (0.086)
Workers evaluation	2.452 (1.579)	0.259 (0.185)
Performance management and retention	3.644 (0.529)	0.098 (0.073)
School image/identity	4.274 (0.557)	0.054 (0.069)
Evaluation Processes	4.091 (0.803)	0.274*** (0.096)
External learning assessment	4.298 (1.020)	0.247** (0.113)
Internal learning indicators	4.346 (0.779)	0.217** (0.094)
School targets	3.630 (1.252)	0.359** (0.152)
Observations	104	297

Notes: This table reports our main estimates of the effect of the *Jovem do Futuro* Program on management practices. Column (1) reports the average and standard deviation for the control group. Column (2) reports the average and standard deviation for treatment units (program effect). The Average Effect Index shows the Average Index based on all thirteen analyzed practices. The first panel describes the program effect estimates for the seven Pedagogical Practices. The second panel describes the program effect estimates for the three Human Resources & School Identity dimensions. Finally, the third panel describes the program effect estimates for the three Evaluation Processes dimensions. Standard errors clustered at the strata-state level are reported in parentheses. ***, **, and * indicate the coefficients significantly different from zero at the 99, 95, and 90 percent confidence levels.

Table 5: Effect of the *Jovem de Futuro* Program on students' learning in 2017 - State evaluations

Variable	(1) Control group average	(2) Program effect
Original impact evaluation		
Management Practices - Average Index	3.981 (0.483)	0.129* (0.071)
Average Portuguese test scores	249.991 (30.623)	5.024*** (1.874)
Average Math test scores	260.256 (29.058)	5.237** (2.083)
Observations	68	183
All Schools		
Management Practices - Average Index	4.036 (0.512)	0.123* (0.063)
Average Portuguese test scores	260.416 (30.352)	2.989** (1.478)
Average Math test scores	270.408 (30.615)	2.695 (1.736)
Observations	102	294

Notes: This table reports our main estimates of the effect of the *Jovem do Futuro* Program on learning outcomes: average Portuguese and Math test scores in state evaluations. For comparison purposes, we present the impacts on the average management index for the subsamples of schools. Panel 'Original impact evaluation' reports the results for the sample of schools that participated in the original impact evaluation of the program by *Instituto Unibanco* (183 out of 297). Panel 'All Schools' reports the results for all schools for which we have available information on state evaluations' performance (294 out of 297). Column (1) reports the average and standard deviation for the control group. Column (2) reports the average and standard deviation for treatment units controlling for strata-state (program effect). Standard errors clustered at the strata-state level are reported in parentheses. ***, ** and * indicate the coefficients significantly different from zero at the 99, 95, and 90 percent confidence levels.

A Appendix Tables

Table A.1: Kappa-statistic measure of interrater agreement

	Kappa	Z-score	P-value
1	0.188	5.900	0.000
2	0.208	5.847	0.000
3	0.203	6.216	0.000
4	0.185	4.242	0.000
5	0.271	8.053	0.000
6	0.070	1.831	0.034
7	0.245	7.249	0.000
8	0.292	8.156	0.000
9	0.322	10.922	0.000
10	0.077	2.118	0.017
11	0.475	16.275	0.000
12	0.209	5.609	0.000
13	0.086	2.051	0.020

Notes: This table reports the kappa test statistic by management practice (13 dimensions), a measure of interrater agreement that compares the agreement between two nonunique raters with what would be expected by chance. In the first column, we present the kappa statistic. In the second column is the z-score, and in the third is the p-value associated with the z-score.

Table A.2: Attrition Test

Variable	(1) In the sample	(2) Out of the sample	(3) Difference
Management Complexity Index	4.882 (1.038)	5.182 (0.751)	0.300 (0.316)
Elementary school	0.734 (0.443)	0.636 (0.505)	-0.098 (0.137)
High school	0.997 (0.058)	1.000 (0.000)	0.003 (0.018)
Professional school	0.007 (0.082)	0.000 (0.000)	-0.007 (0.025)
Number of students	873.047 (393.739)	1,107.727 (450.837)	234.680* (121.508)
Computer lab	0.926 (0.262)	0.727 (0.467)	-0.199** (0.083)
Science lab	0.650 (0.478)	0.545 (0.522)	-0.104 (0.147)
Sports court	0.838 (0.369)	0.909 (0.302)	0.071 (0.113)
Internet	0.990 (0.100)	0.909 (0.302)	-0.081** (0.035)
Rural	0.027 (0.162)	0.000 (0.000)	-0.027 (0.049)
Backup activities	0.010 (0.100)	0.000 (0.000)	-0.010 (0.030)
Share of non-white teachers	0.479 (0.194)	0.623 (0.078)	0.143** (0.059)
Share of female teachers	0.607 (0.134)	0.563 (0.117)	-0.044 (0.041)
Observations	297	11	308
F-test of joint significance (F-stat)			0.654
F-test, number of observations			308

Notes: This table presents statistics of pre-determined characteristics at the school level for schools in the sample and out of the sample. Column (1) reports the average and standard deviation for the 297 schools in the sample. Column (2) reports the average and standard deviation for the schools for which we could not interview principals and, therefore, did not obtain data on management practices. Column (3) reports the unconditional difference between the two groups. We also report the F-test of joint significance in the last rows of the table. Robust standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent critical level.

Table A.3: Minimum detectable effect sizes

	Power		
	70	80	90
Average Index	0.160	0.180	0.208
Pedagogical Practices	0.149	0.168	0.195
Pedagogical Project	0.290	0.327	0.378
Teaching planning process	0.209	0.236	0.273
Teaching and learning customization	0.286	0.323	0.373
Data use for student flow analysis	0.156	0.176	0.203
New teaching practices adoption	0.248	0.279	0.323
Workflow improvement	0.180	0.203	0.234
School leaders definitions and tasks	0.180	0.203	0.234
RH	0.215	0.243	0.281
Workers evaluation	0.489	0.552	0.639
Performance management and retention	0.160	0.180	0.209
School image/identity	0.168	0.190	0.220
Evaluation	0.245	0.276	0.320
External learning assessment	0.315	0.355	0.410
Internal learning indicators	0.239	0.269	0.312
School targets	0.378	0.427	0.494

Notes: This tables presents the minimum detectable effect (MDE) sizes for each response variable we analyze (the overall management index, the subgroups of dimensions and each dimension separately). The MDE represents the smallest parameter value for a given sample size for which we can reject the null hypothesis in a two-sided test at a 10% significance level. Each column represents a different power level: 70%, 80% and 90%.

Table A.4: Heterogeneous effects of the program by Brazilian State

Variable	Espírito Santo		Pará	
	(1) Control group average	(2) Program effect	(3) Control group average	(4) Program effect
Average Index	4.171 (0.449)	0.095 (0.074)	3.757 (0.544)	0.175 (0.121)
Pedagogical Practices	4.383 (0.400)	0.011 (0.069)	3.984 (0.534)	0.155 (0.109)
Pedagogical Project	4.059 (0.987)	0.051 (0.159)	4.167 (0.902)	0.164 (0.162)
Teaching planning process	4.368 (0.501)	0.035 (0.083)	3.847 (0.800)	-0.256* (0.151)
Teaching and learning customization	4.581 (0.583)	-0.087 (0.087)	3.556 (1.094)	0.603** (0.240)
Data use for student flow analysis	4.750 (0.436)	-0.025 (0.066)	4.389 (0.562)	0.158 (0.123)
New teaching practices adoption	4.221 (0.725)	0.044 (0.103)	3.750 (0.866)	0.134 (0.194)
Workflow improvement	4.537 (0.506)	-0.056 (0.094)	4.319 (0.719)	0.067 (0.168)
School leaders definitions and tasks	4.169 (0.564)	0.112 (0.082)	3.861 (0.593)	0.218 (0.135)
Human Resources & School Identity	3.608 (0.713)	0.143 (0.105)	3.171 (0.586)	0.099 (0.159)
Workers evaluation	2.816 (1.648)	0.288 (0.232)	1.764 (1.180)	0.118 (0.308)
Performance management and retention	3.706 (0.483)	0.085 (0.082)	3.528 (0.597)	0.141 (0.152)
School image/identity	4.301 (0.567)	0.054 (0.084)	4.222 (0.540)	0.039 (0.122)
Evaluation Processes	4.238 (0.700)	0.245** (0.109)	3.815 (0.916)	0.297 (0.202)
External learning assessment	4.426 (0.963)	0.278** (0.130)	4.056 (1.094)	0.123 (0.242)
Internal learning indicators	4.529 (0.566)	0.125 (0.099)	4.000 (0.993)	0.395* (0.212)
School targets	3.757 (1.167)	0.333* (0.180)	3.389 (1.384)	0.374 (0.290)
Observations	217		80	

Notes: This table reports the heterogeneous effects of the *Jovem de futuro* program on two different states (Espírito Santo and Pará). Column (1) reports the average score and standard deviation for Espírito Santo's control group. Column (2) reports the program's effect on Espírito Santo. Column (3) reports the average and standard deviation for the Pará's control group. Finally, column (4) reports the program's effect on Pará. The row Average Effect Index presents the Average Index based on all thirteen analyzed practices. The first panel describes the program's heterogeneous effects estimates for the seven Pedagogical Practices. Second panel describes program's heterogeneous effects estimates for the three Human Resources & School Identity dimensions. Finally, the third panel describes the program's heterogeneous effect estimates for the three Evaluation Process dimensions. Standard errors clustered at the strata-state level are reported in parentheses. ***, **, and * indicate the coefficients are statistically different from zero at the 99, 95, and 90 percent confidence levels.

Table A.5: Robustness check: Interviewer Fixed Effect

Variable	(1) Control group average	(2) Program effect
Pedagogical Project	4.097 (1.122)	0.094 (0.115)
Teaching planning process	4.199 (0.811)	-0.008 (0.071)
Teaching and learning customization	4.238 (1.030)	0.174* (0.095)
Data use for student flow analysis	4.621 (0.627)	0.043 (0.054)
New teaching practices adoption	4.058 (0.887)	0.085 (0.100)
Workflow improvement	4.485 (0.795)	-0.015 (0.072)
School leaders definitions and tasks	4.073 (0.814)	0.106 (0.069)
Workers evaluation	2.447 (1.660)	0.275 (0.186)
Performance management and retention	3.646 (0.659)	0.089 (0.067)
School image/identity	4.291 (0.672)	0.079 (0.064)
External learning assessment	4.330 (1.049)	0.234** (0.111)
Internal learning indicators	4.364 (0.910)	0.210** (0.086)
School targets	3.655 (1.355)	0.373*** (0.143)
Observations	206	591
Schools	103	296

Notes: This table reports the effects of the *Jovem de futuro* program on management practices. Each school was rated by two different evaluators and the unit of analysis is the school-score level (2 observations per school). In all regressions, we control for interviewer-order and strata-state fixed effects. We drop 1 school with missing information on the interviewers' identification. Column (1) reports the average and standard deviation for the control group. Column (2) reports the average and standard deviation for treatment units (program effect). The dependent variable are the thirteen analyzed practices. Standard errors clustered at the school level are reported in parentheses. ***, ** and * indicate the coefficients significant different from zero at the 99, 95 and 90 percent confidence level.

Table A.6: Learning subsamples: Effect of the *Jovem de Futuro* Program on management practices

Variable	Original Impact Evaluation		All schools	
	(1) Control group average	(2) Program effect	(3) Control group average	(4) Program effect
Average Index	3.981 (0.483)	0.129* (0.071)	4.036 (0.512)	0.123* (0.063)
Pedagogical	4.202 (0.469)	0.076 (0.067)	4.249 (0.483)	0.063 (0.059)
Pedagogical Project	4.096 (0.923)	0.151 (0.138)	4.088 (0.961)	0.079 (0.122)
Teaching planning process	4.110 (0.657)	-0.039 (0.093)	4.196 (0.661)	-0.027 (0.076)
Teaching and learning customization	4.074 (1.041)	0.235* (0.133)	4.230 (0.930)	0.149 (0.098)
Data use for student flow analysis	4.603 (0.500)	0.044 (0.070)	4.618 (0.513)	0.046 (0.060)
New teaching practices adoption	4.015 (0.801)	-0.002 (0.121)	4.054 (0.811)	0.082 (0.094)
Workflow improvement	4.529 (0.510)	-0.067 (0.088)	4.480 (0.545)	-0.030 (0.081)
School leaders definitions and tasks	3.985 (0.579)	0.211** (0.088)	4.074 (0.586)	0.138* (0.070)
Human Resources & School Identity	3.363 (0.625)	0.130 (0.104)	3.451 (0.698)	0.144 (0.087)
Workers evaluation	2.074 (1.407)	0.426* (0.230)	2.422 (1.573)	0.300 (0.187)
Performance management and retention	3.654 (0.520)	0.047 (0.090)	3.642 (0.533)	0.096 (0.073)
School image/identity	4.360 (0.495)	-0.082 (0.075)	4.289 (0.533)	0.037 (0.067)
Evaluation Processes	4.086 (0.777)	0.249** (0.114)	4.124 (0.764)	0.245*** (0.093)
External learning assessment	4.265 (0.975)	0.235* (0.137)	4.363 (0.918)	0.181* (0.103)
Internal learning indicators	4.338 (0.799)	0.230** (0.113)	4.368 (0.764)	0.198** (0.093)
School targets	3.654 (1.256)	0.284 (0.187)	3.642 (1.230)	0.355** (0.152)
Observations	68	183	102	294

Notes: This table reports our main estimates of the effect of the *Jovem do Futuro* Program on management practices for the subsamples of schools for which we evaluate the impact on learning outcomes. Columns 'Original impact evaluation' reports the results for the sample of schools that participated in the original impact evaluation of the program by *Instituto Unibanco* (183 out of 297). Columns 'All Schools' reports the results for all schools for which we have available information on state evaluations' performance (294 out of 297). The Average Effect Index shows the Average Index based on all thirteen analyzed practices. The first panel describes the program effect estimates for the seven Pedagogical Practices. The second panel describes the program effect estimates for the three Human Resources & School Identity dimensions. The third panel describes the program effect estimates for the three Evaluation Processes dimensions. Column (1) reports the average score and standard deviation for 'Original impact evaluation' control group. Column (2) reports the program's effect in 'Original impact evaluation'. Column (3) reports the average and standard deviation for the 'All Schools' control group. Finally, column (4) reports the program's effect on 'All Schools'. Standard errors clustered at the strata-state level are reported in parentheses. ***, ** and * indicate the coefficients significantly different from zero at the 99, 95, and 90 percent confidence levels.

Online Appendix (For Online Publication)

Evaluating the impact of a principals' professional development program on school management practices: Evidence from Brazil

Bruna Borges, Gabriel Leite, Ricardo Madeira, Luís Meloni.

B *Jovem de Futuro*: Theory of Change

This appendix visually provides the program's Theory of Change presented in [Henriques et al. \(2021\)](#).

Figure B.1: Theory of Change - *Jovem de Futuro* program

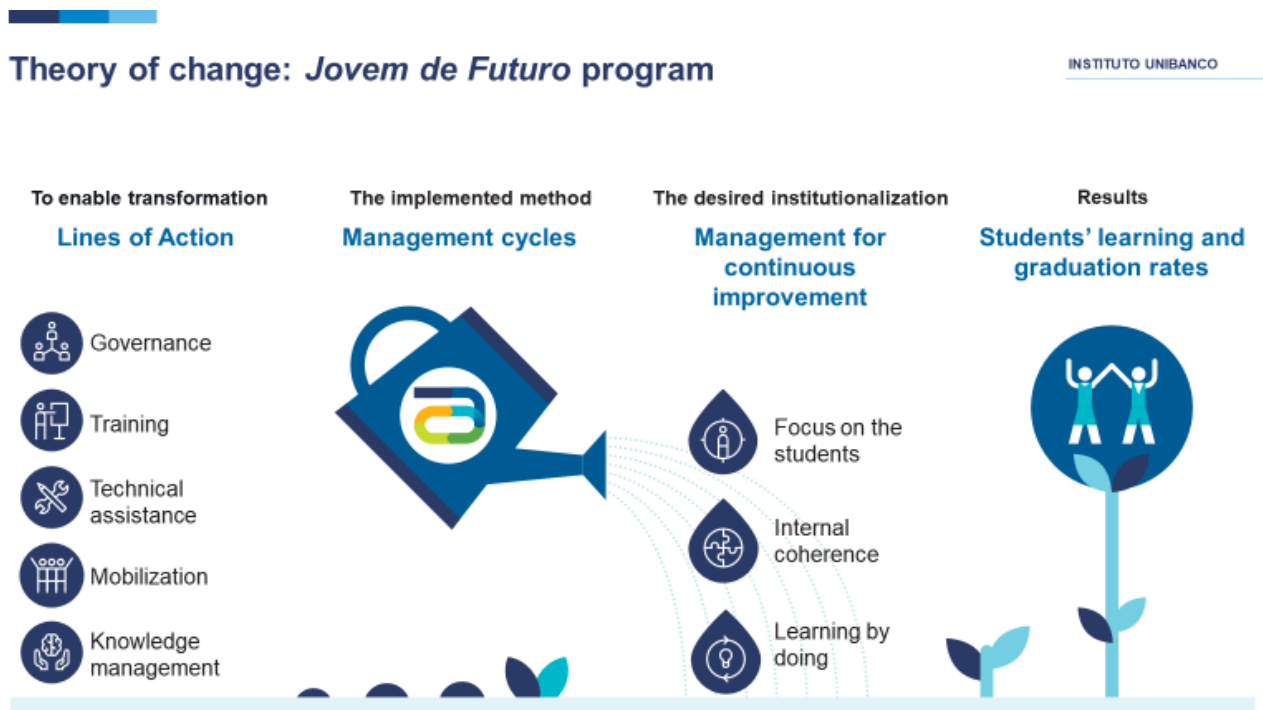


Figure B.2: Effect layers - *Jovem de Futuro* program

Effect layers

INSTITUTO UNIBANCO



C The Management Practices Index

This appendix details which principals' practices are evaluated within each of the thirteen management dimensions. The complete instrument is available upon request.

Dimension 1 – Pedagogical project This dimension evaluates the school elaboration of the Pedagogical Project (PP) and how (and if) it is used to guide school decisions. Firstly, we evaluate if the school effectively has a pedagogical project and when it was established. Then, it analyzes if the document is revised periodically and if the school community is involved in the elaboration and revision processes. Most importantly, this dimension evaluates if the management team makes the pedagogical project publicly available and how (and if) the document is revisited to ensure that the objectives defined by the school are followed.

- Grid:

1. The importance of the Pedagogical Project is unclear to the school management team, and there is no interest in revising it. The management team regards the Pedagogical Project in a bureaucratic sense, merely to follow State or National educational rules.
2. There is some clarity on the importance of the Pedagogical Project by the school management team. The Pedagogical Project is revised sometimes, but not regularly or with a predefined frequency. Those revisions do not democratically involve the whole school community. There is no preoccupation with disseminating the PP and/or making it accessible to others, nor the interest in regularly revisiting the PP over the year to monitor if the school objectives are being accomplished.
3. There is clarity about the importance of the Pedagogical Project by the school management team, which is revised yearly. However, those revisions do not democratically involve the whole school community. There is no preoccupation with disseminating the PP and/or making it accessible to others, nor the interest in regularly revisiting the PP over the year to monitor if the school objectives are being accomplished.
4. There is clarity about the importance of the Pedagogical Project by the school management team, which is revised yearly. All school community is invited

to construct, revise and evaluate the Pedagogical Project. However, there is no preoccupation with disseminating the PP and/or making it accessible to others, nor an interest in regularly revisiting the PP over the year to monitor if the school objectives are being accomplished.

5. There is clarity about the importance of the Pedagogical Project by the school management team, which is revised yearly. All school community is invited to construct, revise and evaluate the Pedagogical Project. The management team and school community regularly revisit the PP in the parents' meetings, teaching planning meetings, and other educational meetings to ensure the school objectives are accomplished.

Dimension 2 – Teaching planning process This dimension evaluates the quality of the pedagogical planning process. Firstly, it evaluates if pedagogical meetings with teachers are regularly carried out. Then, it evaluates how principals organize those meetings and what kind of educational (and training) materials are used (if used) to substantiate discussions and decisions. In particular, the dimension evaluates if principals know their role in this process or if they just follow external guidelines (from State Educational Office, for example). Most importantly, it analyzes if principals focus on pedagogical matters in those meetings.

- Grid:

1. Teaching planning meetings are not held regularly. The school does not adopt materials to support study/discussions at those moments. On the part of the management team, there is no clarity about its role in the pedagogical planning or the importance of this process to the school.
2. Teaching planning meetings are not held regularly. The school does not adopt materials to support study/discussions at those moments. On the part of the management team, there is some clarity about its role in the pedagogical planning process.
3. Teaching planning meetings are held regularly. However, these meetings are not structured clearly, starting from the project elaborated by the management team and/or the pedagogical coordination. The school does not adopt any materials to support study/discussions at those moments, following only the guidelines of the Educational State Office. On the part of the management team, there is clarity about its role in the pedagogical planning process.

4. Teaching planning meetings are held regularly. These meetings are well structured, starting from the project elaborated by the management team and/or by the pedagogical coordination, and have a pedagogical emphasis. The school does not adopt any materials to support study/discussions at those moments, following only the guidelines of the Educational State Office. On the part of the management team, there is clarity about its role in the pedagogical planning process.
5. Teaching planning meetings are held regularly. These meetings are well structured, starting from the project elaborated by the management team and/or by the pedagogical coordination, and have a pedagogical emphasis. The school adopts formative materials to support study/discussions at those moments. On the part of the management team, there is clarity about its role in the pedagogical planning process.

Dimension 3 – Teaching and learning customization This dimension evaluates how principals identify pedagogical strategies to deal with different levels of students' learning. In particular, it evaluates if (and how) principals diagnose those differences, when the diagnostic process starts (at the beginning of the school year, for example), what kind of interventions are designed to deal with the learning gap, and if the results of these actions are monitored throughout the year.

- Grid:

1. The school does not recognize students in a given grade with different learning levels.
2. The school recognizes students in a given grade with different learning levels. However, the school has not a clear mechanism to identify them. The identification of such students occurs over the year according to their performance. There is no interest in elaborating interventions to help those students with learning difficulties beyond the extra and compulsory classes.
3. The school recognizes students in a given grade with different learning levels and tries to identify them from the beginning of the year with diagnostic activities. However, little is done to help those students with learning difficulties beyond the extra and compulsory classes.

4. The school recognizes students in a given grade with different learning levels and tries to identify them from the beginning of the year with diagnostic activities. The school also promotes continuous interventions to help students with learning difficulties. However, there are no regular meetings to discuss the students' improvements or difficulties.
5. The school recognizes students in a given grade with different learning levels and tries to identify them from the beginning of the year with diagnostic activities. The school also promotes continuous interventions to help students with learning difficulties. In addition, there are meetings regularly to discuss the students' improvements or difficulties.

Dimension 4 – Data use for student flow analysis This dimension evaluates how principals deal with absence, repetition, and dropout. First, we evaluate if principals acknowledge the existence of those problems in their schools and how (and if) they identify those issues. Then, we evaluate if (and how) principals collect and organize data on absence, repetition, and dropout, and which actions are developed in the face of this information to mitigate those problems. Most importantly, this dimension analyzes what (and if) individualized and systemic actions (with parents' and teachers' engagement) the principals take to retain students with potential repetition and dropout.

- Grid:

1. The school does not recognize the existence of repetition and dropouts.
2. The school recognizes the existence of repetition and dropouts. However, the school does not have not data on this issue. Also, the school does not develop any action to recover students with a potential dropout/repetition profile.
3. The school recognizes the existence of repetition and dropouts. However, the school does not have organized data on this issue and has no effective management in this aspect. That is, the school does not develop any action to recover students with a potential dropout/repetition profile.
4. The school recognizes the existence of repetition and dropouts. The school has organized data on this issue but lacks effective management. The school does not develop individual or systemic actions to recover students identified with a potential dropout/repetition profile.

5. The school recognizes the existence of repetition and dropouts. The school has organized data on this issue and effectively manages this aspect. The school develops individual or systemic (involving teachers, students, and family) actions to recover students identified with a potential dropout/repetition profile.

Dimension 5 – New teaching practices adoption This dimension evaluates if principals encourage the improvement of teaching practices and the search for innovative learning strategies. In particular, it analyzes if principals monitor the class pedagogical activities and support collaboration and sharing of new practices among teachers and staff members. Most importantly, it evaluates if (and how) principals periodically test and embrace innovative practices and evaluate their efficacy in student learning.

- Grid:

1. Teacher pedagogical activities are little encouraged and monitored by the management team.
2. Teacher pedagogical activities are monitored by the management team, but there is no interest in different and innovative learning methodologies.
3. Teacher pedagogical activities are encouraged and monitored by the management team. There is the adoption of different and innovative learning methodologies, but it is punctual and rarely shared between the teachers. Moreover, the management team does not evaluate the efficacy of those methodologies.
4. Teacher pedagogical activities are encouraged and monitored by the management team. There is the adoption of different and innovative learning methodologies. Teachers frequently share experiences or different learning proposals. The managers (principal and/or pedagogical coordinator) are proactive and provide support for innovations, evaluating the efficacy and encouraging collaboration and sharing of the work between teacher and staff members, acting to engage others in a culture of transformation.
5. Teacher pedagogical activities are encouraged and monitored by the management team. There is the adoption of different and innovative learning methodologies. Teachers frequently share experiences or different learning proposals. The managers (principal and/or pedagogical coordinator) are proactive and provide support for innovations, evaluating the efficacy and encouraging collaboration and sharing of the work between teacher and staff members, acting to

engage others in a culture of transformation. This process occurs regularly in pedagogical meetings.

Dimension 6 – Workflow improvement This dimension evaluates principals' problem-solving process. In particular, it analyzes if principals deal with the school problems only occasionally, through improvisation, or in a structured and proactive manner. Most importantly, we evaluate if organized interventions to tackle the problems are organically embedded in the school management and if the problem-solving process generates learning techniques to address other potential issues. Also, this dimension evaluates if principals foster the engagement of school members in the solution and if (and how) they share the learned lessons with the community.

- Grid:

1. The problems exposure and interventions (to the school, students, teachers, and staff) are not structured and punctual and do not have any further implications for improving the school management process. It is always reactive, and the learning process is restricted to the group or the person who solved the problem.
2. The problems exposure and interventions (to the school, students, teachers, and staff) are not structured and punctual and do not have further implications for improving the school management process. The learning process flows informally between school actors.
3. The problems exposure and interventions (to the school, students, teachers, and staff) are structured, and problem-solving is considered an organic management aspect. However, the process does not involve the school community in resolving more complex problems.
4. The problems exposure and interventions (to the school, students, teachers, and staff) are structured, and problem-solving is considered an organic management aspect. The learning process flows formally, consistent with improvements in school management. However, the process does not involve the school community in resolving more complex problems.
5. The problems exposure and interventions (to the school, students, teacher, and staff) are structured, and problem-solving is considered an organic management aspect. The learning process flows formally and clearly, consistent with improvements in the school management. The school community's involvement

is seen as an integral part of resolving more complex problems. Such problems are shared to engage a cooperative behavior.

Dimension 7 – External learning assessment This dimension evaluates how principals use national and state-level external evaluations to analyze students' learning conditions. In particular, our instrument first evaluates if principals know about standardized tests and how they assess the school's performance. Then, the instrument evaluates what the principal does with the information at hand, what pedagogical actions are triggered in the face of the school performance (if any), and how (and if) those results are shared with the school community.

- Grid:

1. The school does not have/does not know any external evaluation process or does not use external indicators.
2. The school has an external evaluation process and uses the indicators to elaborate some actions, not yet organized or systematized. Results are not disclosed to the school community.
3. The school has an external evaluation process and uses the indicators to elaborate organized and systematized actions, but the results are partially disclosed to the school community. The reflective and diagnostic process does not involve all professionals (teachers and staff members) in a democratic/inclusive way.
4. The school has an external evaluation process and uses the indicators to elaborate organized and systematized actions, and the results are disclosed clearly and objectively to the school community. However, the reflective and diagnostic process does not involve all professionals (teachers and staff members) in a democratic/inclusive way.
5. The school has an external evaluation process and uses the indicators to elaborate organized and systematized actions, and the results are disclosed clearly and objectively to the school community. The reflective and diagnostic process involves and includes all the community (parents, students, teachers, and staff members) to elaborate a school plan.

Dimension 8 – Internal learning indicators This dimension evaluates how the school internally evaluates student performance. Specifically, our instrument first captures the presence of internal mechanisms to assess student learning and its frequency. Then, most importantly, we evaluate if principals contrast these indicators with the external assessments, the actions taken in the face of students’ performance, and how they share these results with the school community.

- Grid:

1. The school does not have its own internal mechanism for the performance evaluation of students and does not use data from external indicators.
2. The school has its own internal performance evaluation mechanism for students but it is held without regularity. The school does not use external indicators to evaluate its performance or does so in a poor manner without comparing them with internal data. Results from these evaluations are restricted to the management team and some teachers.
3. The school has its own internal mechanism for the performance evaluation of students, and it is held regularly. However, the school does not use external indicators to evaluate its performance or does poorly without comparing them with internal data to elaborate an action plan. Results from these evaluations are shared in a broad, clear, and objective way with school professionals to engage them in the improvement management process.
4. The school has its own internal mechanism for the performance evaluation of students, and it is held regularly. The school monitors the external indicators to evaluate its performance but does not compare them with internal data to elaborate an action plan. Results from these evaluations are shared in a broad, clear, and objective way with school professionals to engage them in the improvement management process.
5. The school has its own internal mechanism for the performance evaluation of students, and it is held regularly. The school monitors the external indicators to evaluate its performance and compares them with internal data to elaborate an action plan. Results from these evaluations are shared in a broad, clear, and objective way with the whole school community (parents, students, teachers, managers, and staff members) to engage them in the improvement management process.

Dimension 9 – School targets This dimension evaluates the relevance of establishing targets for student learning. Firstly, our instrument evaluates if principals define internal targets for their schools and how they use the externally established targets (by the Regional or State Educational Office, for example). Most importantly, we ask who is involved in elaborating those targets, what actions (if any) are taken to pursue and achieve them, and the frequency (if any) of target monitoring and revision.

- Grid:

1. The school does not have its own targets (or does not know external targets) or knows only performance targets externally established (by the State Educational Office or Ministry of Education) without any further consideration about their meaning.
2. The school knows external targets established but does not elaborate any specific internal targets. There are no considerations about their meaning. The school community has no involvement in the discussions and reflections about the targets.
3. The school knows the external targets and elaborates internal targets only for performance results. The construction of the targets intends to discuss their meaning, but actions are taken without the effective participation of the school community. The accomplishment of those targets is monitored sporadically.
4. The school knows the external targets and elaborates internal targets. The construction of the targets intends to discuss their meaning, and actions are elaborated with the effective participation of the school community. The accomplishment of those targets is monitored sporadically.
5. The school knows the external targets and elaborates internal targets to the school and its employees (management team, teachers, and staff). Targets include performance improvement but are broader in dimensions and relevance to the institution and well-being of all (e.g., improve the student-teacher relationship, decrease violence, improve infrastructure, etc.). The construction of the targets intends to discuss their meaning, and actions are elaborated with the effective participation of the school community. The accomplishment of those targets is monitored regularly and systematically, and revised to re-elaborate the action plan in case of non-accomplishment.

Dimension 10 – School leaders’ definitions and tasks This dimension evaluates how principals identify school leaders and attribute responsibilities to the management team and teachers. Our instrument evaluates how clearly the principals define the staff members’ roles and how principals address changes in attributions from time to time. We also evaluate what kind of tasks are mostly performed by the principals (bureaucratic vs. pedagogical).

- Grid:

1. School professionals’ roles are not clearly defined. The principal’s main attributions are bureaucratic, without further implications for students learning.
2. School professionals’ roles are not clearly defined. The principal’s main attributions are bureaucratic but have some implications for students learning.
3. School professionals’ roles are not clearly defined or are defined only by the regulations of the State Educational Office, but they are not developed by school professionals or revised by the management team.
4. Principals understand their role in the pedagogical and administrative management of the school and their association with student learning. Attributions, responsibilities, and skills desired for teachers and other leaders are clearly defined. These definitions are based on factors aimed at improving the pedagogical work and the well-being of all school professionals, as well as the organization and better functioning of the school.
5. Principals understand their role in the pedagogical and administrative management of the school and their association with student learning. Attributions, responsibilities, and skills desired for teachers and other leaders are clearly defined. These definitions are based on factors aimed at improving the pedagogical work and the well-being of all school professionals, as well as the organization and better functioning of the school. Attributions and responsibilities are developed with school professionals and revised whenever required.

Dimension 11 – Workers evaluation This dimension identifies how principals evaluate the performance of school members (staff and teachers). Firstly, it analyzes if principals have a performance evaluation system. Then, it investigates if this system is informal (or externally established by the State Educational Office, for example), i.e.,

occasionally applied without further consequences, or formally built to give periodical feedback. Most importantly, it also evaluates how (and if) principals work with their employees individually and collectively to improve overall performance.

- Grid:

1. Professionals are not evaluated because the school does not have its own evaluation system to assess the team's performance.
2. There is an informal evaluation system or an evaluation system elaborated by the State Educational Office (or Directory), which is sporadically applied by the management team. Some professionals are praised and recognized for their performance as general feedback without further developments or revisions.
3. There is a formal evaluation system or one improved upon the system of the State Educational Office (or Directory), which is regularly applied by the management team. Some professionals are praised and recognized for their performance. Each professional receives individual feedback but without further developments or revisions.
4. There is a formal evaluation system or one improved upon the system of the State Educational Office (or Directory), which is regularly applied by the management team. Some professionals are praised and recognized for their performance. Each professional receives individual feedback. Aspects that should be improved are not individually developed but collectively discussed.
5. There is a formal evaluation system or one improved upon the system of the State Educational Office (or Directory), which is regularly applied by the management team. Some professionals are praised and recognized for their performance. Each professional receives individual feedback. Aspects that should be improved are individually developed and collectively discussed in a framework of collective and respectful progress. Results are analyzed to revise attributions or adjust profiles whenever necessary.

Dimension 12 – Performance management and retention This dimension analyzes how (and if) principals deal with both great and poor staff performances. In particular, it evaluates what actions the principals take to correct poor performance/behavior or value the good ones. It also evaluates how systemic and diverse these strategies are and how principals assess the efficacy of those actions. Most importantly, this dimension analyses if principals make every possible effort to retain their best professionals.

- Grid:

1. Bad behavior is not corrected or corrected inconsistently. The school leadership rarely reprimands individuals who have compromised performance. Good professionals are not appreciated, and the school can hardly maintain them.
2. Bad behavior is reactively or impulsively corrected without any monitoring by the management team regarding improving actions. Good professionals are little appreciated, and the school has no mechanism to maintain those who want to leave.
3. Bad behavior is corrected with a limited range of actions such as warnings, reprimands, individual training, or professional relocation. Good professionals are occasionally and superficially appreciated. However, the school has difficulties maintaining excellent professionals.
4. A frequent bad behavior is corrected with targeted interventions and multiple methods (such as work observation, orientation meetings, dialog groups, and professional support). The school monitors the efficacy of the corrective actions. Good professionals are appreciated through a systematic process of good work recognition, but the management team invests little effort in removals cases.
5. A frequent bad behavior is corrected with targeted interventions and multiple methods (such as work observation, orientation meetings, dialog groups, and professionals' support). The school monitors the efficacy of the corrective actions. Good professionals are appreciated through a systematic process of good work recognition. The management team invests huge efforts and considers all the possible ways to maintain excellent professionals, even resorting to superior instances to seek support.

Dimension 13 – School Image/identity This dimension evaluates if principals care about creating an identity for the school. First, we evaluate if principals know how the community perceives their school identity. Then, it investigates if principals recognize the pros and cons of working in their schools. Most importantly, we evaluate what kind of practices (if any) they held to improve the school identity and their actions to disseminate the school value to the community. It also analyses if those actions are organically planned as part of the school objectives.

- Grid:

1. On the part of the management team, there is no preoccupation with creating a school's identity to make the institution recognized and appreciated by the whole school community.
2. On the part of the management team, there is an interest in creating a school's identity to make the institution recognized and appreciated by the whole school community. Still, no actions are taken in this regard. The management team cannot acknowledge the pros and cons of teaching in their school and does not clearly understand how the school community sees the institution.
3. On the part of the management team, there is an interest in creating a school's identity to make the institution recognized and appreciated by the whole school community. Still, few actions have been taken in this regard. The management team cannot acknowledge the pros and cons of teaching in their school and does not clearly understand how the school community sees the institution.
4. On the part of the management team, there is an interest in creating a school's identity to make the institution recognized and appreciated by the whole school community. Actions are taken to disseminate it. The management team clearly understands how the school community sees the institution but needs a systematic action plan to continuously improve the school's identity. The management team acknowledges the pros and cons of teaching in their school.
5. On the part of the management team, there is an interest in creating a school's identity to make the institution recognized and appreciated by the whole school community. Actions are taken to disseminate it. The management team clearly understands how the school community sees the institution and has a systematic action plan to continuously improve the school's identity by perfecting the quality (pedagogical, climate, relationships, infrastructure). The management team acknowledges the pros and cons of teaching in their school.

D Official Statement from State Secretary of Education of Espírito Santo and Pará

Dear Principal,

The University of São Paulo - USP requested our collaboration to realize research with the State High Schools. Restating our commitment to educational public policy improvement in our management, we have given the Institution permission to proceed and contact our schools.

Your school will participate in this research, which addresses topics related to School Management Aspects. The Institution will provide data access to the State Secretary of Education without any school identification to preserve the confidentiality of respondents.

The research about School Management Aspects aims at understanding some topics of school management, and it will be an excellent opportunity for you to reveal your considerations about the subject and contribute to the improvement of public policies in education. The research will be carried out by telephone and conducted by *OPE-Sociais* consulting company, headquartered in Vitória, ES, with a chat format. The conversation is expected to take approximately 1 hour.

Therefore, we ask you to take a quiet moment to contribute with your detailed reflections about aspects of the management routine of your school.

The research coordinators will get in contact to schedule and explain the proceedings, and it will be realized between August and September.

Certain of your collaboration, we count on the dedication and commitment of all for the success of this action.

Best regards,

State Secretary of Education

E Independent Instrument - Vinha et al. (2017)

In this Appendix, we list the questions related to school management dimensions included in Vinha et al. (2017)'s survey instrument. These items are used in Figure 4 to provide additional validation of our Management Practices Index. All questions are answered on an agreement scale: (1) Do not agree; (2) Somewhat agree; (3) Agree; (4) Strongly agree.

Principals:

1. In this school, different kinds of information are shared and flow fast among parents, students, teachers, and other employees.
2. I engage the school community to identify problems and search for solutions.
3. This school is receptive to critics, suggestions, opinions, and contributions from teachers and other employees.
4. When school problems arise, small commissions are organized to solve them.
5. This school offers opportunities to study, professional training, and qualification for our professionals.
6. The pedagogical coordinator guides the teacher on how to improve classes.
7. This school has institutional evaluation instruments (for professionals' performance and relationship, school's functioning and structure).
8. Different school representatives participate in the institutional evaluation process.
9. We take external evaluation results into account to identify problems and search for solutions.
10. We take institutional evaluation results into account to identify problems and search for solutions.
11. The pedagogical project was elaborated with different segments of the school community.
12. The pedagogical project was based on a systematic diagnosis of the school's context.
13. The pedagogical project is revised periodically by the school community.

14. There is a mismatch between what is proposed in the pedagogical project and what is practiced in the school.
15. The National Curricular Guidelines (*Diretrizes Curriculares Nacionais*) are well-known by the management team and teachers.
16. The School's Regime is updated periodically.
17. I feel that I can rely on the help of other management team members when needed.
18. I feel that I can rely on the help of higher education instances (e.g. Directory and Superintendency of Education) when needed.
19. I spend much of my time dealing with bureaucratic issues.
20. New teachers are oriented and accompanied by school management when they start to teach at school.
21. The collective work meetings are held weekly or biweekly.
22. The school teachers trust me as a manager.
23. Most teachers respect the management teams' opinions and decisions.
24. The teachers clearly know the needs and problems of this school.
25. The teachers can rely on the support to enable their projects and activities (materials, workspaces, resources, etc.).
26. The number of teachers and employees is adequate to ensure the proper functioning of the school.
27. The teachers' working hours cover the time to participate in work meetings, teaching planning, and correcting activities.

Teachers:

1. The collective work meetings are held weekly or biweekly.
2. The management team encourages the participation of teachers in decisions related to the school's life.
3. The information flows adequately, facilitating the participation of all school professionals.

4. The management team clearly knows the school's problems and needs.
5. New teachers are oriented and accompanied by school management when they start to teach at school.
6. The principal is always at school.
7. The management has an authoritarian style.
8. Most teachers respect the management teams' opinions and decisions.
9. The teachers can rely on the support to enable their projects and activities (materials, workspaces, resources, etc.).
10. The number of teachers and employees is adequate to ensure the proper functioning of the school.
11. The teachers' working hours cover the time to participate in work meetings, teaching planning, and correcting activities.
12. The pedagogical coordinator guides the teacher on how to improve classes.
13. This school offers opportunities to study, professional training, and qualification for our professionals.

F *Jovem de Futuro* Program: Randomization Process

In this Appendix, we summarize the technical report from Instituto Unibanco, authored by *OPE-Sociais*, that describes the randomization process of the *Jovem de Futuro* program for schools that joined the program in 2015.

The program follows a phase-in strategy, where schools in the treatment group begin participating in the program in the first year. In contrast, schools selected for the control group will only participate in the program's third year. Thus, for the schools in the experiment used in this impact evaluation, the baseline year is 2014, the years of intervention are 2015, 2016, and 2017, and the impact evaluation of the program takes place in 2017.

Instituto Unibanco employs an experimental method to select schools for the treatment and control groups. The definition of the universe of schools that will receive the program in the first year (treated schools) and that will receive the program at the end of the evaluation window (control schools) is a decision taken by each State Department of Education, according to its management capacity and priorities. The only counterpart required by *Instituto Unibanco* is to maintain a minimum number of schools in the control group to enable the impact evaluation of *Jovem de Futuro*.

The randomization process is carried out within strata, stratifying by territorial characteristics of the schools (such as municipalities or regions to which they belong) and the schools' social vulnerability level. Given the universe of schools that will be able to receive the program, they group comparable schools to guarantee the homogeneity of the groups. First, schools are grouped according to the predicted evolution in performance in the absence of the program, controlling for the average language and mathematics score at baseline and the school's socioeconomic level index (*Indicador de Nível Socioeconômico – INSE*).

As we cannot observe the evolution of the schools' performance in the program's first two years in the absence of the intervention, they estimate the evolution using data from years before the baseline. Specifically, they use the growth between 2010 and 2012 to predict the growth between 2012 and 2014, comparing the actual change observed with the predicted change. The method selected for pairing the schools is the one that guarantees the best model for predicting the outcomes' evolution.

With this model estimated for each state participating in the program, they carry out the randomization process by stratum. The predicted evolution of performance is used to rank the schools within each stratum and group them. Then, after grouping

schools, they randomize similar schools within strata to the treatment or control status.

When randomly assigning schools to the treatment and control statuses, they ensured that each group (pair, trios, or quartets) had at least one control school. To check the validity of the randomization design, they employed tests for equality of means (t-test) and distribution (Kolmogorov-Smirnov). They verified the similarity of schools in the control and treatment groups in terms of their response variables (performance in Portuguese and Mathematics) at the baseline.

As reported in Table 1, for Pará state, there were 87 schools eligible to participate in the program, 45 assigned to the treatment group and 42 to the control group. They divided the schools into four groups according to their municipalities: Belém, Ananindeua, Marabá, and Santarém. In Pará, for reasons related to the Secretariat staff capacity, there was no stratification of schools by social vulnerability, only by territory.

In Espírito Santo, there were 221 eligible schools, 151 allocated to the treatment group and 70 to the control group. The eligibility criteria used by the Department of Education when delimiting the universe of schools were to select schools that, at the beginning of the intervention (the year 2015), had at least 120 students enrolled in the secondary school. They divided the schools into four groups for stratification based on the instruction regions. In addition, there was a second stratification level, considering vulnerable schools. The Department of Education indicated that the program could include 30 schools with the highest socioeconomic vulnerability. In the randomization process of the 221 schools, 188 schools were matched according to the predicted change in school performance between 2014 and 2016, 30 were matched based on the vulnerability criterion, and three were drawn in a simple randomization process.