

## **Raising Academic Achievement in Under-Performing Schools through Improved Management: A Research-Practice Partnership**

University of Toronto and Puerto Rico Department of Education

University of Toronto lead: Gustavo Bobonis

Puerto Rico Department of Education lead: Damarys Valera Velez

Amount: \$650,000 (2019-2022, co-funded by the William T. Grant Foundation and the Spencer Foundation), \$350,000 (2022-2024, funded by the William T. Grant Foundation)

### **III. Joint Research Agenda to Reduce Youth Inequality**

The Puerto Rico Department of Education's (PRDE) commitment to advancing the life prospects of youth in Puerto Rico has resulted in the development of a number of initiatives based on the improvement of school and system-wide management practices. Specifically, the Department wishes to establish a systematic approach to learn from the experiences and practices of better-performing schools in the system and adopt successful strategies based on rigorous evidence—from Puerto Rico and elsewhere—to institute these in underperforming schools. These efforts are thus intended to generate their greatest impacts among the underperforming public schools and students and disadvantaged students' educational outcomes.

Improvements in management practices have been credited with generating stunning productivity improvement in the private sector.<sup>6</sup> Although their adoption in the education sector has been slow, a growing literature suggests that management-level interventions can lead to significant improvements in managerial performance and learning. Principals have the capacity to influence students' academic achievement through a variety of mechanisms. They set goals and share common expectations, oversee and improve academic operations by making instruction decisions, manage personnel, and are often responsible for their evaluation and professional support. Also, principals can influence school culture by sanctioning norms of conduct and behavior, and by promoting certain habits and mindsets (see Schleicher, 2015).

*While most ICG grants involve partnerships located in the same city or state, sometimes ICG partnerships are between institutions located in different geographic areas to facilitate a strong match between each institution's needs and assets. In this case, the University of Toronto partnered with the Puerto Rico Department of Education. Read more about this partnership [here](#) and [here](#).*

*This research agenda describes two distinct quantitative research projects, designed in response to PRDE's desire to transform schools and improve educational outcomes.*

*This section establishes the theoretical and empirical rationale for the first project: an evaluation of a professional development program for principals.*

Recent efforts to explain differences in academic performance between schools, districts, and states have focused attention on the role and influence of school principals. In assessing school principal quality and its relationship to education outcomes in New York City, Clark, Martorell and Rockoff (2009) conclude that some principal's observable characteristics, such as educational attainment, have no effect on students' academic achievement. Yet, the study finds that experience, measured in years as a principal, has a positive impact on standardized test scores in mathematics and on attendance.

The importance of principals' experience is supported by research that finds that schools affected by a departing principal have, on average, lower student achievement gains and higher teacher turnover rates (Béteille, Kalogrides, and Loeb, 2012). Following a complementary approach, research inspired

by teacher “value-added” models confirms the importance of principal quality. Using the Texas Education Agency’s statewide database, Branch, Hanushek, and Rivkin (2012) find an increase of between 0.05 and 0.21 standard deviations in student standardized test scores for each standard deviation difference in a measure of school principal quality. Likewise, using a similar approach with data from the province of British Columbia, Dhuey and Smith (2014) find effects of 0.29 and 0.41 standard deviations on reading and mathematics test scores, respectively.

Perhaps the most compelling evidence to date on the importance of school principals is found in the seminal work of Bloom, Lemos, Sadun, and Van Reenen (2015). These researchers adapted the World Management Survey, a tool designed to measure management skills in the private sector, to the reality of the public sector and applicable to the field of education. The survey gathers information regarding twenty (20) school management practices that fall under the following categories: (a) operations (establishment of curricular planning mechanisms, personalization of teaching, and using data to guide instruction); (b) monitoring (overseeing, analysis of findings, and decision-making based on performance); (c) target (goal setting that is relevant, specific and clear for the whole school community); and (d) personnel management (conducive to attraction and retention of the best available resources). The World Management Survey for the education sector was conducted in eight (8) countries and more than 1,800 schools, allowing the researchers to construct an international school management index.

The results of the field work showed both clear differences in management practices between and within countries, and also a suggestive association between management skills and student outcomes. Bloom et al. (2015)’s analysis estimates that a one standard deviation increase in the school management index is associated with a 0.23 standard deviation increase in students’ standardized test scores. Another significant finding is the identification of two determinants of better scores in the managerial index: a school’s governance structure and the leadership capacity of the school principal. These results suggest that principals working under clear accountability schemes, and with appropriate performance evaluation systems, tend to obtain better results in the management index. A similar relationship is found when principals have the opportunity and the ability to effectively lead the school (i.e., contribute to the teaching-learning component, make decisions about the use of resources, determine vocational training needs). Cases in which the school director has greater leadership capacity are positively correlated with higher scores in the management index.

*Research conducted in partnerships are valuable for answering relevant to practitioners and policymakers. For the ICG application, proposed research studies should also be grounded in and add to relevant literature.*

Rigorous evidence on the impact of professional development programs and training focused on school principals is scarce. To date, there is only one study with an experimental design that identifies the causal impact of this type of intervention (Fryer, 2017). In the city of Houston, 29 school principals were randomly selected to participate in a two-year intensive training program (300 hours of training), focusing on three areas: instructional planning, data-driven instruction, and classroom observation and feedback. After two years, teachers in treatment

schools received detailed feedback on classroom work every two months, compared to once a year in control group schools. In terms of impact on outcomes, one year after participation in the training program, students from treatment schools outperformed their peers in control group schools by 0.10 standard deviations in a composite index of high-stakes tests of language and mathematics. Results in a low-stakes test index of language, mathematics, natural sciences, and social sciences were also positive, with an estimated effect of 0.19 standard deviations in favor of the treatment group.

This body of research points to the possibility that promoting *long-term* sustained improvements in school management practices is an amenable area for policy intervention, given the need to identify ways to achieve this *sustainably* and in *large scale* in the context of government-run schools. Both dimensions – scalability and sustainability – are very relevant to the context of this project, given evidence that training effects in the public sector is often not sustained in the long run. For example, Fryer (2017) reports that one year after the principals’ training program, students from treatment schools outperformed their peers in control group schools in academic achievement, but that the gains were not generally sustained in a subsequent year.<sup>7</sup> In a managerial upgrading experiment focusing on primary health centers in Nigeria, for example, Dunsch et al. (2017) find that a management consulting treatment improved practices substantially in the short run, but that these were effectively abandoned after a year. A similar management intervention involving manufacturing firms in India, however, found that some practices adopted in the short run were in fact retained, though there was a preference for “visible” practices (Bloom et al., 2017).

Evidence points to the difficulty of changing behavioral patterns as well as the importance of reminders, monitoring, and external reinforcement (Frey and Rogers, 2014). One type of intervention that has recently been used successfully to improve external reinforcement involves the use of text messages to nudge behavior of key individuals.<sup>8</sup>

### ***A. Improving and Sustaining Management Practices in Public Schools***

Based on this understanding of the importance of management and in alignment with the agency’s strategic goal of promoting the professional development of school leaders, PRDE will implement Puerto Rico Principals’ Academy, a *large-scale* management training program that will impact all PRDE school directors over a three-year period. The Academy will provide school principals basic management training in Fall 2018, focused on administrative and operations content (24 hours of professional development). Beginning in Spring 2018, and continuing for the next three academic years, randomly selected cohorts of approximately 33 percent of school principals will participate in an *intensive* management training focused on personnel management, instruction, planning and target setting, monitoring, school culture, and personal leadership (140 hours of professional development plus monthly follow-up support throughout the academic year). The Academy will also train central and regional office staff who will have the responsibility of overseeing and supporting school principals’ work. The training will be conducted by a high-quality provider with a great deal of experience in training and supporting school principals through the use of a curriculum based on impacting: learning and teaching, school culture, talent management, planning and operations and personal leadership. The curriculum closely parallels the categories studied by Bloom et al. (2015).

*The applicants do a nice job of explaining and contextualizing the intervention: the Puerto Rico Principals’ Academy.*

*Relevance and rigor are not mutually exclusive. In this case, the rigorous design is an impact evaluation of an intervention's effectiveness. All types of research designs – from qualitative to quantitative to mixed methods – can be rigorous. What matters, as illustrated in this proposal, is that the proposed design answers the questions asked and provides key details about the plans for data collection and analysis so that reviewers can judge the rigor and feasibility.*

We propose an impact evaluation of *Puerto Rico Principals' Academy* – via a randomized controlled trial – that will lead to a rich understanding of ways to improve principals' quality, understood in terms of management skills and practice, and their effects on improving the academic outcomes of students in the public school system. The evaluation project will collect detailed longitudinal data on managerial practices across academic years using an internationally validated instrument, the Development World Management Survey methodology (Lemos and Scur, 2017; Bloom et al., 2015) at baseline and at the end of each academic year over the course

of the study, with the aim of identifying the impact of the training program on managerial practices. Understanding how principals manage operations, set monitoring practices, plan and target-set, and manage personnel, will allow policymakers and researchers to identify the levers through which improved results might be achieved. In addition, we will use existing school (principal, teachers) administrative performance and student-level standardized test scores data to evaluate effects on measurable school inputs and the *distribution* of students' academic achievement.

PRDE's commitment to invest in personnel training offers an unparalleled opportunity to learn about the effectiveness of professional development programs targeted at school principals and their potential benefits on students. By introducing a rigorous evaluation strategy to the program design, PRDE will benefit from learning if the Principal's Academy program works and why, while also promoting a culture of evidence-based policy making and transparency of results within the agency. From an academic standpoint, the literature on this topic is incipient, so far suggesting that (1) it is possible for schools to adopt good practices, (2) that these practices have effects on student outcomes, but (3) that sustaining the *practices over the long term* is a challenging endeavor. To the best of our knowledge, no project to date focuses on understanding the mechanisms behind adoption and retention of best practices. We aim to add to this research with a focus on opening the black box of management technology adoption, evaluating the contexts that favor organizational change. This evaluation project will thus be a path-breaking contribution in our understanding of whether training programs can be successfully scaled up to a system-wide setting and whether improved practices can be sustained over the long term. The combination of rich data together with the use of a rigorous evaluation will enable both policymakers and researchers to answer high-priority questions about the initiative's effectiveness and how they help close gaps in student outcomes.

### ***Description of the Intervention***

PRDE's Principal's Academy program has four components:

*Basic Management Training:* All school principals in Puerto Rico will participate in a program that will involve 24 hours of introductory training, delivered in six 4-hour workshops. Content will be delivered by a provider selected by the PRDE with input from the research team.

Workshops will provide instruction in: complying with local and federal regulations, monitoring operations and administrative systems, developing and maintaining data systems, developing and managing all schedules, managing time efficiently, and leading effective meetings. Principals will receive these trainings in cohorts of 30-35, organized by school and geographical area. The purpose of this training is to ensure that all principals have basic tools to effectively manage schools in the near-term, as more intensive training will roll out over several years.

*Intensive Management Training:* Principals will participate in an intensive management training program that will be delivered across three academic years beginning in Spring 2018. Participants will be randomly assigned to one of three treatment arms, following a stratified sample selection by school type (i.e., vocational, elementary, high school) and school region. Approximately one third of PRDE’s principals will participate in the first intensive training cohort, followed in subsequent academic years by cohorts of similar size.

The program will consist of 80 hours of training over two weeks. The selection of topics of discussion will be informed by the results of the implementation of the Development World Management Survey by the research team. These will likely include one or several of the following: personnel management, instruction (teaching and learning), planning and goal setting, monitoring, school culture and personal leadership. An additional 60 hours of training will be delivered in workshops during the academic year to help reinforce concepts, address challenges, and make actionable adjustments in schools.<sup>9</sup>

*Train the Trainer:* the provider will train approximately 35 staff members from PRDE’s central and regional offices with the objective of providing them the tools and developing the skills to coordinate, monitor, and coach school principals being trained in the Academy. These staff members will participate in the regular school principal workshops and will also receive specific training in accordance to their expected roles. Trainers will have the responsibility to provide support to school principals as the program is rolled out.

*Adoption Reminders:* Principals in the treatment groups will also receive targeted messages reminding them of the management practices as well as tailored tips based on school management baseline data. All others will receive a placebo message. Depending on the connectivity at the time of implementation, we will experiment with individualized text messages or group text messages that will also allow measurement of network formation and dissemination of practices.

### ***Evaluation Methodology***

The project will rigorously evaluate PRDE’s Principal’s Academy to accurately measure its impact on a set of key outcome variables. Researchers will conduct both interim and long-term evaluations to help inform program design, estimate intervention effects, and understand underlying mechanisms that explain observed results.

#### ***Survey of Managerial Practices***

To obtain a comprehensive baseline of school leaders’ attitudes and managerial practices, we will survey a representative sample of 400 school principals in the island using the Development World Management Survey (D-WMS). The D-WMS expands the original WMS adapted by Bloom et al. (2015) in three important ways. First, it identifies three different management activities – implementation, usage and monitoring – that should be considered in assessing the strength of management practices within schools. Second, it maps these activities into the original surveys’ questions, resulting in an expansion of scored items. This reduces measurement error and allows for more precise interpretation of findings. Third, the D-WMS also introduced a modification in the scoring system. By allowing interviewers to assign “half-scores” within the original scale, much finer data is collected (Lemos and Scur, 2017). These characteristics make the D-WMS a very relevant and useful tool for the Puerto Rican context.

We will partner with a local firm or organization with a proven surveying track record to conduct face to face or phone interviews across the island. The D-WMS content will be adapted, piloted and tested to guarantee that it is context-sensible and correctly understood by both school principals and enumerators. After the survey is conducted, each principal will obtain a

management score and results will also serve to identify strengths and weaknesses across the island in operations, monitoring, target setting, and people management.

This information will be highly valuable for designing the curriculum of the intensive training program and will be merged with school level data to analyze correlations between manager quality and school quality, as in Bloom et al. (2015) and Lemos and Scur (2017). This exercise will inform which aspects of management quality are most important in the context of Puerto Rico's schools. The D-WMS will also be conducted annually to track changes in school management practices induced by the Principal's Academy program.

### *Phase-in Randomized Control Trial*

Exploiting the cohort-by-year program design, we will use a phase-in randomized control trial methodology to identify the causal effects of the intensive management training on a variety of student and school outcomes. To guarantee balance among groups on key variables, increase statistical power, and allow analysis of impacts by subgroups, randomization of school principals into treatment arms will be stratified by school type, the use of management practices at baseline, and geographic region.

*Randomization:* Approximately 819 PRDE schools and their principals will be stratified by type of school, academic district/region, and adoption of modern management practices at baseline. All schools within each stratum will be randomly assigned to one of three (3) experimental arms:

*The proposal describes the evaluation design, clearly explaining treatment and control arms. It provides important information about sample size and data collection. In this case, the evaluation includes all school principals in Puerto Rico, as noted above. If a proposed study employs a sampling framework, please provide a rationale for the framework: what determines the number and type of cases?*

*Treatment Arm 1* – Intensive Management Training (IMT) (Spring 2019, AY 2019-20) + Adoption Reminders (AYs 2019-20, 2020-21)

*Treatment Arm 2* – IMT (Fall 2020, AY 2020-21) + Adoption Reminders (AY 2020-21)

*Control Arm* – Status Quo (Placebo)<sup>10</sup>

Each of the arms of this stepped-wedge design will be composed of 273 schools/principals, with principals in all schools receiving the IMT by the end of the third year.

### *Data*

*Administrative Records:* To understand the impact of the treatments on principal, teacher, and student performance, we will obtain access to PRDE administrative records, including student enrollment, demographics, grades, and standardized test scores (Grades 3-8, 11). Teacher-level data will include credentials, experience, position, college major, and attendance record. School-level information will consist of course catalogue, size and type of school, budget, and technological infrastructure.

*Principal Tests, Time Use Diary, and Other Data:* We will also employ several strategies for measuring the extent of implementation of improved management practices. First, to test for basic content learning, we will conduct pre-post surveys for each of the trainings, including the ones carried out as part of the basic management training. Second, considering the training will promote administrators to maintain database systems that track teacher evaluations, teacher performance indicators, physical condition of schools, we will use this information to collect data on the consequences of improved management practices. Third, on two randomly selected

days each term, principals will be asked to complete a diary of time use to measure time allocation to distinct management tasks.

### *Analysis*

We will quantify effects on several measures of school quality and student performance by comparing school, principal, and student-based longitudinal outcomes. Inferential data analysis will rely on comparisons of means and distributions for outcomes of interest at baseline and at the various points during follow-up across each of the treatment and control arms.

The following results will be reported:

- By the end of academic year 2018-19, by comparing Treatment Arm 1 with Treatment Arm 2 and the Control Group combined, we can estimate the effect of one semester-long intensive management training.
- By the end of academic year 2019-20, by comparing Treatment Arm 2 and the Control group, we can estimate the effect of one year-long intensive management training.
- By the end of academic year 2019-20, by comparing Treatment Arm 1 and the Control group, we can test for post-intervention persistence or fade-out effects.

*Notice here that the applicant addresses how the analyses will provide evidence on reducing inequality. Usually, applicants should specify for which groups: that is, will analyses show that outcomes are improved for a groups or groups, or will analyses show inequality is reduced for one group in comparison to another?*

Stratification of treatment by quality of management practices at baseline will contribute to an accurate assessment of potential heterogeneity in responses to the intervention. An answer to the question of the type of school that benefits most from management training is important, as it will determine whether training reduces or increases inequality in school-based outcomes. That is, inequality can be expected to decline if underperforming schools benefit disproportionately from the

intervention, but it can increase if better-performing schools benefit most. However, a definitive answer to the question of the initiative's impact on (between and within-school) inequality in outcomes will be based on distributional differences between treatment and control group students.

Interim evaluations will focus on outcomes such as student and teacher attendance, principal turnover, teacher turnover, share of teachers submitting lesson plans, frequency of observation and feedback, frequency of data analysis meetings with teachers, physical conditions of the school, and community engagement activities. Longer-term analyses will use assessment data to track test score gains that resulted from the principal training, and use interim outcomes, along with data from the D-WMS, to explain contributing factors.

The randomization of treatment will allow us to not only identify causal estimates of the effect of the program, but also to understand what drives differences in the program impact across principals. Specifically, our design will achieve an unbiased estimation of *intent-to-treat* (ITT) program impacts. In order to minimize bias in ITT impacts due to sorting of students, principals, or other staff, we will focus the analysis on pre-existing students and keep individuals as part of the school assignment at baseline even if they move schools. In addition, we will study heterogeneous impacts, stratifying on baseline characteristics (with endogenous stratification corrections, as in Abadie et al., 2017).

*Power Calculations:* We have conducted power calculations for students' subject-specific proficiency rates based on standardized test scores, by academic level (i.e., Grades 3-8, Grade 11), using school-level administrative data for AY 2015-16 & 2016-17. Considering the cluster-randomized design with measurements of student test scores at baseline and end-line, the power calculations (power=0.9,  $\alpha=0.05$ ) suggest that with 237 elementary/middle schools per arm (assuming equal treatment and control group sample sizes), it will be possible to detect a year-specific increase in Grades 3-8 students' academic proficiency in Spanish, English, and Math scores of respectively  $0.11\alpha$ ,  $0.10\alpha$ , and  $0.09\alpha$ . This is similar to existing short-term impacts of principal training programs in the US using student-level outcomes (e.g., Fryer, 2017, Gates et al., 2014). Comparable MDEs for Grade 11 test scores are somewhat larger – in the range of  $0.27\alpha$  (in English) to  $0.31\alpha$  (in Spanish) – given the more limited number of high schools (61 high schools per experimental arm). We expect to gain statistical power at both grade level groups since we will have access to individual students' (continuous) standardized test scores as opposed to only (binary) proficiency levels, individual-level baseline outcomes, and a larger sample of schools in the control group in Year 1 [e.g., 474 (= 273 x 2) K-8 schools and 122 (= 61 x 2) high schools].

*Where appropriate, please provide a formal power analysis. Reviewers will want to see that your study is adequately powered to detect plausible effects that are well-justified. Finally, make sure power is calculated at the appropriate level (e.g., the level of random assignment).*

### **Significance**

A phase-in randomized control trial methodology offers an unparalleled opportunity to conduct a rigorous impact evaluation of the Principal's Academy program to evaluate it during a three-year period. There are a number of benefits associated with this evaluation. First, PRDE will have clear, transparent findings about the program's effectiveness. It will learn if this investment works, and why. Second, results from interim and long-term evaluation results will help inform program design, in particular, and decision making within the agency in general. This will help advance a culture of evidence-based policy making and transparency of results. Third, results will contribute to a growing knowledge base about effective education policy interventions and will provide valuable insights on a field that remains understudied with the use of experimental designs. Fourth, this work will provide meaningful benefits to the residents of Puerto Rico, whose children, families, and schools have suffered from both the financial crisis and the recent devastation of Hurricane María. Finally, the large-scale nature of the Principal's Academy initiative in terms of participating schools and principals will allow the assessment of the question of whether improved managerial practices can reduce between-school inequality in school outcomes, a phenomenon that can be expected if training programs have disproportionately beneficial impacts on schools with poorer management practices. Similarly, these initiatives may reduce within-school inequality in the distribution of student outcomes. This may be the case if best managerial practices have larger impacts among underperforming students who, for example, may rely more on schools than family environments for learning.



## ***B. Understanding the Impacts of School Consolidation Policies***

*As noted above, this partnership proposed two distinct studies. Applicants should work closely in partnership to determine what research questions and analyses are most relevant to support the use of research evidence to improve youth outcomes.*

As a result of out-migration and declining fertility, public school enrollment in Puerto Rico has declined sharply from a peak reach in 1980 (Ladd and Rivera Batiz, 2006). Alongside the diminishing number of students in the island, a concomitant reduction of available funds has also stricken the system. One response by current and past PRDE administrations has been the consolidation of public schools. Albeit unpopular, school closures can free up resources that can be channeled by school system managers and school principals to more

effective uses. In fact, large scale school closings have taken place in Chicago, Pittsburgh, Philadelphia, Detroit, and Buffalo following this theory of action. While many have occurred as a result of low performance and guidelines under the No Child Left Behind (NCLB) and ESSA Acts, infrastructure underutilization has been a significant driver in many contexts (Engberg et al., 2012). However, school closures entail disruption costs among students, parents, and teachers, and evidence of its impact on student level outcomes is mixed, as effects seem to be contingent on variables such as the quality of the consolidated and receiving schools, and the percentage of students and teachers that are relocated to the same school (Brummet, 2014, Engberg et al., 2012, de la Torre and Gwynne, 2009).

This decision to consolidate schools is aligned with the PRDE's strategic goal of reducing the size of the system and making a more effective use of resources. Since 2010 the PRDE has closed more than 700 schools, with school closings occurring almost every single academic year. Fiscal constraints and shifting demographics that left many schools significantly underutilized have been driving these decisions. Despite the magnitude of the policy, the impacts on the distribution of students' academic achievement remain unknown.

Research on the impacts of large-scale school closures has often found little if any effects on student achievement outcomes, on average, with small negative or positive effects depending on the circumstances of the closure and the characteristics of the student (Brummet, 2014; de la Torre and Gwynne, 2009; Engberg et al., 2012). Nonetheless, teachers and parents remain wary of decisions to consolidate underpopulated schools. Qualitative studies report emotional and academic challenges faced by students displaced by school closings (Kirshner et al., 2010). The resulting uncertainty about the potential effects of school closings on education outcomes makes the task of conducting a rigorous evaluation of the school consolidation process in Puerto Rico a highly relevant endeavor.

There are several potential channels through which students may be affected by school consolidations. Understanding the relative importance of each of these channels will provide a more complete picture of the costs and benefits of closing schools, aside from the clear fiscal gains arising from more efficient infrastructure usage. Further, insights into the challenges and advantages associated with school consolidations will aid the PRDE in facilitating smooth and constructive transitions for students when making decisions surrounding future closures.

Significant prior research has analyzed various outcomes associated with school closures, often with mixed results. The most apparent consequence of closing a school is the disturbance to students displaced by the closure. Brummet (2014) as well as de la Torres and Gwynne (2009) analyze school closures in Michigan and Chicago, respectively. Both studies find that being displaced by a school closing had no lasting effects on student test scores. One might expect displacement to affect students through adjusting to new routines, higher transportation costs, and changes to peer and teacher networks. While they do not isolate the influence of routines and transportation, these studies do find evidence that relocation to higher performing schools improved student test scores markedly. Using data on test scores and student and teacher relocations, our analysis will similarly quantify how peers as well as teachers may ease the transition for displaced students. Supplementing these data with information gathered from interviews will shed light on how students adapt to new routines and environments. In a similar vein, an important aspect of our evaluation will involve looking at students and teachers in receiving schools to understand how they adapt to the influx of new students.

*Again, the proposal situates the second study in existing research. While the ICG application is complex, bear in mind that review committee members are also gauging the proposed research for its potential contribution to existing scholarship. Make a strong case that your research is shaped by and will contribute to academic literature.*

Another clear impact of school consolidation is changes to school and class sizes. While intuition might dictate that smaller class sizes improve instructional quality by affording greater attention to individual students and more effective classroom management, research quantifying these benefits lacks consensus. Disagreement also exists in research on school size, with some studies citing student alienation in larger schools (Strang, 1987) and others quantifying large economies of scale in school size with the optimal high school having around 1,500 students (e.g., Chakraborty, 2000; Colegrave et al., 2008). Recent reports on developing countries highlight the contextual nature of class size benefits (Woessmann and West, 2006). Perhaps most relevant to the Puerto Rico setting, Coupé et al. (2015) investigate large-scale school closures in Ukraine driven by demographic shifts. They find that student achievement did not suffer from increased class and school sizes resulting from the consolidation of underpopulated schools. Detailed enrollment data (especially at the class or teacher level) will allow us to evaluate the importance of these channels in Puerto Rico's school consolidations.

Lastly, a less obvious potential benefit of school closures is improvements in facilities and infrastructure. Studies on developing countries, and on Latin America and the Caribbean in particular, have found significant benefits of school facility quality on student achievement and attendance (Cuesta et al., 2016). It is possible that relocating students and teachers from older buildings to newer ones and concentrating resources for facility maintenance could boost both student and teacher performance by improving attendance, enthusiasm, and focus. Analysis of school closures will evaluate the role of changes to the physical school environment in affecting student outcomes. Understanding the role of education facilities will be especially important going forward as the PRDE makes decisions around infrastructure damage caused by recent hurricanes.

The second project in this collaborative agenda will evaluate the impact of school closings on student achievement and a range of other outcomes, taking advantage of newly-available administrative longitudinal data linking information on students, teachers, and schools. Results stemming from this research will help shed light on the effects of a policy that remains highly controversial, and will help inform decision-making processes involving potential school closures in the future. The research design will allow the study of peer effects, differences in effects on students being relocated by the closure and students in receiving schools, and influence on siblings unaffected directly by the policy. As such, factors that can help minimize negative impacts on student learning will be identified so that adequate policy measures can be tailored to address any challenges observed in particular schools or settings. Equally important will be the assessment of potentially positive effects associated with transferring students to better-performing schools as a strategy for improving academic outcomes and reducing inequality in their distribution.

### ***Evaluation Methodology***

#### *Data*

The quantitative analysis, based primarily on the PRDE Data Warehouse Databases, contains detailed information on PRDE students, teachers, and schools. Specifically, the database contains longitudinal student-level outcomes such as: date of birth, gender, ethnicity, attendance, special education/needs status, poverty status, parents' names and addresses, school assignment, class schedules, attendance, transcripts, standardized tests scores (Spanish, English, Mathematics, for Grades 3-8, Grade 11; Science, Grades 8 and 11), and disciplinary incidents. Information on teachers includes class schedules, attendance, and professional background, such as university degree(s), major, and the name of the university that granted the degree(s). School-related variables include geographical location, course catalogs, number of classrooms, NCLB/ESSA Acts accountability status, annual measurable goals, graduation and retention rates, among others. Database relationship keys allow student, teacher, and school records to be linked to all others, so that students can be linked to their teachers, who can be connected to the schools where they work. Unique staff, student, and school identification numbers also allow records to be linked across academic years.

*Across all proposed studies associated with this RPP, the team is clear about the data sources, variables, and outcomes.*

School-level data and student-level data aggregated at the school level will be matched to publicly-available municipality-level and neighborhood-level 2000 and 2010 Puerto Rico Census of Population and local government data providing information on the socio-economic of the local population and local government characteristics of surrounding geographical areas. This will be used to study the geographic correlates of specific school consolidation decisions.

#### *Analysis*

Using the data described above, we will analyze the frequency of school consolidations, as well as school, neighborhood, and municipality-level correlates of specific school consolidations. This will help contextualize the research question, by using empirical evidence to understand the criteria and correlates that informed school consolidation decisions across the distinct academic/fiscal years.

Following the descriptive analysis, the Research Team will turn to the primary issue of determining the impact of school closures on a range of student outcomes of children in consolidated schools, as well as these outcomes on students in receiving schools, such as:

- distance to school (a measure of cost of going to school);
- decisions to drop out of public school or the public education system;
- school attendance;
- academic performance as reflected by student transcripts;
- Spanish, English, Mathematics (and Science, when available) standardized tests scores;
- number and gravity of discipline incidents;
- grade promotion;
- requests for specialized education plans;
- persistence to high school graduation;
- curriculum choices (among high school students, given the potential impact on college access).

Similarly, the Research Team will study the indirect consequences of the school consolidations on children attending receiving and non-receiving schools who are siblings of students directly impacted by the school consolidations. This will allow the researchers to measure intra-family spillovers effects of the relocation of students across schools.

Finally, the Research Team will study specific outcomes for relocated teachers' performance:

- decisions to quit/leave the work in the public education system;
- teacher attendance;
- academic performance of students assigned to these teachers (as measured by student transcripts and standardized test scores).

To arrive at estimates of these effects, the Research Team will rely on multivariate regression models that regress the various outcome variables on: (a) measures of the incidence or size of influxes of displaced students and teachers into receiving schools or measures of other factors that can smooth or disrupt transitions among sending and receiving school students; (b) a set of variables indicating the time that has elapsed (if any) since a student or teacher changed schools due to a closure; and (c) student, parental, and school characteristics relevant to the determination of the outcomes in question. Under certain statistical assumptions, the coefficient estimates from the school closure indicators can reflect causal effects, to the extent that they quantify the difference between observed outcomes and those that would have been observed in the absence of a school transition due to a school closure.

In addition, we will examine the distributional effects of the policies by constructing robust semi-parametric counterfactuals of the distribution of students' educational outcomes employing the methods developed in the distribution decomposition literature (DiNardo, Fortin, and Lemieux, 1996; Fortin, Lemieux, and Firpo, 2010). Convincing counterfactuals can be obtained as a result of the large number of controls group individuals and schools that are contained in the administrative dataset, the rich set of characteristics for regression-adjustment in the empirical analysis, as well as the fact that the longitudinal nature of the data allows extrapolation of student and teacher performance on the basis of their characteristics and outcomes in the years prior to a school closure.

### ***Significance***

The Puerto Rico Department of Education will continue to face several fiscal, operational and administrative challenges in the near future, all of which will be exacerbated by the devastation brought upon by the passing of hurricanes Irma and Maria. Closing underpopulated, damaged or chronically underperforming schools will remain a part of the policy alternatives the department will have at hand to achieve its goal of reducing the size of a system which is highly inefficient. The results of this study will help in understanding the effects of school closings and,

consequently, provide insights on the best way to select schools for future consolidations. It will also allow for identifying critical areas in which future interventions can help smooth the potential challenges associated with closures.

Finally, the qualitative component of the study will also create positive spillovers for the government by providing information about students', teachers' and parents' experiences in a context of recovery after the passing of the storms. This information can be used by the PRDE or other local agencies to design specific interventions aimed to improve the lives of the communities that have been most affected by these natural disasters.

### ***C. Other Proposed Evaluations***

In addition, the partners will establish a systematic approach to develop additional joint research- practice projects to inform Department practices aimed at improving the school achievement of underperforming schools. For example, the territory recently enacted a new

*Not all strong proposals must describe every study or project the RPP will undertake together. This proposal clearly lays out two studies but acknowledges they will develop additional studies as needed by PRDE. The partnership later pursued three additional studies on management training, computer assisted learning, and teacher professional development.*

education reform law making the introduction of charter schools a key component of the strategy aimed at raising students' academic achievement and reducing inequality in the educational outcomes of low- income public school students.<sup>11</sup> PRDE is thus interested in monitoring the performance of charter schools with respect to students' educational outcomes. Conversations have begun with the research team to design an impact evaluation of the pilot program together with the first charter school in the territory (Vimenti School). The partnership will help support an effective research-practice collaboration for this and other Department evaluation initiatives.

- 
- 5 The Research Teams shall submit to PRDE for review and comment any Publications intended for publication, release, and/or dissemination. PRDE will have fifteen (15) days for its review of the Publication to ensure that no Data is released that permits direct or indirect identification of any individual. PRDE will make every effort to work with the Research Team to resolve outstanding questions.
  - 6 See Bloom, Lemos, Sadun, Scur and Van Reenen (2014) for a survey.
  - 7 The average effects during the second year were statistically indistinguishable from zero. Fryer (2017) documents that the differentiated findings across years can be partially explained by principal turnover and lack of fidelity in the implementation of the program.
  - 8 See Frey and Rogers (2014). Text messages have been used to improve kindergarten children's development level (Doss et al., 2017), preventing dropouts (Kraft and Rogers, 2015), and improving parent engagement (Bergman and Chan, 2017).
  - 9 In order to maximize treatment effectiveness, the Research Team and PRDE will jointly develop plans to (a) create a state-of- the-art curriculum that is appropriate for the PRDE context; (b) create mechanisms for selecting a high-quality and proven service provider; and (c) invest in PRDE and Research Team capacity to monitor compliance with the intervention.
  - 10 The Control Arm will receive Intensive Management Training during AY 2020-21.
  - 11 Greater accountability and flexibility in hiring policies and other practices can make charter schools more effective instruments for increasing learning through improved management. For instance, Bloom et al. (2015) show that this is indeed the case, in establishing that better management shown by autonomous government schools like UK academies or US charters, relative to private or regular public schools, is associated with greater accountability to stakeholders and leadership on the part of school principals.

## References

- Abadie, A., Chingos, M.M., and West, M.R. (2017). "Endogenous Stratification in Randomized Experiments." Unpublished manuscript, Harvard University.
- Abdulkadiroglu, A., Angrist, J., Dynarski, S., Kane, T.J., and Pathak, P. (2011), "Accountability in Public Schools: Evidence from Boston's Charters and Pilots", *Quarterly Journal of Economics*, 126(2), 699-748.
- Bambrick-Santoyo, Paul (2012). *Leverage Leadership: A Practical Guide to Building Exceptional Schools*. San Francisco: Josey-Bass.
- Bates, M.A., and Glennester, R. (2017). "The Generalizability Puzzle", *Stanford Social Innovation Review*. Summer 2017, 50-54.
- Bergman, P. and Chan, E. (2017). "Leveraging Parents through Technology: The Impact of High-Frequency Information on Student Achievement." Unpublished manuscript, Columbia University.
- Béteille, T., Kalogrides, D., and Loeb, S. (2012). "Stepping stones: Principal career paths and school outcomes", *Social Science Research*, 41, 904-919.
- Bloom, N., Lemos, R., Sadun, R., Scur, D., and Van Reenen, J. (2014). "The New Empirical Economics of Management", *Journal of the European Economic Association*, 12(4), 835-876.
- Bloom, N., Lemos, R., Sadun, R., and Van Reenen, J. (2015). "Does Management Matter in Schools?", *The Economic Journal*, 125(May): 647-674.
- Bloom, N., Mahajan, A., McKenzie, D., and Roberts, J. (2017). "Do Management Interventions Last? Evidence from India." Unpublished manuscript, Stanford University.
- Branch, G. F., Hanushek, E. A. and Rivkin, S. G. (2012). "Estimating the effect of leaders on public sector productivity: The case of school principals." NBER Working Paper No. 17803.
- Brummet, Q. (2014). "The Effect of school closings on student achievement.", *Journal of Public Economics*, 119, 108-124.
- Chakraborty, K., Biswas, B., and Lewis, W. C. (2000). Economies of scale in public education: an econometric analysis. *Contemporary Economic Policy*, 18(2), 238-247.
- Chan, J.Y., Lemanski, N., Perkins, R., Katsberg, D., Roey, S., and Ruess, R. (2014). Performance of Puerto Rico 15-Year-Old Students in Mathematics, Science, and Reading Literacy in an International Context. Programme for International Student Assessment 2012.
- Clark, D., Martorell, P. and Rockoff, J.E. (2009). "School Principals and School Performance." Working Paper No. 38, Center for Analysis of Longitudinal Data in Education Research.
- Coelli, M., and Green, D.A. (2012). "Leadership effects: school principals and student outcomes.", *Economics of Education Review*, 31(1), 92-109.
- Colegrave, A. D., and Giles, M. J. (2008). "School cost functions: A meta-regression analysis." *Economics of Education Review*, 27(6), 688-696.
- College Board (2013). The 2013 SAT Report on College & Career Readiness.
- College Board (2018). Tendencias en los promedios de las pruebas de razonamiento (PAA) y aprovechamiento académico (PACH) de la Prueba de Evaluación y Admisión Universitaria (PEAU) para las clases graduandas 2006-2016 y la administración de la PAA revisada en diciembre de 2017 (datos preliminares). Internal Report.
- Coupé, T., Olefir, A., and Alonso, J. D. (2015). "Class Size, School Size and the Size of the School Network." *Education Economics*, 24(3), 329-351.

- Cuesta, A., Glewwe, P. and Krause, B. (2016). "School Infrastructure and Educational Outcomes: A Literature Review, with Special Reference to Latin America.", *Economía*, 17(1), 95-130.
- De la Torre, M., and Gwynne, J. (2009). "When Schools Close: Effects on Displaced Students in Chicago Public Schools." Consortium on Chicago School Research, Chicago, IL.
- de Hoyos, R., Lemos, R., and Scur, D. (in progress). "How to Improve Management: Evidence from Mexico on Different Delivery Approaches."
- DiNardo, J., Fortin, N. M., and Lemieux, T. (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach.", *Econometrica*, 64(5), 1001-44.
- Dhuey, E., and Smith, J. (2014). "How Important are School Principals in the Production of Student Achievement?" *Canadian Journal of Economics*, 47(2), 634-663.
- Dobbie, W., and Fryer, R.G. (2013), "Getting Beneath the Veil of Effective Schools: Evidence from New York City.", *American Economic Journal: Applied Economics*, 5(4), 28-60.
- Doss, C., Fahle, E., Loeb, S., and York, B. (2017). "Supporting Parenting through Differentiated and Personalized Text-Messaging: Testing Effects on Learning During Kindergarten." CEPA Working Paper No. 16-18.
- Dunsch, F.A., Evans, D.K., Eze-Ajoku, E., and Macis, M. (2017). "Management, Supervision, and Health Care: A Field Experiment." NBER Working Paper No. 23749.
- Engberg, J., Gill, B., Zamarro, G., and Zimmer, R. (2012). "Closing schools in a shrinking district: do students outcomes depend on which schools are closed?", *Journal of Urban Economics*, (71)2, 189-203.
- Fortin, N., Lemieux, T., and Firpo, S. (2010). "Decomposition Methods in Economics." NBER Working Paper 16045.
- Frey, E., and Rogers, T. (2014). "Persistence: How Treatment Effects Persist After Interventions Stop." *Policy Insights from the Behavioral and Brain Sciences*, 1(1), 172-179.
- Fryer, R.G. (2014). "Injecting Charter School Best Practices into Traditional Public Schools: Evidence from Field Experiments.", *The Quarterly Journal of Economics*, 129(3), 1355-1407.
- Fryer, R.G. (2017). 'Management and Student Achievement: Evidence from a Randomized Experiment', NBER Working Paper 23437.
- Gates, S.M., Hamilton, L.S., Martorell, P., Burkhauser, S., Heaton, P., Pierson, A., Baird, M., Vuollo, M., Li, J., Lavery, D.C., Harvey, M., and Gu, K. (2014). *Preparing Principals to Raise Student Achievement: Implementation and Effects of the New Leaders Program in Ten Districts*. Rand Corporation.
- Hanushek, E. A. (2003). "The Failure of Input-Based Schooling Policies.", *The Economic Journal*, 113(485), F64-F98.
- J-PAL: The Abdul Latif Jameel Poverty Action Lab (2017). Ev-Labs: Evidencia Científica e Innovación para Políticas Públicas de Alto Impacto. Internal Document.
- Kraft, M.A., and Rogers, T. (2015). "The Underutilized Potential of Teacher-to-Parent Communication: Evidence from a Field Experiment.", *Economics of Education Review*, 47, 49-63.
- Kirshner, B., Gaertner, M., and Pozzoboni, K. (2010). "Tracing Transitions: The Effect of High School Closure on Displaced Students.", *Education Evaluation and Policy Analysis*, 32(3), 407-429.
- Krueger, A. B. (2003). "Economic Considerations and Class Size.", *The Economic Journal*, 113(485), F34-F63.
- Ladd, H.F., and Rivera-Batiz, F.L. (2006). "Education and Economic Development." Chapter 6



- in: Collins, S.M., Bosworth, B.P., Soto-Class, M.A. (Eds.): *The Economy of Puerto Rico: Restoring Growth*. Brookings Institution Press.
- Lemos, R., Muralidharan, K., and Scur, D. (2017). "School Management and Productivity in the Public Sector: Evidence from India." Unpublished manuscript, Oxford University.
- Lemos, R., and Scur, D. (2017). "Developing Management: An Expanded Evaluation Tool for Developing Countries." Unpublished manuscript, Oxford University.
- National Center for Education Statistics (2015). *The Nation's Report Card: 2015 Mathematics in Puerto Rico*.
- Puerto Rico Department of Education (2017). *Strategic Plan 2017-2020*.
- Puerto Rico Department of Education (2018). *Consolidated State Plan: The Elementary and Secondary Education Act of 1965, as Amended by the Every Student Succeeds Act*.
- Rockoff, J. E., Staiger, D.O., Kane, T.J., and Taylor, E.S. (2012), "Information and Employee Evaluation: Evidence from a Randomized Intervention in Public Schools.", *The American Economic Review*, 102(7), 3184-3213.
- Schleicher, A. (2015). *Schools for 21st-Century Learners: Strong Leaders, Confident Teachers, Innovative Approaches*. International Summit on the Teaching Profession, OECD Publishing.
- Strang, D. (1987). "The Administrative Transformation of American Education: School District Consolidation, 1938-1980.", *Administrative Science Quarterly*, 32(3), 352-366.
- Taylor, E. S. and Tyler, J.H. (2012), "The Effect of Evaluation on Teacher Performance.", *American Economic Review*, 102(7), 3628-3651.
- Woessmann, L. and West, M. (2006). "Class-size effects in school systems around the world: Evidence from between-grade variations in TIMSS.", *European Economic Review*, 50, 695-736.