Can Principal Training Improve School Management? Short-Term Experimental Evidence from Argentina*

Alejandro J. Ganimian†
New York University

Samuel Hansen Freel‡
New York University

September 26, 2020

Abstract

This paper presents one of the first evaluations of a principal-training program in a developing country. We randomly assigned 100 public primary schools in the Province of Salta, Argentina to a treatment group in which school principals attended a six-week, intensive, training workshop provided by an international foundation, or to a business-as-usual control group. Take up of the program was lower than expected: of the 52 schools assigned to the workshop, only 37 had a staff member sign up, 36 had someone attend, and the average participant attended 19 of the 26 sessions. Yet, satisfaction among participants was high: 99% of attendants reported the workshop was similar or better to other opportunities and 97% would recommend it to a peer. One month after the workshop, the program did not improve the availability or adequacy of school facilities, but it increased the links between schools and community organizations (e.g., teacher-training institutions and universities), as reported by principals. It had no impact on school climate, as reported by students, or on students' performance in math and language in the national assessment. Less than three months after the workshop, the program had no effect on students' performance in school. In fact, we can rule out small-to-moderate positive effects. The null results illustrate the difficulty of observing impacts of principal-training interventions in the short run, given that they often require non-trivial changes in management and instruction.

JEL codes: C93, I21, I22, I25

Keywords: school management, capacity-building, Argentina

*We gratefully acknowledge the funding provided by the Government Partnership Initiative (GPI) at the Abdul Latif Jameel Poverty Action Lab (J-PAL) to set up the Innovation and Evaluation Hub at the Ministry of Education, Science, and Technology of Salta. We thank Malena Araneo, Analía Berruezo, Gloria Crespo, Leonardo Ferrario, Gabriela Guerrero, Miguel Guitérrez, and Agustín Porres for making this study possible. We also thank María Cortelezzi, Fernando Giménez Zapiola, and Ismael Sanz Labrador for comments that informed this draft. Nicolás Buchbinder provided excellent research assistance. The usual disclaimers apply. This study was registered with the AEA Trial Registry (RCT ID: AEARCTR-0005361). All views expressed are our own and not of any institutions with which we are affiliated.

†Assistant Professor of Applied Psychology and Economics, New York University Steinhardt School of Culture, Education, and Human Development. E-mail: alejandro.ganimian@nyu.edu.

‡Doctoral Candidate, Psychology and Social Intervention, New York University Steinhardt School of Culture, Education, and Human Development. E-mail: shf281@nyu.edu.

1 Introduction

There is growing evidence indicating that school principals influence students' achievement. In the United States and Canada, several studies have found that principals' "value-added" (i.e., school-level year-on-year gains in student achievement, accounting for student and principal characteristics) varies widely across and within schools over time, suggesting that management practices matter for students' opportunity to learn (see, e.g., Branch, Hanushek and Rivkin 2012; Coelli and Green 2012; Dhuey and Smith 2014, 2018; Lipscomb, Chiang and Gill 2012).[1] These value-added estimates correlate with other measures of principals' performance (e.g., district evaluations) and time allocation (e.g., on management v. instructional tasks), suggesting that differences between/within schools are not merely due to selection or inputs (see, e.g., Grissom, Kalogrides and Loeb 2015; Horng, Klasik and Loeb 2010).[2] Further, principals make multiple decisions that impact the recruitment, effort, and professional development of teachers—the most important school-based determinant of student learning (see, e.g., Husain, Matsa and Miller 2018; Jacob 2011; Jacob and Lefgren 2005). This body of research is consistent with international evidence on the association between educational outcomes and school-management practices (see, e.g., Bloom et al. 2015; Woessmann 2016).

Yet, in spite of the importance of school principals, we know little about how to improve their effectiveness in low- and middle-income countries (for an evidence review, see Mbiti 2016). Experiments with "school-based management" (i.e., giving school management committees that include parents authority over budgetary, infrastructure, and personnel decisions) have yielded mixed results, working only when parents already possess high levels of capacity (see, e.g., Banerjee et al. 2010; Beasley and Huillery 2016; Blimpo, Evans and Lahire 2015; Duflo, Dupas and Kremer 2015). Attempts to provide principals with diagnostic (i.e., low-stakes) information on student performance to make them more responsive to students' needs have floundered in lower-middle income countries (Muralidharan and Sundararaman 2010) and produced one-off improvements in upper-middle income contexts (de Hoyos, Ganimian and Holland 2019). Yet, we still do not know how to raise principals' capacity. Experiments on

[1] In fact, one study estimates that highly effective principals raise the achievement of a typical student in their schools by between two and seven months of learning (Branch, Hanushek and Rivkin 2012). This makes principals the second most important school-based determinant of student achievement, after teachers. Importantly, however, several studies have highlighted the fact that the match between principals and schools accounts for a non-trivial share of school-level achievement gains, indicating that principal quality is not easily transferrable (see, e.g., Cannon, Figlio and Sass 2012; Dhuey and Smith 2018).

[2] As Chiang, Lipscomb and Gill (2016) note, however, principal value-added can only be estimated for schools that experience a transition in leadership. School value-added, which can be estimated for all schools, is weakly associated with principal value-added.

this more pressing question have mostly focused on *sui generis* interventions with limited potential to be taken-up elsewhere (see, e.g., Glewwe and Maïga 2011; Lassibille et al. 2010).[3]

This paper presents one of the first evaluations of a principal-training program in a developing country. We randomly assigned 100 public primary schools in the Province of Salta, Argentina to a "treatment" group, in which principals attended a six-week, intensive, training workshop offered by an international foundation, or to a "control" group that did not receive the training. Take up was lower than expected: of the 52 schools assigned to the workshop, 37 had a staff member enroll, 36 had someone attend, and the average participant attended 19 of 26 sessions.[4] Yet, satisfaction among participants was very high: 99% of attendants reported the workshop was similar or better to other opportunities and 97% would recommend it to a peer.

We examine the short-term impact of the program, a month after the end of the workshop, leveraging principal and student surveys and achievement data from the national assessment. We find the program was more successful at changing aspects of school management that can be easily influenced over the short-run (e.g., establishing links with community organizations, such as teacher-training institutions and universities) than at influencing aspects that require longer-lasting changes (e.g., the availability and adequacy of school facilities or school climate). We find no impact on math or reading achievement either, as measured by the assessment.

We explore the effect of the program on students' school performance by the end of the year, less than three months after the end of the workshop, leveraging the annual census of schools. We find the program had no effect on the number of students enrolled at school, or passing, failure, or dropout rates during the 2018 school year. In fact, all of these effects are precisely estimated around zero, and in most cases, allow us to rule out small-to-moderate impacts. Together, these null results illustrate the difficulty of observing impacts of principal-training interventions in the short run, given that they often need to be followed by non-trivial changes in management and instruction to influence student-level outcomes.

This study makes three contributions to research on school leadership in developing countries. First, it illustrates the challenge of seeking to improve the daily experiences of children at school by trying to influence the capacity of their principals. Such an approach requires that principals take up opportunities for professional development, that such opportunities equip them with the necessary knowledge and skills to implement meaningful changes at their school, that their teachers embrace their proposed changes, and that students respond as expected. We recommend that future studies collect data on each step of this complex theory of change. Second, our study also highlights the importance of understanding the nature of the changes

[3]One notable exception is an ongoing evaluation of an intervention designed to increase principals' role in supporting and supervising differentiated instruction in Ghana (Beg et al. 2019).

[4]The low take-up may have been in part due to problems in the sample drawn by the education ministry. Attendance rates should be interpreted as a lower-bound estimate due to errors in the program's record-keeping. We discuss both of these issues in greater detail in section 5.

that principals undertake, the obstacles they face most frequently, and their degree of success. We propose one way in which researchers can liaise with practitioners to collect such data. Third, our study speaks to the limitations of workshop-based approaches to capacity-building. The null short-term results that we find for this component of the program are consistent with those of similar initiatives for teachers (see, e.g., Angrist and Lavy 2001; Yoshikawa et al. 2015; Zhang et al. 2013) and suggest that researchers should investigate practice-based approaches for principals that resemble those that have proven effective at raising teachers' capacity (e.g., Bruns, Costa and Cunha 2017; Cilliers et al. 2019, 2020; Lucas et al. 2014).[5]

The paper is structured as follows. Section 2 presents the context, study design, and intervention. Section 3 describes the data. Section 4 discusses the empirical strategy. Section 5 reports the results. Section 6 discusses implications for future research.

2 Experiment

2.1 Context

Schooling in Argentina is compulsory and free from age 4 until the end of secondary school. In 12 out the 24 provinces including Salta, primary education runs from grades 1 to 7 and secondary education from grades 8 to 12 (DiNIECE 2013).[6] The Argentine school system serves 11.4 million students: 1.8 million in pre-school, 4.8 million in primary school, and 3.7 million in secondary school (DIEE 2016). The school year runs from February to December.

Most children in Argentina enroll in school: in 2017, the most recent year with available data, 98% of five-year-olds were in kindergarten and 97% of six-year-olds in grade 1 (DIEE 2019). Further, nearly all primary-school students graduate from this level. Yet, the achievement of these students lags behind curricular expectations: in 2018, the most recent year in which all grade 6 students were assessed, one in four students performed below expectations in language and nearly one in two did so in math (SEE-MEDN 2019*b*).[7] In fact, the relative standing of Argentine students in Latin America has deteriorated (see Ganimian 2014).

The Province of Salta is the eighth-largest sub-national primary school system in Argentina: in 2017, it served 28,204 students across 838 schools at that level (DIEE 2019). It performs similarly to the rest of the country: in 2018, 25% of grade 6 students scored in the lowest two levels of the national assessment in language and 40% did so in math (SEE-MEDN 2019*a*).

[5] As we discuss in section 6, the other components of the program that we study (e.g., school-based support in the implementation of school innovation projects) offer opportunities to pursue such research.

[6] In the other 12 provinces, primary runs from grades 1 to 6 and secondary from grades 7 to 12.

[7] These figures refer to students in the "basic" and "below basic" levels, which are below the "satisfactory" and "advanced" levels of each assessment.

2.2 Sample

The sample for the study includes 100 public primary schools in Salta, the capital city, and Tartagal, a remote urban area. We selected this sample in collaboration with the Ministry of Education, Science, and Technology (MECyT) of Salta as follows. First, the MECyT excluded secondary schools because we wanted to assess the intervention using the national assessment, and on the year of the study, only primary schools were assessed using census-based testing (SEE-MEDN 2019*b*). Second, it left out all private primary schools because we were interested in understanding the effect of the program on principals' capacity in the public school system. Third, it dropped all public primary schools that had already participated in the training program because we wanted to estimate the effect of receiving the training for the first time. Finally, among the remaining public primary schools, we selected the final 100 schools based on four eligibility criteria for principals to participate in the program (see section 2.4).

In-sample schools differ from out-of-sample schools, regardless of whether we compare them to all out-of-sample primary schools, out-of-sample public primary schools, or out-of-sample public primary schools in Salta and Tartagal. Specifically, in-sample schools enroll more students, have higher failure rates (across all primary-school grades), and lower overage rates than all three out-of-sample groups (Table A.1 in Appendix A). Additionally, the students in in-sample schools are more likely to have attended pre-K, less likely to be "chronically absent" (defined as being absent 15 or more days per year), perform more poorly in the national student assessment of 2017 in natural and social sciences, and have lower socio-economic status than their counterparts in out-of-sample schools (Table A.2).

2.3 Randomization

We randomly assigned the 100 sampled schools to one of two experimental groups, stratifying our randomization by geographic location and school category to maximize statistical power.[8] First, we grouped schools into six strata based on these two variables. Then, we randomly assigned schools in each stratum to: (a) a "treatment" group, whose principals were invited to participate in the program in 2018; or (b) a "control" group, which were invited to participate in 2019, after our study. This process resulted in 52 treatment and 48 control schools.

Control and treatment schools were comparable before the intervention was implemented, suggesting that the randomization worked as expected. We find no statistically significant

[8]In Salta, the MECyT classifies schools into three categories based on their geographic location and management staff: "category 1" schools are located in urban or semi-urban areas and have both a principal and a vice-principal; "category 2" schools are also located in urban or semi-urban areas, but they only have a principal; and "category 3" schools are located in rural areas and they also only have a principal.

differences on internal efficiency, either across all primary-school grades or grade 6 (Table A.3). We do not find any differences on student characteristics or achievement either (Table A.4).

2.4 Intervention

The Program on Leadership and Innovation in Education (PLIE) is an in-service professional development initiative for principals, vice-principals, and teachers designed by the Varkey Foundation (VF), a charity headquartered in the United Kingdom, in 2016. It was adapted for the Argentine context in conjunction with four sub-national ministries of education, a local non-profit, and domestic and international researchers. It has been implemented in the provinces of Corrientes, Jujuy, Mendoza, and Salta with support from the central government.[9] To date, PLIE has reached 6,544 principals, vice-principals, and teachers across 3,591 schools, and the VF plans to reach 15,000 school officials across the country.[10]

In Salta, to participate in the PLIE, principals have to: (a) be tenured public officials (i.e., they cannot be interim or substitute principals); (b) not be close to the age of retirement; (c) not be the only teacher at their school; and (d) be able to take six weeks off regular duties. Principals may send a vice-principal on their behalf, and they may also invite a teacher.

The program has four main components: (a) a six-week, in-person workshop; (b) a "school innovation project", which participants develop during the workshop and implement in the six months after the workshop; (c) an online portal, to which participants are granted access after the workshop; and (d) school visits by VF staff in the six months after the workshop.

The workshop includes six modules: (a) education leadership for organizational development and school reform; (b) managing technology integration; (c) leading and managing learning, creativity, and curriculum innovation; (d) educational leadership for quality assurance and to improve performance in the teaching and learning process; (e) leading teacher professional development;[11] and (f) leading and developing community relations. Participants are grouped with others working in different roles, education levels, academic programs, and school types. In the morning sessions, participants engage in lectures and discussions. In the afternoon, they work in teams to solve specific problems of practice (Alonso 2018). They receive support from facilitators, who lead sessions, and tutors, who provide one-on-one and small-group guidance.[12]

During the workshop, participants develop a school innovation project to improve their school. These projects start with a diagnosis of the school, which is then followed by the rationale,

[9] For a detailed account of the development of the program in Argentina, see Alonso (2018).

[10] Varkey Foundation's website: `https://bit.ly/3ci4YRS`. Last accessed: June 1, 2020.

[11] During this module, participants visit a school being led by one of their peers. In this visit, participants try to reconcile what they have learned with the day-to-day realities of the school.

[12] Facilitators and tutors are required to have a graduate degree in education or a related field. Typically, facilitators have more work experience and have served as school principals.

objectives, action steps, evaluation plan, and timeline for a set of improvements to the school. Participants are expected to implement the project in the six months after the workshop. During these months, VF representatives (called *seguidores*) visit schools to check on the progress of each project and lend additional support with implementation (VF 2019).[13]

After the workshop, participants can join the "Varkey Academy", an online portal to take courses, complete activities, join discussion groups, and access curated content (VF 2019). The aim of this portal is to support a network of principals seeking to improve their schools.

Table 1 shows the theory of change of the intervention. Importantly, in the present study, we only observe some of the target outcomes a month after the workshop, so our results capture the short-term effect of the workshop on these outcomes, not of the program as a whole.[14] Only the effects on school performance, which were estimated roughly three months after the workshop, can provide suggestive evidence on the value-added of the other components.

3 Data

As Table 2 shows, we obtained access to: (a) all data from the annual census of schools for the year prior to the program (2017) and the year in which it was implemented (2018); (b) all available data for the national assessment for the same years, which includes students' performance on the test as well as student and principal surveys; (c) principals' enrollment and attendance to the training, to verify that they participated; and (d) principals' surveys at the end of the program to measure their satisfaction.

3.1 Annual census of schools

The MECyT granted us access to all data collected during the annual census of schools (e.g., passing and dropout rates). We use the data for 2017, the year prior to the intervention, to compare in- and out-of-sample schools and check balance across experimental groups and the data for 2018, the year of the intervention, to estimate the program's impact. These data are collected at the school level, so we report results for all primary-school grades and grade 6. The 2018 data are particularly useful because they are the most distal set of indicators on which we can assess impact; they capture students' performance by the end of the year.[15]

[13]These meetings may be individual- or group-based, and they may be held in person or virtually.

[14]PLIE underwent an external evaluation from 2016 to 2018, but several aspects of the design of this study do not allow for causal inferences about the impact of the program (see VF 2019).

[15]In fact, as Table 2 shows, these data were finalized in April of the *following* school year to account for any students who may passed the grade after taking make-up exams in December of 2018 or March of 2019.

3.2 National student assessment

The MECyT also granted us access to all data in the national, census-based student assessment of grade 6. These include students' achievement in natural and social sciences (in 2017) and in math and language (in 2018), as well as responses to student and principal surveys.[16] We use the 2017 data for balancing checks and the 2018 data to estimate the program's impact. The national ministry of education scaled all test scores using a two-parameter logistic Item Response Theory (IRT) model (Yen and Fitzpatrick 2006), so all effects in this paper are expressed with respect to the overall national distribution. This feature of the achievement data sets this study apart from most evaluations of educational interventions, which rely on assessments designed by researchers and administered over a convenience sample.

3.3 Principals' enrollment in and attendance to the program

The VF provided us with the roster of school officials (principals and accompanying members) who enrolled in the program and with the number of days each of these individuals attended. We use these data, available for 2018, to check the intervention was delivered as intended.[17]

3.4 Principals' satisfaction with the program

The VF also gave us access to the surveys administered to workshop participants. We use these data, available for 2018, to measure whether participants were satisfied with their training.

4 Empirical strategy

We estimate the effect of the offer of the intervention (i.e., the intent-to-treat or ITT effect) by fitting the following model:

$$Y_{is}^{t} = \alpha_{r(s)} + \gamma \bar{Y}_{s}^{t-1} + \beta T_s + \varepsilon_{is}^{t}, \tag{1}$$

where Y_{is}^{t} is an outcome for student i in school s and year t, $r(s)$ is the randomization stratum of school s and $\alpha_{r(s)}$ is a stratum fixed effect, $\bar{Y}_{s}^{t-1}$ is the school-level average of the same

[16] The student surveys can be accessed at: `https://bit.ly/2TSwzUn` (2018). The principal surveys can be accessed at: `https://bit.ly/2RfSbZ3` (2018).

[17] The VF also administers surveys to all workshop participants upon entering the program, to understand their expectations for the training, and at the end of each week of the workshop, to assess their satisfaction. The weekly surveys ask similar questions as the end-of-workshop survey that we use (VF 2019).

outcome for year $t-1$,[18] and T_s is an indicator variable for random assignment to treatment. The parameter of interest is β, which captures the causal effect of the intervention. We use cluster-robust standard errors to account for within-school correlations across students in outcomes. We also test the sensitivity of our estimates to the inclusion of $\bar{Y}_s^{t-1}$. We fit variations of this model that interact the treatment dummy with school characteristics (indexes of prior-year achievement, resources, and supports) to estimate the heterogeneous effects of the intervention on these sub-groups.[19]

5 Results

5.1 Enrollment in and attendance to the workshop

Enrollment in the program was lower than expected. According to the enrollment records provided by the VF, only 37 of the 52 treatment schools (71% of this group) had one or more members (principal, vice-principal, or teacher) sign up for the training.[20] Of those treatment schools with enrollment data, only 16 of them (31%) had a principal signed up for the training; 13 of them (25%) had a vice-principal instead. Twenty-nine schools (56% of the treatment group) had a teacher sign up for the training: in 9 schools, teachers were alone; in 20 schools, they accompanied a principal or vice-principal. According to the VF, the low take-up of this cohort was atypical. In the previous three cohorts of primary schools that were invited to participate in the program in 2018, the take-up rates were 72%, 63%, and 85%, respectively.[21]

Attendance to the workshop was also below expectations. According to attendance records provided by the VF, only 36 of the 52 treatment schools (69% of this group) had one or more members attend the training. Among these schools, the average participant attended 19 of 26 possible sessions (73%). Principals, vice-principals, and teachers had similar average attendance rates: 19, 19, and 20, respectively. According to the VF, the attendance figures should be understood as a lower-bound estimate of actual attendance to the workshop. Due to errors in record-keeping, some participants were not marked as present even if they attended.[22]

[18] The national census of schools and national assessment are repeated cross-sections of students, so we do not observe each student's prior-year outcome.

[19] The index of prior-year achievement is the school-level average score in the 2017 national assessment, which covered natural and social sciences. The indexes of school resources and supports are the first principal components from principal component analyses of questions in the 2017 survey of principals on the resources and supports for low-performing students at the school, respectively.

[20] Additionally, three control schools, which were not supposed to receive the training, had also signed up.

[21] We cannot determine what led to the low take-up for the evaluation sample, but it seems to be at least in part due to a problem in the identification of eligible schools by the MECyT.

[22] For the purposes of this paper, we treat those "unregistered" cases as absent.

Nearly all schools (97%) had completed at least one project as part of the training.[23] In fact, the average treatment school had completed two projects by the end of the workshop.

5.2 Satisfaction with the workshop

Principals, vice-principals, and teachers who participated in the workshop were very satisfied. We have participants' responses to a survey administered on the last day of the workshop.[24] More than 99% of survey respondents reported that the workshop was similar (16%) or better (83%) than other professional-development opportunities in which they had participated. Further, more than 97% indicated that they would likely (11%), very likely (12%), or definitely (74%) recommend the workshop to others. Participants also gave instructors very high ratings: 94% rated tutors as excellent or very good and 99% rated facilitators as excellent or very good.

The aspects of the workshop that were most valued by the participants varied widely. The most valued aspect was the quality of instruction (79%), followed by the use of technology (76%), facilitators (65%), content (64%), support (54%), tutors (52%), and school visits (39%). The workshop units that participants valued also varied: 30% rated all units as useful, but the three units most frequently rated as useful were those covering managing technology integration (29%), leading and managing learning, creativity, and curriculum innovation (16%), and educational leadership for quality assurance and improvement of teaching and learning (11%). These responses broadly coincided with those on the units considered as most relevant.

5.3 Availability and adequacy of school facilities

We do not find evidence that the program led to increased availability of school facilities. As Table 3 indicates, principals in treatment schools on average reported having similar availability of school facilities (classroom heating, disability access, electricity, full restrooms, paved access, a sewage system, or running water) as those in control schools. In fact, the effect of the program on the availability of these facilities is consistently estimated around zero.[25] According to the VF, this null effect on infrastructure is to be expected, given that participants are encouraged to focus on aspects that do not require additional funding (which depends on the federal and province governments) to be implemented in their school innovation projects.

[23] According to the VF, the only school that did not completed a project did so after the training.

[24] Surveys were anonymous, so we cannot identify which respondents belong to each school or match them to enrollment or attendance data.

[25] There is only one exception. treatment principals were 13 percentage points (pp.) *less* likely than their control counterparts to report having heating in their classrooms, but it is possible that this difference emerged due to multiple-hypothesis testing.

We do not find that the program improved the adequacy of facilities either. In fact, as Table 4 shows, principals in treatment schools were generally *less* likely to report that their school had adequate facilities than those in the control group. Specifically, they were 16 pp. less likely to report having adequate classrooms and 17 pp. less likely to report having an adequate multi-use room (although both differences were only marginally statistically significant). It is possible that participants paid greater attention to other aspects of school management. Yet, it seems more likely that they suddenly came into contact with a different reference group, which in turn affected their subjective appraisal of the adequacy of their own facilities.[26]

5.4 Links with community organizations

The program increased school's propensity to establish links with organizations in their communities. As Table 5 indicates, treatment principals are consistently more likely than their control peers to report that their school has links with community organizations. Specifically, they were 13 pp. more likely to report having links with cultural centers, 20 pp. more likely to report having links with teacher-training institutions, and 9 pp. more likely to report having links with universities. The estimates on all other community links are statistically insignificant, but positive. Given that establishing links with organizations is something that principals can do in the short-run, it is perhaps not surprising that the program most clearly impacted this set of outcomes. The survey provides no information, however, on the nature of such links (e.g., whether they focus on one-time events or longer-lasting activities).

5.5 School climate

We find that the intervention had no impact on student-reported school climate indicators. As Table 6 shows, students in school assigned to the workshop were no more likely, on average, to report that they enjoyed going to school, got along with peers, or that students at their school were bullied, discriminated against, or damaged school property. The point estimates across all variables are consistently estimated around zero, and across all outcomes, we can rule out small positive effects (larger than 7 pp.) This survey of students was administered as part of the 2018 national assessment, a month after the end of the workshop, so principals in treatment schools may not have had enough time to address these issues after the program.

[26]This seems particularly likely given that, as stated in section 2.4, participants visited several schools during the workshop.

5.6 Students' performance in school

The program had no discernible effects on students' performance in school. As Table 7 shows, schools assigned to the program did not differ in the number of enrolled students, passing or failure rates, or dropout rates from those not assigned to the program. This pattern of results holds across all primary grades and in grade 6, the target grade for the national assessment. Once again, all effects are consistently estimated around zero, allowing us to rule out effects larger than 5 pp. across primary and effects larger than 8 pp. in grade 6.

Note that data on students' performance in school is finalized in April of the following year (in this case, 2019). This means that the program produced no improvements in students' propensity to do well in school even six months after the end of the workshop.

5.7 Students' achievement

We do not find that the intervention improved students' achievement in math or language. As Table 8 indicates, students enrolled in schools assigned to the workshop performed, on average, 0.01 standard deviations (SDs) *below* their counterparts in control schools in math and 0.11 SDs below in language, but neither difference was statistically significant. We cannot rule out small-to-moderate positive or negative effects in either subject, but the difference between treatment and control schools actually widens when we account for schools' prior-year performance in the national assessment, which focused on natural and social sciences, providing further evidence of the advantage of control schools.

As stated above, the national assessment was administered one month after the workshop, so we did not necessarily expect to see impacts on student achievement over such a short term. Yet, our results raise the possibility that taking principals out of schools for six weeks might have temporarily adversely influenced learning.

6 Conclusion

The present study offers several lessons for research on principals' professional development in low- and middle-income countries. First, it illustrates the challenge of seeking to improve the daily experiences of children at school by trying to influence the capacity of their principals. Such a process requires not only that principals enroll and attend the workshops and that they conduct work during the workshops with potential to meaningfully influence student learning, but also that the changes that they propose (in this case, in the school innovation projects) are willingly embraced and aptly implemented by teachers and that students respond as expected.

This is a risky proposition: if any of those steps do not occur as expected, the intervention is unlikely to improve student-level outcomes (e.g., performance in school and achievement). This should not dissuade school systems from the worthy goal of raising principals' capacity, but it draws attention to the multiple points at which this theory of change may break down. One important limitation of our study is that we were only able to obtain data on parts of this hypothesized causal chain. Future studies should try, to the extent that it is possible, to collect data on all of the crucial steps that are necessary to make the intervention work.

Second, our study also highlights the importance of understanding the nature of the changes that principals seek to undertake. Our evaluation of the workshop can tell us whether training principals on management practices and helping them develop, implement, and monitor changes in their schools improves management, instruction, and student learning. Yet, it cannot shed light on the types of changes that principals are more likely to pursue, the most frequent roadblocks they face (e.g., gaps in their training, resistance from colleagues), and the extent to which some changes are more likely to be enacted successfully than others. Further research on these questions is needed to better understand the types of support that principals need to pursue and accomplish meaningful changes. One way in which this could be done is by working alongside training providers to codify aspects of the plans developed by principals and to use check-in visits to collect data on their implementation fidelity.

Third, our evaluation also speaks to the limitations of workshop-based approaches to training. As we clarify above, we did not evaluate the full program developed by the Varkey Foundation, which includes the workshop, a school innovation project, school visits, and an online portal; we only studied the short-term effect of the first and part of the second of these components.[27] Our null results are consistent with those of lecture-style in-service training for teachers in other developing countries, which also indicate this type of training is insufficient to change trainees' practices (see, e.g., Angrist and Lavy 2001; Yoshikawa et al. 2015; Zhang et al. 2013). In light of the far more encouraging results of practice-based approaches to teacher training (e.g., Bruns, Costa and Cunha 2017; Cilliers et al. 2019, 2020; Lucas et al. 2014), the other components of the leadership program by the Varkey Foundation hold greater promise.

[27] While our outcomes were measured after the development of the school innovation projects, principals only had two months to start implementing them.

References

Alonso, A. A. 2018. "PLIE: Improving the capacity of school leaders in Argentina." *Public Education Leadership Project at Harvard University*. Cambridge, MA: Harvard Graduate School of Education (HGSE).

Angrist, J. D., and V. Lavy. 2001. "Does teacher training affect pupil learning? Evidence from matched comparisons in Jerusalem public schools." *Journal of Labor Economics*, 19(2).

Banerjee, A. V., R. Banerji, E. Duflo, R. Glennerster, and S. Khemani. 2010. "Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India." *American Economic Journal: Economic Policy*, 2(1-30).

Beasley, E., and E. Huillery. 2016. "Willing but unable? Short-term experimental evidence on parent empowerment and school quality." *The World Bank Economic Review*, 31(2): 531–552.

Beg, S., A. Fitzpatrick, A. M. Lucas, E. Tsinigo, and H. Atimone. 2019. "Strengthening teacher accountability to reach all students (STARS)." (World Bank SIEF). Milestone 3: Observational survey field report.

Blimpo, M. P., D. K. Evans, and N. Lahire. 2015. "Parental human capital and effective school management: Evidence from the Gambia." (Policy Research Working Paper No. 7238). Washington, DC: The World Bank.

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

Branch, G. F., E. A. Hanushek, and S. G. Rivkin. 2012. "Estimating the effect of leaders on public sector productivity: The case of school principals." National Bureau of Economic Research.

Bruns, B., L. Costa, and N. Cunha. 2017. "Through the looking glass: Can classroom observation and coaching improve teacher performance in Brazil?" *Unpublished manuscript*. Washington, DC: The World Bank.

Cannon, S., D. Figlio, and T. Sass. 2012. "Principal quality and the persistence of school policies." *Unpublished paper, Northwestern University*.

Chiang, H., S. Lipscomb, and B. Gill. 2016. "Is school value added indicative of principal quality?" *Education Finance and Policy*, 11(3): 283–309.

Cilliers, J., B. Fleisch, C. Prinsloo, and S. Taylor. 2019. "How to improve teaching practice? An experimental comparison of centralized training and in-classroom coaching." *Journal of Human Resources*, 0618–9538R1.

Cilliers, J., B. Fleisch, J. Kotze, N. Mohohlwane, S. Taylor, and T. Thulare. 2020. "Can virtual replace in-person coaching? Experimental evidence on teacher professional development and student learning in South Africa." *Unpublished manuscript*. Washington, DC: Georgetown University.

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

de Hoyos, R., A. J. Ganimian, and P. A. Holland. 2019. "Teaching with the test: Experimental evidence on diagnostic feedback and capacity-building for schools in Argentina." *World Bank Economic Research*.

Dhuey, E., and J. Smith. 2014. "How important are school principals in the production of student achievement?" *Canadian Journal of Economics/Revue canadienne d'économique*, 47(2): 634–663.

Dhuey, E., and J. Smith. 2018. "How school principals influence student learning." *Empirical Economics*, 54(2): 851–882.

DIEE. 2016. "Anuario estadístico 2016." Buenos Aires, Argentina: Dirección de Investigación y Estadística Educativa (DIEE).

DIEE. 2019. "Anuario estadístico 2018." Buenos Aires, Argentina: Dirección de Investigación y Estadística Educativa (DIEE).

DiNIECE. 2013. "Redefiniciones normativas y desafíos de la educación secundaria en Argentina. Acuerdos federales en un sistema descentralizado." La educación en debate. Buenos Aires, Argentina: Dirección Nacional de Información y Evaluación de la Calidad Educativa (DiNIECE).

Duflo, E., P. Dupas, and M. Kremer. 2015. "School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from Kenyan primary schools." *Journal of Public Economics*, 123: 92–110.

Ganimian, A. J. 2014. "Avances y desafíos pendientes: Informe sobre el desempeño de Argentina en el Tercer Estudio Regional Comparativo y Explicativo (TERCE) del 2013." Buenos Aires, Argentina: Proyecto Educar 2050.

Glewwe, P., and E.W.H. Maïga. 2011. "The impacts of school management reforms in Madagascar: do the impacts vary by teacher type?" *Journal of Development Effectiveness*, 3(4): 435–469.

Grissom, J. A., D. Kalogrides, and S. Loeb. 2015. "Using student test scores to measure principal performance." *Educational Evaluation and Policy Analysis*, 37(1): 3–28.

Horng, E. L., D. Klasik, and S. Loeb. 2010. "Principal's time use and school effectiveness." *American Journal of Education*, 116(4): 491–523.

Husain, A. N., D. A. Matsa, and A. R. Miller. 2018. "Do male workers prefer male leaders? An analysis of principals' effects on teacher retention." National Bureau of Economic Research.

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

Jacob, B. A., and L. Lefgren. 2005. "Principals as agents: Subjective performance measurement in education." National Bureau of Economic Research.

Lassibille, G., J. Tan, C. Jesse, and T. Van Nguyen. 2010. "Managing for results in primary education in Madagascar: Evaluating the impact of selected workflow interventions." *The World Bank Economic Review*, 24(2): 303–329.

Lipscomb, S., H. Chiang, and B. Gill. 2012. "Value-added estimates for Phase 1 of the Pennsylvania Teacher and Principal Evaluation Pilot. Full Report." *Mathematica Policy Research, Inc.*

Lucas, Adrienne M, Patrick J McEwan, Moses Ngware, and Moses Oketch. 2014. "Improving early-grade literacy in East Africa: Experimental evidence from Kenya and Uganda." *Journal of Policy Analysis and Management*, 33(4): 950–976.

Mbiti, I. M. 2016. "The need for accountability in education in developing countries." *Journal of Economic Perspectives*, 30(3): 109–32.

Muralidharan, K., and V. Sundararaman. 2010. "The impact of diagnostic feedback to teachers on student learning: Experimental evidence from India." *The Economic Journal*, 120(F187-F203).

SEE-MEDN. 2019*a*. "Aprender 2018: Informe de resultados, Primaria - 6to año, Provincia de Salta." Ciudad Autónoma de Buenos Aires: Secretaría de Evaluación Educativa. Ministerio de Educación y Deportes de la Nación.

SEE-MEDN. 2019*b*. "Aprender 2018: Informe nacional de resultados, 6to año nivel primario." Ciudad Autónoma de Buenos Aires: Secretaría de Evaluación Educativa. Ministerio de Educación y Deportes de la Nación.

VF. 2019. "Evaluación de resultados intermedios: Presentación del impacto del Programa de Liderazgo e Innovación Educativa (PLIE)." Buenos Aires, Argentina: Varkey Foundation (VF).

Woessmann, L. 2016. "The importance of school systems: Evidence from international differences in student achievement." *Journal of Economic Perspectives*, 30(3): 3–32.

Yen, W. M., and A. R. Fitzpatrick. 2006. "Item response theory." In Brennan, R. (Ed.) *Educational measurement* (4th ed.). Westport, CT: American Council on Education and Praeger Publishers.

Yoshikawa, H., D. Leyva, C. E. Snow, E. Treviño, M. C. Arbour, M. C. Barata, C. Weiland, C. Gómez, L. Moreno, A. Rolla, and N. D'Sa. 2015. "Experimental impacts of a teacher professional development program in Chile on preschool classroom quality and child outcomes." *Journal of Developmental Psychology*, 51: 309–322.

Zhang, L., F. Lai, X. Pang, H. Yi, and S. Rozelle. 2013. "The impact of teacher training on teacher and student outcomes: Evidence from a randomised experiment in Beijing migrant schools." *Journal of Development Effectiveness*, 5(3): 339–358.

Table 1: Theory of change of the Education Leadership and Innovation Program in Salta

(1) **Need**	(2) **Inputs/Activities**	(3) **Outputs**	(4) **Outcomes**	(5) **Impact**
• **Principals do not receive adequate training on school management**, which results in them performing poorly in key tasks (e.g., managing teachers, reforming curricula, etc.)	• **VF provides six-week in-person workshop** for school principals • **VF supports development of school innovation project** for workshop participants • **VF offers online portal** to workshop alumni • **VF representatives visit schools** of participants to check on school innovation project implementation	• **Principals attend the workshop** or send a vice-principal, accompanied by a teacher • **Principals develop a school innovation project** during the workshop • **Principals implement the school innovation project** during six months after the workshop	• **Principals run their schools in alignment with its vision/mission**, guided by clear values • **Principals focus on improving student learning** at their schools • **Principals manage human/physical resources efficiently** at their schools • **Principals foster community engagement** at their schools (e.g., from nonprofits, universities)	• **School resources are used more efficiently** (e.g., higher incidence of links with community organizations) • **School climate improves** (e.g., lower incidence of bullying) • **Teachers' instruction improves** (e.g., higher prevalence of practices mindful of students' needs) • **Students' achievement improves** (e.g., higher scores on standardized tests)
Assumptions:	• Principals believe they need to improve their management practices • Principals see the program as addressing their training needs	• Principals can take time off to attend workshops • Principals have Internet to access online portal • Principals have time to medium- and long-term projects while coping with day-to-day issues	• Principals do not face major resource constraints • Principals secure the support of their teachers	• Principals act on aspects that influence teachers' instruction and/or students' learning

Notes: Authors' adaptation of VF (2019).

Table 2: Data collection timeline of the study

(1)	(2)	(3) School participation rates	(4)
Month	Event	Control schools	Treatment schools
A. 2017			
February	School year starts		
November	National assessment of grade 6 students (principal surveys, student surveys, and tests of natural and social sciences)	98%	100%
December	School year ends		
B. 2018			
February	School year starts		
April	Annual census of schools for 2017 school year	100%	100%
August	Training participants enrollment data	-	65%
October	Training participants attendance data	-	72%
November	National assessment of grade 6 students (principal surveys, student surveys, and tests of math and language)	100%	94%
December	School year ends		
C. 2019			
February	School year starts		
April	Annual census of schools for 2018 school year	100%	100%

Notes: The table shows the timeline for the interventions and rounds of data collection for the study, including the month in which each event occurred (column 1), a brief description of the event (column 2), and the percentage of schools that participated in each event by experimental group (columns 3-4).

Table 3: ITT effect on principal-reported availability of school facilities (2018)

	(1) Classroom heating	(2) Disability access	(3) Electricity	(4) Back-up electricity	(5) Restroom	(6) Paved access	(7) Sewage system	(8) Running water
Treatment	-0.133**	-0.026	0.025	0.027	-0.003	0.060	0.011	0.025
	(0.056)	(0.095)	(0.022)	(0.054)	(0.031)	(0.088)	(0.081)	(0.054)
N (schools)	86	86	92	82	92	90	89	92
Control mean	0.143	0.349	0.978	0.050	0.978	0.467	0.500	0.913

Notes: (1) The table shows the intent-to-treat effect of the intervention on the availability of school facilities reported by principals, as measured during the national assessment in 2018, which was administered a month after the end of the workshop. (2) Principals were asked whether their school had classroom heating, disability access, electricity, back-up electricity, a full restroom, paved access, a sewage system, and running water (in all cases, these were yes/no questions). (3) The control mean shows the share of principals in the control group who reported that their school had a given facility. For example, 14% of control principals reported that their school had heating for classrooms. (4) * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: ITT effect on principal-reported adequacy of school facilities (2018)

	(1) School building	(2) Classrooms	(3) Library	(4) Playground	(5) Multi-use room
Treatment	-0.055	-0.157*	0.019	-0.055	-0.172*
	(0.073)	(0.084)	(0.097)	(0.091)	(0.092)
N (schools)	90	90	88	89	92
Control mean	0.889	0.889	0.467	0.778	0.348

Notes: (1) The table shows the intent-to-treat effect of the intervention on the adequacy of school facilities reported by principals, as measured during the national assessment in 2018, which was administered a month after the end of the workshop. (2) Principals were asked whether their school had *adequate* building, classrooms, library, playground, and multi-use room (in all cases, these were yes/no questions). (3) The control mean shows the share of principals in the control group who reported that their school had an adequate facility. For example, 89% of control principals reported that their school had an adequate building. (4) * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: ITT effect on principal-reported links with community organizations (2018)

	School has links with...								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Community spaces	Cultural centers	Museums	Non-profit organizations	Other schools	Religious institutions	Sport centers	Teacher training institutions	Universities
Treatment	0.095	0.128*	0.094	0.045	0.113	0.003	0.021	0.200**	0.089*
	(0.092)	(0.074)	(0.069)	(0.086)	(0.099)	(0.094)	(0.074)	(0.093)	(0.047)
N (schools)	96	96	96	96	96	96	96	96	96
Control mean	0.312	0.125	0.104	0.208	0.458	0.312	0.146	0.333	0.021

Notes: (1) The table shows the intent-to-treat effect of the intervention on the links between the school and its community reported by principals, as measured during the national assessment in 2018, which was administered a month after the end of the workshop. (2) Principals were asked whether their school had links with community spaces, cultural centers, museums, non-profits, other schools, religious institutions, sport centers, teacher-training institutions, or universities (in all cases, these were yes/no questions). (3) The control mean shows the share of principals in the control group who reported that their school had an adequate facility. For example, 89% of control principals reported that their school had an adequate building. (4) * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: ITT effect on student-reported school climate (2018)

	(1)	(2)	(3)	(4)	(5)
	Enjoys going to school	Gets along with peers	Students are bullied often	Students are discriminated against often	Students damage school property often
Treatment	-0.010	-0.006	-0.010	0.016	0.011
	(0.012)	(0.011)	(0.028)	(0.029)	(0.016)
N (children)	4766	4909	4909	4909	4909
Control mean	0.933	0.888	0.472	0.348	0.150

Notes: (1) The table shows the intent-to-treat effect of the intervention on indicators of school climate reported by grade 6 students, as measured during the national assessment in 2018, which was administered a month after the end of the workshop. (2) Students were asked whether they enjoy going to school (yes/no question); whether they get along with peers (Likert-type question, here displaying students who reported they get along with some, most, or all peers); the frequency with which students are bullied due to school performance (Likert-type question, here displaying those who report students are bullied often or always for getting good grades or repeating a grade); the frequency with which students are discriminated against (Likert-type question, here displaying those who report students are discriminated against often due to disabilities, nationality, physical appearance, or religion); and the frequency of student vandalism (Likert-type question, here displaying those who report students damage school property often or always). (3) The control mean shows the share of students in the control group who agreed with the statement, as indicated in the previous note. For example, 93% of control students enjoys going to school. (4) * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: ITT effect on students' performance in school (2018)

	(1) Enrolled students	(2) Passing rate	(3) Failure rate	(4) Dropout rate
A. Primary school				
Treatment	-3.579	-0.017	0.005	-0.005
	(27.326)	(0.017)	(0.006)	(0.006)
N (schools)	99	99	99	99
Control mean	411.604	0.956	0.020	0.011
B. Grade 6				
Treatment	1.298	-0.023	0.008	-0.009
	(3.956)	(0.029)	(0.009)	(0.007)
N (schools)	99	99	99	99
Control mean	54.812	0.953	0.013	0.014

Notes: (1) The table shows the intent-to-treat effect of the intervention on student achievement of grade 6 students, as measured by the national assessment in 2018, which was administered a month after the end of the workshop. (2) The prior-year school index is the first principal component from a principal component analysis that included the school-level average score of grade 6 students in the national assessment achievement in 2017, which focused on natural and social sciences. (3) The control mean is expressed with respect to the national achievement distribution. For example, the average control school performed .018 standard deviations above the national mean in grade 6 math in 2018. (4) * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8: ITT effect on student achievement (2018)

	Math		Language	
	(1)	(2)	(3)	(4)
Treatment	-0.014	-0.026	-0.117	-0.142
	(0.091)	(0.083)	(0.125)	(0.112)
Prior-year school index		0.290***		0.467***
		(0.102)		(0.113)
N (children)	4730	4725	4692	4686
Control mean	0.018		0.064	

Notes: (1) The table shows the intent-to-treat effect of the intervention on student achievement of grade 6 students, as measured by the national assessment in 2018, which was administered a month after the end of the workshop. (2) The prior-year school index is the first principal component from a principal component analysis that included the school-level average score of grade 6 students in the national assessment achievement in 2017, which focused on natural and social sciences. (3) The control mean is expressed with respect to the national achievement distribution. For example, the average control school performed .018 standard deviations above the national mean in grade 6 math in 2018. (4) * significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix A Additional graphs and tables

Table A.1: Comparison between in- and out-of-sample schools on internal efficiency (2017)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Out-of-sample schools						
	All schools	All	Public	Public and Salta or Tartagal	In-sample schools	Col.(5)-Col.(2)	Col.(5)-Col.(3)	Col.(5)-Col.(4)
A. Primary school								
Enrolled students	221.654	196.734	180.764	270.142	405.56	208.826***	224.796***	147.325***
	(278.765)	(268.737)	(272.335)	(306.202)	(283.863)	(28.832)	(29.427)	(33.399)
Passing rate	96.753	96.833	96.536	97.044	96.162	-.672	-.375	-.819
	(7.292)	(7.502)	(7.928)	(6.277)	(5.492)	(.777)	(.822)	(.701)
Failure rate	1.977	1.867	2.001	1.925	2.783	.916**	.783*	.883**
	(3.969)	(4.04)	(4.252)	(3.656)	(3.308)	(.422)	(.445)	(.404)
Repetition rate	2.675	2.61	2.772	2.462	3.16	.55	.388	.663
	(4.937)	(5.034)	(5.272)	(4.25)	(4.14)	(.526)	(.552)	(.476)
Overage rate	9.138	9.844	10.472	7.282	5.512	-4.332***	-4.96***	-2.271**
	(11.76)	(12.344)	(12.567)	(9.687)	(7.133)	(1.315)	(1.34)	(1.122)
Dropout rate	1.27	1.299	1.463	1.031	1.055	-.245	-.408	-.063
	(5.658)	(5.826)	(6.193)	(4.717)	(4.244)	(.603)	(.642)	(.533)
N (schools)	1111	738	648	644	100	838	748	744
B. Grade 6								
Enrolled students	32.084	28.733	26.622	38.459	54.57	25.837***	27.948***	17.553***
	(38.726)	(37.504)	(37.957)	(42.243)	(39.475)	(4.048)	(4.13)	(4.653)
Passing rate	97.001	97.122	96.911	97.251	96.182	-.94	-.73	-.95
	(9.941)	(9.672)	(10.238)	(8.662)	(11.631)	(1.065)	(1.129)	(1.07)
Failure rate	1.554	1.484	1.513	1.423	2.023	.539	.51	.607
	(5.577)	(5.774)	(6.056)	(4.35)	(3.977)	(.597)	(.627)	(.488)
Repetition rate	2.85	2.72	2.729	2.697	3.716	.996	.988	.912
	(9.284)	(9.393)	(9.801)	(7.823)	(8.505)	(.995)	(1.041)	(.914)
Overage rate	13.502	14.37	15.24	12.211	9.236	-5.135**	-6.004**	-3.634
	(20.958)	(22.144)	(22.7)	(19.54)	(12.981)	(2.364)	(2.425)	(2.243)
Dropout rate	1.445	1.394	1.576	1.326	1.795	.402	.22	.343
	(8.248)	(7.775)	(8.287)	(7.636)	(10.977)	(.884)	(.943)	(.966)
N (schools)	1111	738	648	644	100	838	748	744

Notes: (1) The table shows the means and standard deviations of all primary schools in Salta (column 1), non-RCT schools (columns 2-4), and RCT schools (column 5). It also tests for differences between each group of non-RCT schools and RCT schools (columns 6-8). Panel A shows results for all primary school students and Panel B for grade 6 students. (2) Dropout rates should be interpreted as a upper-bound estimate, as they actually refer to the percentage of students who leave their schools without asking for a pass to another school. (3) * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.2: Comparison between in- and out-of-sample schools on students' characteristics and achievement (2017)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Out-of-sample schools						
	All schools	All	Public	Public and Salta or Tartagal	In-sample schools	Col.(5)-Col.(2)	Col.(5)-Col.(3)	Col.(5)-Col.(4)
Age	2.585	2.581	2.619	2.563	2.603	.022	-.016	.041**
	(.699)	(.698)	(.726)	(.674)	(.702)	(.019)	(.02)	(.019)
Female	.498	.5	.495	.502	.491	-.01	-.005	-.012
	(.5)	(.5)	(.5)	(.5)	(.5)	(.01)	(.009)	(.011)
Attended pre-K	5.866	5.836	5.894	5.847	5.982	.146***	.088***	.156***
	(.993)	(.973)	(1.012)	(.983)	(1.057)	(.032)	(.032)	(.034)
Repeated a grade	.155	.154	.18	.146	.162	.009	-.018	.015
	(.362)	(.361)	(.384)	(.354)	(.369)	(.011)	(.011)	(.011)
Absent 15 or more times	.121	.124	.125	.127	.107	-.018**	-.018**	-.024***
	(.326)	(.33)	(.331)	(.333)	(.309)	(.007)	(.007)	(.007)
Natural sciences score	495.528	499.385	486.262	498.453	480.984	-18.401***	-5.278	-20.06***
	(96.552)	(97.205)	(93.954)	(96.873)	(92.624)	(4.053)	(3.854)	(4.406)
Social sciences score	498.963	502.516	489.805	500.88	485.513	-17.002***	-4.292	-18.32***
	(99.211)	(99.022)	(96.777)	(98.242)	(98.78)	(5.477)	(5.325)	(5.737)
Household asset index	0	.07	-.28	.15	-.298	-.368***	-.019	-.496***
	(1.716)	(1.736)	(1.767)	(1.616)	(1.593)	(.097)	(.098)	(.099)
N (students)	23846	18832	15099	17049	5014	23846	20113	22063

Notes: (1) The table shows the means and standard deviations of all primary schools in Salta (column 1), non-RCT schools (columns 2-4), and RCT schools (column 5). It also tests for differences between each group of non-RCT schools and RCT schools (columns 6-8). Panel A shows results for all primary school students and Panel B for grade 6 students. (2) Dropout rates should be interpreted as a upper-bound estimate, as they actually refer to the percentage of students who leave their schools without asking for a pass to another school. (3) * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.3: Balancing checks on schools' internal efficiency (2017)

Variable	(1) Control	(2) Treatment	(3) Difference
A. Primary school			
Enrolled students	412.625	399.038	-7.596
	(303.821)	(266.943)	(27.766)
Passing rate	96.209	96.118	-0.083
	(5.770)	(5.279)	(1.042)
Failure rate	2.434	3.106	0.661
	(2.832)	(3.692)	(0.635)
Repetition rate	2.866	3.431	0.573
	(3.594)	(4.606)	(0.815)
Overage rate	6.211	4.869	-1.151
	(8.943)	(4.937)	(1.368)
Dropout rate	1.357	0.776	-0.578
	(5.087)	(3.309)	(0.857)
N (schools)	48	52	100
B. Grade 6			
Enrolled students	54.708	54.442	0.624
	(42.214)	(37.183)	(3.983)
Passing rate	94.791	97.466	2.681
	(15.055)	(7.089)	(2.351)
Failure rate	2.464	1.616	-0.845
	(4.857)	(2.933)	(0.784)
Repetition rate	4.544	2.953	-1.571
	(10.529)	(6.084)	(1.753)
Overage rate	10.163	8.384	-1.550
	(16.756)	(8.220)	(2.619)
Dropout rate	2.746	0.918	-1.836
	(14.547)	(6.104)	(2.294)
N (schools)	48	52	100

Notes: (1) This table compares the internal efficiency of control and treatment schools on the year before the program. It shows the means and corresponding standard deviations for each group (columns 1-2) and tests for differences between groups including randomization fixed effects (column 3). Panel A displays figures for all primary school students and Panel B only for grade 6 students. (2) * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.4: Balancing checks on students' characteristics and achievement (2017)

	(1) Control	(2) Treatment	(3) Difference
Age	2.591	2.615	0.016
	(0.687)	(0.716)	(0.031)
Female	0.499	0.482	-0.017
	(0.500)	(0.500)	(0.016)
Attended pre-K	5.983	5.981	0.004
	(1.033)	(1.080)	(0.058)
Repeated a grade	0.155	0.169	0.010
	(0.362)	(0.375)	(0.017)
Absent 15 or more times	0.100	0.114	0.015
	(0.300)	(0.317)	(0.012)
Natural sciences score	478.882	483.003	4.693
	(90.907)	(94.219)	(6.222)
Social sciences score	484.983	486.027	1.860
	(99.662)	(97.936)	(9.635)
Household asset index	-0.281	-0.314	0.043
	(1.607)	(1.580)	(0.152)
N (students)	2,458	2,557	5,015

Notes: (1) This table compares the characteristics and achievement students in control and treatment schools on the year before the program. It shows the means and corresponding standard deviations for each group (columns 1-2) and tests for differences between groups including randomization fixed effects (column 3). (2) * significant at 10%; ** significant at 5%; *** significant at 1%.