

School-Based Management and Learning Outcomes

Experimental Evidence from Colima, Mexico

Vicente Garcia-Moreno

Paul Gertler

Harry Anthony Patrinos



WORLD BANK GROUP

Education Global Practice

June 2019

Abstract

A school-based management program was implemented in Mexico in 2001 and continued until 2014. This national program, *Programa Escuelas de Calidad*, was considered a key intervention to improve learning outcomes. In 2006, the national program was evaluated in the Mexican state of Colima, being the first experimental evaluation of the national program. All schools were invited to participate in the program; a random selection was performed to select the treatment and control groups among all the applicants. An intent-to-treat approach did not detect any impact on learning outcomes; a formal school-based management

intervention plus a monetary grant was not enough to improve learning outcomes. First, the schools in the evaluation sample, control and treatment, were schools with high learning outcomes. Second, these schools had experienced some years of regular school-based management practices before the evaluation. A difference-in-difference design is used to identify heterogeneous effects of the program on learning outcomes. The difference-in-difference approach shows that the intensity of treatment increased test scores during the first year of the intervention.

This paper is a product of the Education Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://www.worldbank.org/prwp>. The authors may be contacted at athpatrinos@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

School-Based Management and Learning Outcomes: Experimental Evidence from Colima, Mexico¹

Vicente Garcia-Moreno
XABER

Paul Gertler
University of California at Berkeley

Harry Anthony Patrinos
World Bank

JEL Codes: I20, I21, I28

Keywords: School-based management, impact evaluation, Mexico

¹ The authors gratefully acknowledge funding from the World Bank's Research Support Budget and the comments of participants at seminars. Thanks to Oscar Hernández and staff at the Secretariat of Education of Colima, Mexico, for all the support provided during the evaluation. Views expressed are those of the authors and should not be attributed to the World Bank Group or to their respective organizations. Comments from Marta Rubio are gratefully acknowledged. We also appreciate the implementation support from Stefan Metzger.

INTRODUCTION

At first, school policies focused on the supply side of education—more schools, more classrooms, more textbooks, more of everything that could be physically bought—as if somehow children were going to learn, and teachers to teach, just by having the tools without the right incentives. Over time, centralized management policies created inefficiencies at the school level (King and Cordeiro-Guerra 2005). In this line of thought, Hanushek and Woessmann (2007) identify three main institutional features that affect learning outcomes – choice and competition, school autonomy, and school accountability. At the local level, school actors may be able to observe, and presumably to encourage, the efforts to improve learning while the central authority only observes noisy signals on whole-school performance which reflected broad solutions with low impact on learning.

School-based management (SBM) is defined as control over responsibility of school administration by the school actors, which in most cases are principals, teachers, parents, sometimes students, and other school community members (Burns, Filmer and Patrinos 2011; Hanushek and Woessmann 2013; Bloom, Lemos, Sadun and Van Reenen 2015). In other words, the decisions are made by local agents. SBM reforms generally allow schools to allocate their budget as they prefer, to hire and fire staff, to develop their own curriculum, to procure educational material, and to monitor and evaluate teachers' performance and students' learning outcomes. The potential benefits of SBM are: increase in inputs and resources from parents, more effective use of resources, higher quality of education, more open school environment, an increase in participation of local stakeholders, and improved student performance.

Advocates of SBM assert that this idea may improve educational outcomes for several reasons: For example, SBM intervention can: (1) increase administrative efficiency and tighten professional control; (2) address local needs through enhanced community control; and (3) balance decision-making between parents and teachers, who are the main stakeholders in any school (Burns, Filmer and Patrinos 2011). Nevertheless, SBM can also lead to a change in the rules of interaction among school actors. The framework of SBM, the responsibility and decision-making, vary upon the many different forms of the different programs.

Several studies in different countries have been made to draw some conclusions about the importance of SBM. The main conclusions explored in the literature can be assigned to three categories: the first category groups results involving the dynamics of teacher and parental behavior. Regarding this point, King and Ozler (1998) and Gunnarsson et al. (2004) conclude that SBM help to modify school dynamics either due to an increase in parental involvement or changes in teacher behavior. The second category synthesizes results on the impact of SBM on repetition and dropout rates, studies such as Gertler et al. (2006) and Bando (2010) show that SBM lowers both repetition and dropout rates due to incentives alignment. Finally, some of the literature explores the relation between SBM and students' performance on standardized tests. King and Ozler (1998), Sawada and Ragatz (2005), and Lopez-Calva and Espinosa (2006) show mixed evidence about the relation of SBM and standardized test scores.

Due to the mixed evidence presented in the last-mentioned category, in recent years, most of the studies have focused on the impact of SBM implementation on student performance at the micro

level. For example, Abdulkadiroglu et al. (2011) show that charter schools in Boston with binding lotteries generate larger gains in students' standardized test scores. Khattri et al. (2013) show that the introduction of SBM has a positive impact on the school-level test scores in 23 school districts in the Philippines, and Hanushek et al. (2013) show with over one million student observations for 42 countries using the PISA standardized test that SBM reforms in well-developed institutions may be conducive to student achievement, while in low-performing institutions the opposite effect will occur.

Improving school performance, especially in poor communities, remains a challenge facing most countries. One policy being examined by many developing countries is SBM; Hong Kong SAR, China; Indonesia; Kenya; Kyrgyz Republic; Nepal; Paraguay; and Mexico are among the main countries implementing the program. To date, the empirical evidence on the effects of SBM interventions on student academic performance in developing countries is quite limited. Several studies rely on cross-sectional variation, ex-post propensity score matching and exclusion restrictions – either using functional forms or weak instrumental variables – thus leaving their ability to establish causality open to question (see, for example, the works of Jimenez and Sawada 1999, 2003 on El Salvador's EDUCO; DiGropello and Marshall 2005 on the effects of the Honduras PROHECO program; King and Ozler 1998, King et al. 1999 and Parker 2005 on Nicaragua's Autonomous School program; or López-Calva and Espinosa 2006 on the impacts of AGE as well as the other Compensatory Program supports on test scores). Notable quasi-experimental exceptions include Shapiro and Skoufias (2005) and Murnane et al. (2006) who use difference-in-differences models to estimate the impact of the PEC national intervention on drop out, repetition and failure rates. Gertler et al. (2008a) combine quantitative (difference-in-differences) and qualitative methods to examine the effects of an initiative that involves parents directly in the management of school grants transferred to the parent associations in schools in highly disadvantaged rural communities.

Duflo, Dupas and Kremer (2015) is one of the few studies that exploits a controlled randomized experiment to examine the impacts of SBM. The authors evaluate the effects of monetary empowerment of local school management committees to monitor and train teachers combined with contract teacher hiring in primary schools in Kenya. They show that combining class size reduction with improved incentives – by either hiring contract teachers (as opposed to civil servants) or increasing parental oversight – leads to significantly larger increases in test scores.

This analysis explores the impact of a school-based management program in Mexico, where the government attempts to empower the school community, and to promote school autonomy and accountability. This experiment is based on the federal education SBM program, which began in 2001. The final objective of this initiative is to encourage parents, teachers, and directors to design and carryout school strategies and transformation plans to respond to the needs of schools and their students. For this purpose, the SBM program awards selected schools with annual monetary grants and provides technical assistance to further improve the educational plans. All submitted school plans are evaluated and ranked by a technical committee. Only the higher quality proposals are eligible for SBM benefits.

This research evaluates these issues through the empirical study of the Quality School Program (*Programa Escuelas de Calidad* or PEC) in the Mexican state of Colima using a control

randomized evaluation design. PEC is a major school-based management program launched in 2001 by the Mexican federal government to improve the quality of public schools in urban areas. In Colima, the state government enjoyed certain flexibility in the implementation of the intervention.

Our study contributes to the previous literature in two ways. First, the random allocation of PEC benefits to schools permits the unbiased identification of the effect of this SBM initiative on education outcomes. We can isolate the impact of PEC from that of other educational reform initiatives that may be simultaneously operating in the school. Second, the short- and medium-term impacts of the program can be analyzed due to 12 years of information on the program in Colima and eight years of student test scores.

We take advantage of these factors to assess the short- and medium-term impacts of PEC over eight school years – from school year 2006-07 to school year 2012-13 – on school performance and student learning. More specifically, we follow a sample of 98 experimental primary schools where we have randomized the allocation of benefits. The present study follows two alternative strategies to identify the effects of PEC on test scores. The first compares treatment and control groups exogenously determined by a randomization. The second strategy exploits the difference-in-difference approach to investigate the intensity of the treatment over time on test scores.

The comparison of means between Treatment and Control schools did not detect any impact of the intervention on learning outcomes. However, the difference-in-difference shows that the intensity of treatment increased test scores during the time of the intervention and after. The intuition behind these results is that the participating schools, control and treatment, were SBM schools before, during and after the intervention; a formal SBM intervention with a cash grant was not enough to improve learning outcomes in schools with regular SBM practices. However, the program has a heterogeneous impact on test scores conditional on the level of intensity-in terms of years participating in the program-of the treatment.

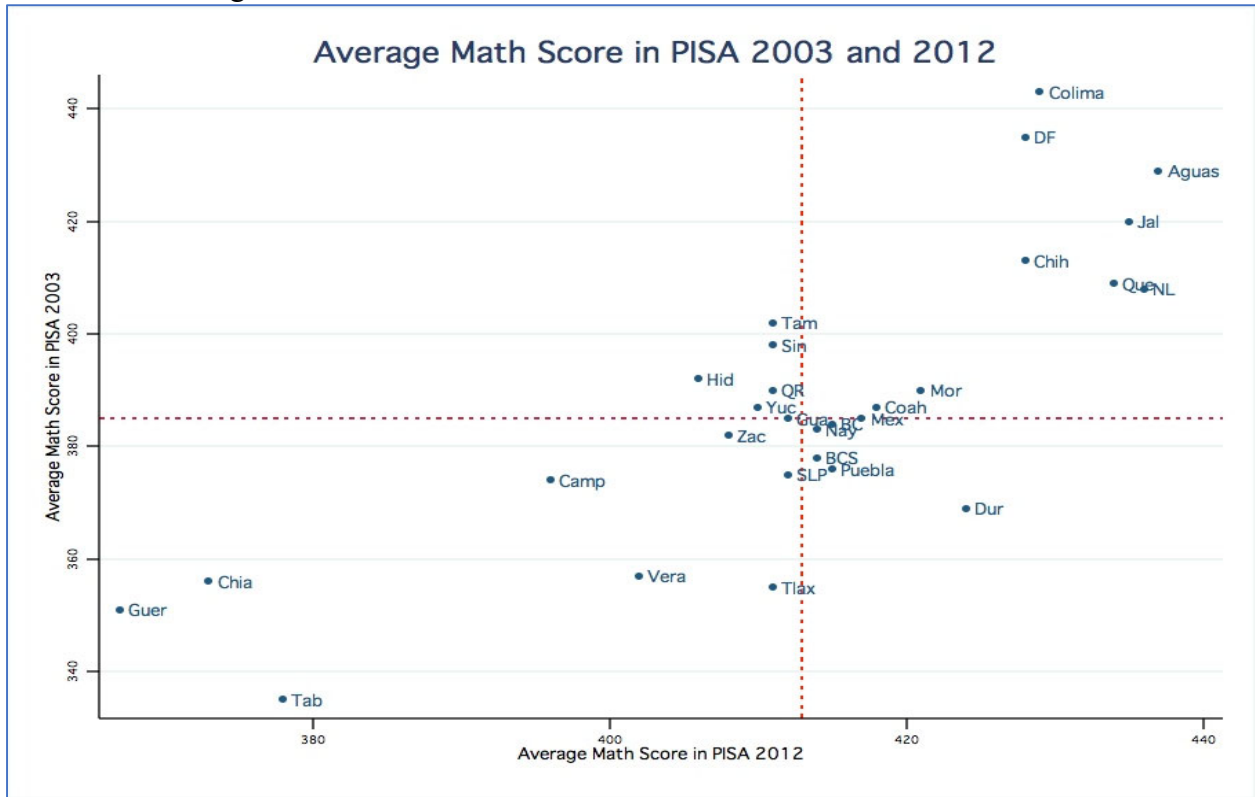
BACKGROUND AND RECENT TRENDS

A small, well-connected state, with just two major urban centers, Colima has undoubtedly benefited from these characteristics in its efforts to improve educational outcomes. Within Mexico, the state of Colima has been a champion in building an efficient educational program to their specific educational development (De Hoyos et al. 2017). In 2006, a World Bank report (2006) analyzed the relationship between test scores and institutional reforms in Mexico at the state level. The report concluded that states with the best academic results are characterized by: state-wide examinations (including own assessment systems), decentralized decision-making to the local or school level, and campaigns to persuade teachers, teacher representatives (unions) and civil society of the necessity and effectiveness of their reforms. All these elements have been in operation in Colima since the late 1990s.

Colima ranks first among all Mexican states in the international achievement test, PISA, which is organized by the OECD. In many ways, this achievement is the culmination of a combination of both traditional and more innovative efforts to improve educational outcomes since the late 1990s.

Figure 1 shows the relative performance of Colima with respect to Mexican states in PISA 2003 and 2012.

Figure 1: PISA Math Score across Mexican States, 2003 and 2012



Source: Authors' calculation using PISA 2003 and PISA 2012.

Anecdotal evidence suggests that the continuity of high rank officials in office and their strong influence in the federal government have also contributed. Moreover, Colima possesses well-functioning information and planning schemes that benefit from unique files for each student and modern communications and computation systems. This has allowed for important decentralization to the municipal level with 10 units providing administrative functions in personnel management, buildings, equipment and materials; and four regional centers for teacher training – *Centros de Maestros Regionales* or CMR – which also provide technical assistance.

Supporting this local initiative is a rather well-developed system of assessment – the School Quality Competition (*Concurso Escuelas de Calidad* or CEC) – which annually tests students at five different grade levels in every school in the state. Results from these assessments are available to teachers and principals for the subsequent school year for each student, grade, school and municipality and provide for accountability and the targeting of incentives tied to performance. They also provide feedback for quality enhancement activities undertaken by the school. Schools can choose whether to disseminate results to parents and the wider community. Often, only the best performing schools choose to do so.

Colima education authorities also attribute their performance to two other innovations: competitive testing for half of all new teachers and an agreement with the teacher union not to rotate teachers during the school year. In other states, such rotations start a chain reaction that can result in some classes having 3-4 different teachers during the academic year.

School-Based Management Program in Mexico: “*Programa Escuelas de Calidad*” (PEC)

In 2001, the Mexican federal government introduced PEC to empower the school community and promote school autonomy and accountability. The main objective of this initiative was to encourage parents, teachers and directors to design and carry out School Strategic Transformation Plans that responded to the needs of the school and its students. For this purpose, PEC yearly awards a monetary grant and technical assistance to implement these improvement plans in selected schools. While it builds on previous small-scale experiences that decentralized limited decision-making powers to the school, PEC was the most extensive SBM program implemented in Mexico before 2013 (see Schmelkes 2001 for more details on other SBM initiatives in Mexico prior to PEC). PEC also became one of the flagship programs of the Fox administration (2000 - 2006) and it expanded rapidly, from 2,239 schools included in the program in 2001 to around 55,432 schools at 2014. PEC was canceled in 2014, after 14 years of implementation; a new SBM program was initiated called “*Escuelas al Centro*” (Schools at the Center).

The national PEC intervention has been subject to evaluation. Murnane et al. (2006) use nationally representative data and difference-in-difference estimation techniques to compare PEC schools – defined as schools that joined the program in its second year of operation (2002) – to non-PEC schools. They validate the equality of pre-intervention trends between both sets of schools. However, the authors reject the equality of trends in the pre-period when schools that received benefits during the first year of operation of the intervention are included in the treatment group. These schools were improving outcomes more rapidly in the pre-PEC years than comparison schools which suggests that any PEC evaluation is likely to suffer from self-selection bias, at least on its initial stages – that is, better performing schools had a higher probability to apply for and receive benefits in the first years of program operation. The authors find that the program reduced dropout rates by 0.27 percentage point but find no impact on failure rates.

Skoufias and Shapiro (2006) build on the standard difference-in-difference methodology in Murnane et al. (2006) to evaluate PEC on school averaged repetition, failure and dropout rates over the first three years of PEC benefits. Their approach combines difference-in-difference techniques and propensity score matching. In this way, each PEC school is matched to a set of comparison schools with similar observed characteristics and located in similar communities, rather than to all comparison schools. However, because they only have data on two periods, the authors are forced to assume that pre-intervention trends between treatment and comparison schools were the same. Their findings suggest that participation in PEC decreases dropout rates by 0.24 points, failure rates by 0.24 points and repetition rates by 0.31 points.

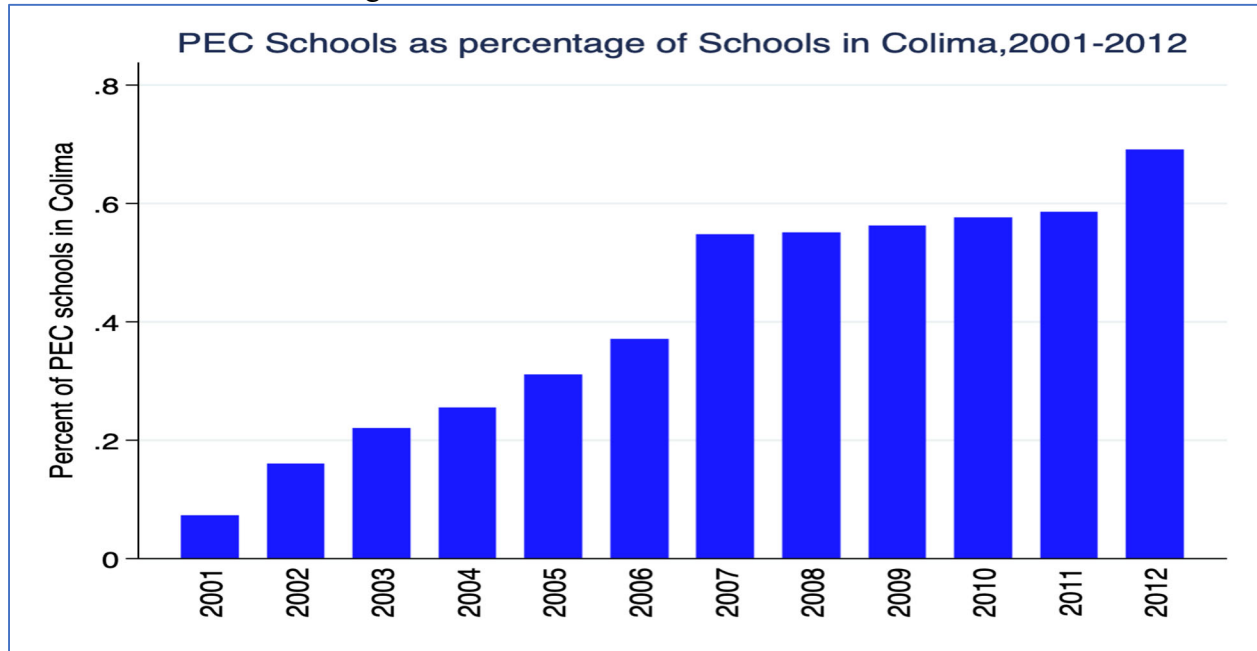
Loera (2005) and Loera and Cazares (2005) use achievement and coverage data over school years 2001-02 to 2003-04 and report a statistically significant positive correlation of 0.90 point between student test scores and PEC benefits. Bando (2010) exploited the exposure to the program across cohorts and schools over time; Bando finds that the program impacted decreasing failure and drop-

out rates but there is no evidence on test scores. These results are larger in schools with more time with the program. Recently, significant evidence on the significant effects of the PEC program in learning outcomes (Santibañez, Abreu and O'Donoghue 2014). Cabrera Hernández and Perez Campuzano (2018) evaluated PEC from 2008 to 2013 using a DD strategy. They find both that PEC has a positive impact on learning and more years in the program have more of an impact, particularly for relatively better-off schools. Santibañez et al. (2014) show that, under certain conditions (mainly much bigger size grants), PEC can also improve learning outcomes as measured by the national standardized test, ENLACE (Santibañez et al. 2014).

School-Based Management Programs in Colima

In 1998, Colima created a mandatory program of school-based management named “*Proyecto de Gestión Escolar*” for all schools (except private) in the state and it was independent from the start of PEC in 2001. The PGE provided tools for diagnostics and activity planning and involved all school agents (teachers, principals, parents and students) in the design and implementation of yearly school plans. The PGE aimed at improving school functioning and performance by devolving increased decision-making power and voice to the school community, based on the premise that schools can better identify their needs and determine the most effective way to use resources to address them. The PGE was designed to support school autonomy to improve learning outcomes in the state. Before 2001, almost all schools in the state had benefited for several years from an initiative to promote a formal school autonomy, the School Management Project (*Proyecto Gestión Escolar* or PGE). In 2001, Colima introduced PEC with grants to support their PGE program incentivizing the empowerment of the school community and promoting school autonomy and accountability. Figure 2 shows the number of schools that at least had one year of PEC during 2001 to 2012 out of all public schools. PEC covered a great proportion of all public schools.

Figure 2: PEC Schools in Colima, 2001-2012



Source: Administrative data from PEC.

These grants were jointly financed by the Federal Secretariat of Public Education and by the state governments. On average, the federal investment in the program leverages up to half as much as the state contributions. In addition, PEC raises contributions from local school communities, where schools mobilize parents, municipal governments and private organizations to support their school improvement plans. Although local (school and municipality) cost sharing of the school improvement plan is encouraged to match the state contributions, it is not required. Usually, national regulations govern the implementation process and the federal government is responsible for monitoring the implementation of the intervention and suggesting corrective adjustments to the state authorities.

The program is open to all public basic education schools although priority is given to disadvantaged schools located in poor urban areas, as these schools have the most need and could benefit more from the program. To participate in the program, schools voluntarily submit their improvement plan – the School Strategic Transformation Plan or in the case of Colima, the PGE – along with a needs assessment and a list of actions to address these needs. Proposals are submitted early in the school year, typically by the end of October. In the state of Colima, the preparation of a PGE is mandatory for all schools. Moreover, all schools receive information on the PEC call for proposals along with support from the CMR (teacher training and development centers) to write their school improvement plans. However, not all schools choose to submit their school plan and apply for PEC benefits.

The Intervention

In 2006, local educational authorities in Colima started receiving applications for PEC-2006, and for the first time, an impact evaluation of the program was planned. The randomized allocation of PEC benefits to schools permits the unbiased identification of the effect of this SBM initiative on

education outcomes. A Randomized Control Trial (RCT) was used to answer the question: does a formal School-Based Management program increase school performance (as measured by test scores)?

Using this approach allows us to isolate the impact of PEC from that of other educational reform initiatives that may be simultaneously operating in the school. In Colima, these include the following:

- (i) policies that enhance the demand for schooling like the conditional cash transfer program, *Oportunidades*
- (ii) policies that strengthen education supply such as the Compensatory Program and its SBM component, the AGE
- (iii) other policies that develop school autonomy, such as the School Management Project (*Proyecto Gestión Escolar* or PGE)
- (iv) information for accountability systems, like the School Quality Competition (*Concurso Escuelas de Calidad* or CEC).

Figure 3: Intervention Timeline

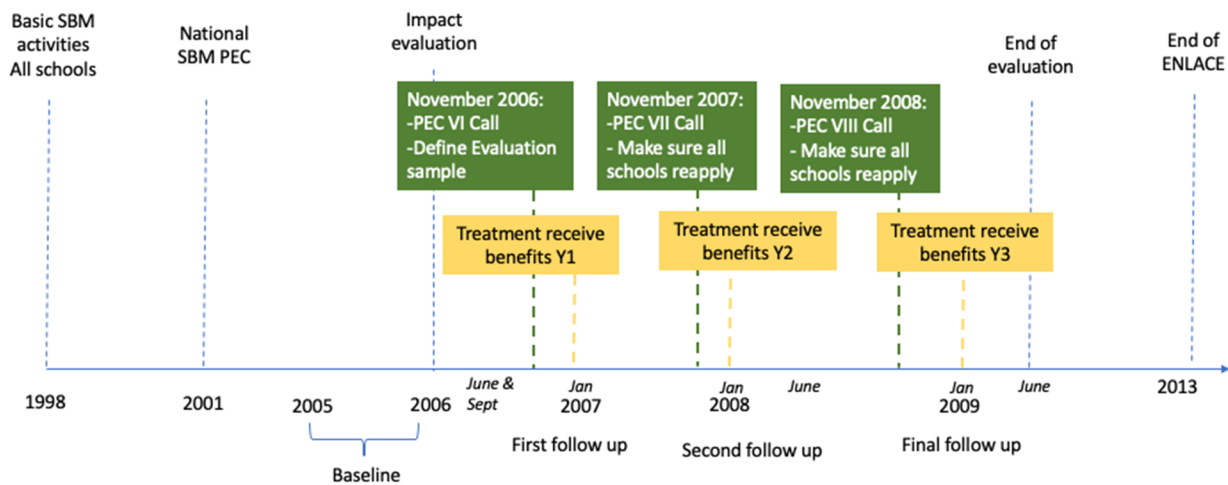


Figure 3 presents the intervention timeline. Of the 435 primary schools in Colima in academic year 2006-2007, 39 were private and 88 were one- and two-teacher schools. In November of 2006, all submitted school plans were evaluated and ranked by a technical committee. Higher quality proposals were eligible for PEC benefits. The list of winning schools was made public on the internet in early December and benefits usually started being disbursed in early January.

In Colima, the evaluation committee was composed by CMR staff and assessors, who were themselves teachers. It was required that the assessors review proposals from schools outside the area of influence of their CMR. There were specific criteria to mark and evaluate proposals established by the state education authority. These criteria included:

- (i) whether the proposal identifies a specific problematic in the school,
- (ii) the sources of information used to identify such problematic – for example, standardized examination results, consultation, census data on performance, etc.; and
- (iii) how accurate and appropriate the work plan laid out is to address the problematic identified.

For each of these criteria, a score was assigned according to an equivalent quality scale. The total score of the proposal was the weighted average of the scores obtained for each criterion.

In principle, the schools enjoyed relative freedom in determining whether the school transformation plans should fund school maintenance and repairs, construction of new physical infrastructure, acquisition of educational materials or professional development for school staff. In practice, program norms place a cap on the percentage of resources that can be devoted to these different activities. In many states, schools were eligible for benefits over five consecutive school years. Initially, during the first four years of participation in the program, at least 80 percent of the resources received must be spent on school infrastructure and maintenance, and in the acquisition of pedagogical materials. As of school year 2006-07, the percentage to be spent on infrastructure for schools newly incorporated into the program declined to 70 percent. The remaining 20 percent of the funds can be spent on training for teachers, parents, or the school director. This amount is increased up to 50 percent during the last year of participation in PEC.

In Colima, however, schools receive PEC benefits for one year. Beneficiary schools must re-apply and submit a new (original) winning proposal to receive benefits for two (or more) consecutive years. Nonetheless, schools in the program in a given year are given priority over non-beneficiary schools if they re-apply. Extra criteria, such as the originality of the new proposal and the appropriate use of the funds granted during the previous school year, apply to these schools. For schools that receive benefits over various years, similar rules to those existing in other states govern the allocation of the funds. During the first three years of benefits, schools are required to spend 70 percent of the grant in equipment and infrastructure, and the remainder of the grant in pedagogical development. During years four and five, these percentages are half and half. As in all other states, no school in Colima can receive benefits for more than five years (consecutive or not).

Each school gets at least 50,000 Mexican pesos (or about \$2,500) per year, which amounts to the federal contribution. The average national grant in 2001 amounted to 220,411 pesos (or about

\$11,000) although it had diminished to 55,691 pesos (\$2,800) by 2006. In Colima, the municipality and the school are encouraged to raise another 20,000 and 30,000 pesos (\$1,000 and \$1,500). The state government matches each peso raised by the local and school communities to contribute to the development of the improvement plan. Hence, the maximum benefit a school can receive in a year amounts to 150,000 pesos (\$7,600).

The monetary benefits are complemented with permanent support from the state government in the implementation of the school improvement plan and in the management of the funds. In Colima, this support is channeled through the four regional teacher centers, the CMR. These centers are also in charge of providing adequate professional development through workshops, courses and seminars for teachers and principals. Similarly, the CMR run workshops and distribute leaflets and information materials to encourage parental participation in school matters and activities. Formal rules mandating the participation of parents in the school improvement plan are a requisite for school participation in the program.

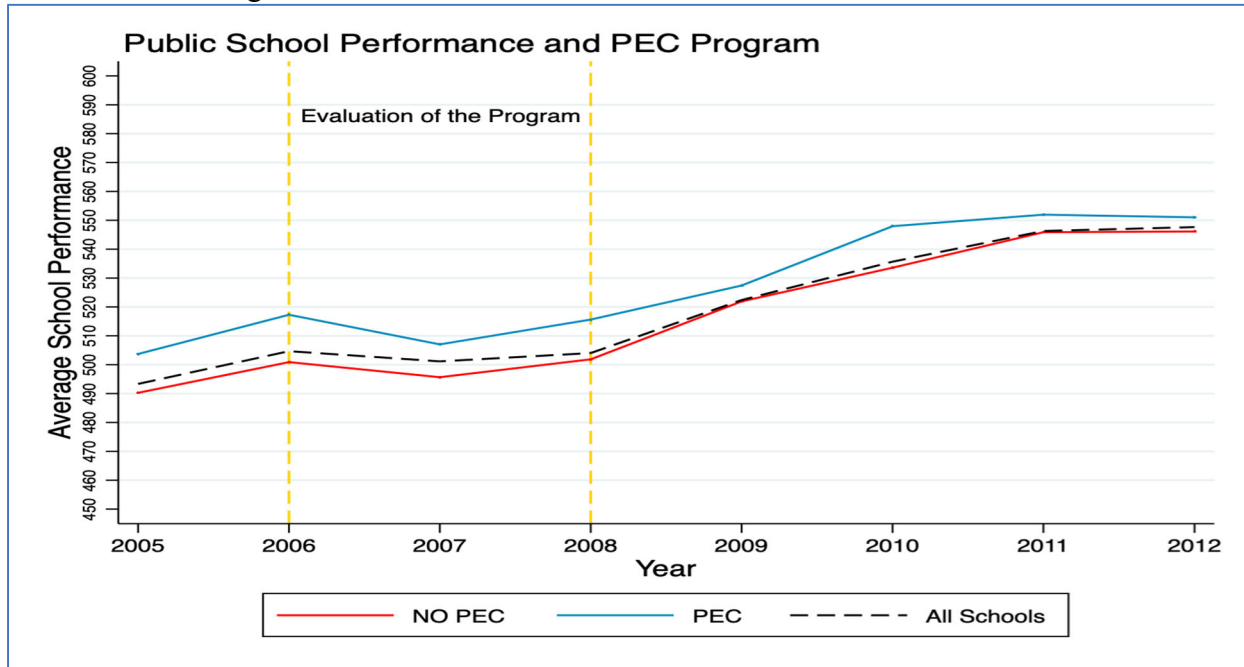
Data and Trends

This analysis uses data collected for the RCT and five additional sources of data: the National Evaluation of Academic Achievement in School Centers (ENLACE), the administrative PEC data, local administrative achievement data and the administrative School Census data (SCD- 911). The data set utilized in this analysis includes Math and Spanish test scores from ENLACE, PEC program information, and school inputs from SCD-911 in the school years 2001-2002 through 2012-2013. The Ministry of Education (ME) produces these sources of data. The last data set is the index of disadvantage (*indice de marginalidad*), which is a weighted average of literacy, access to basic public utilities, household infrastructure and average wages. All localities are ranked with ratings from very high, high, medium, low, and very low disadvantage. The National Population Council (*Consejo Nacional de Población*, CONAPO) constructs this index.

Since the school year 2005-2006, the Ministry of Education has administered the National Evaluation of Academic Achievement in School Centers (ENLACE) nationwide. ENLACE was administered annually in April to early June. ENLACE has limitations for between grade comparison; for this reason, the analysis is based on within grade comparison in repeated cross sections.

Figure 4 depicts average school performance, defined as the average Math and Spanish scores from 3rd grade to 6th grade, from 2005 to 2012, schools participating in PEC are above the average performance of public general schools in Colima; non-PEC schools have a relatively lower performance than the average in the first years of the program.

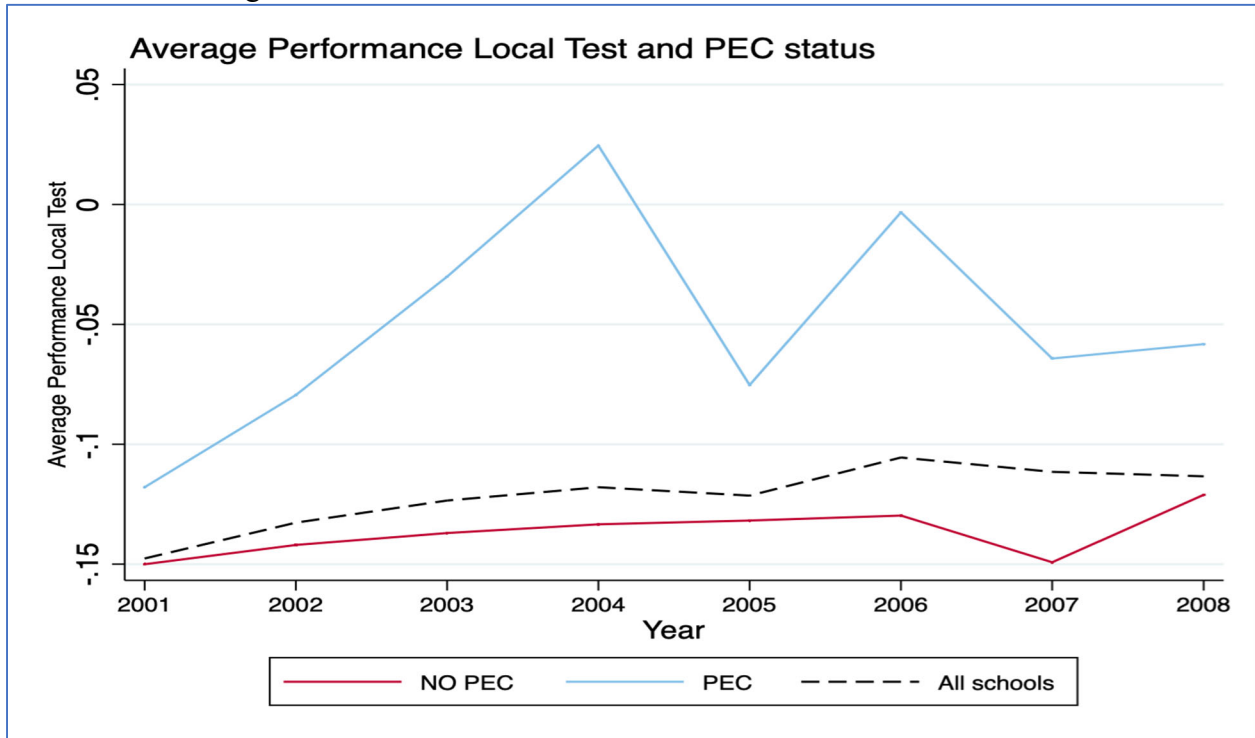
Figure 4: ENLACE School Performance in Colima, 2005-2012



Source: Author's calculation using ENLACE and the administrative data from PEC.

Since 1997, Colima has had an evaluation system which publicly rewards the schools with the highest achievement rankings. Contest of Quality Schools “*Concurso de Escuelas de Calidad*” (CEC) is an annual test for students at five different grade levels in every school in the state. Figure 5 shows the Z-scores between schools with the PEC and without PEC conditional that the composition of the schools changed every year. However, PEC schools had always the best results in ENLACE. The dash line represents the average performance of all schools in Colima. This result confirms the same trend as in the national examination.

Figure 5: Local Test School Performance in Colima, 2005-2012



Source: Authors' calculation using the local achievement test and administrative data from PEC.

At baseline, 98 schools applied for the program; 49 schools were selected to get the treatment and 49 schools were selected as a control. In the second year of the intervention, only 44 schools in the treatment group applied; while 39 control schools applied. Moreover, 40 treatment schools got the PEC for the second year (80 percent) and control schools got the PEC grant. In the following school year, only 43 treatment schools out of 49 applied to get the grant. This number was 36 schools in the control group. However, 31 treatment schools out of 43 received the PEC and 2 control schools out of 50 schools in initial control group.

Table 1 depicts mean average school performance between control and treatment schools in Colima from 2005 to 2012. Before the program, both groups have no statistical differences in test scores. These differences between control and treatment are not significant at the first year of the program. In 2007 and 2008, there is a significant test score gap between groups. After 2009, control schools have better performance than treatment. During the program, these differences were not significant.

Table 1: Average School Achievement, evaluation sample

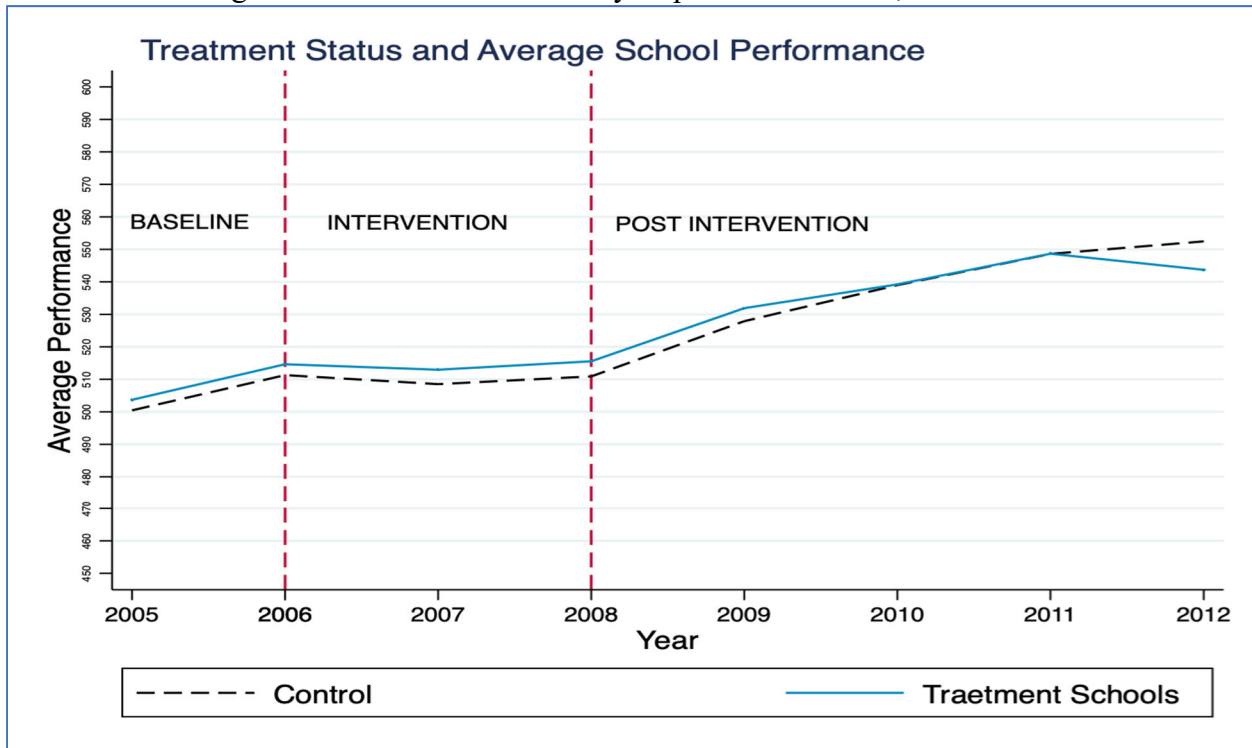
	Control	Treatment	diff	S.E.
2005	500	503	-2.45	1.60
2006	511	514	-2.56	1.53
2007	509	512	-3.58	1.69
2008	511	515	-4.14	1.75
2009	528	531	-2.67	1.78
2010	539	539	-0.10	1.78
2011	549	549	-0.17	1.80
2012	552	545	7.31	1.79

Source: ENLACE and Evaluation roster

In a graphical analysis, the figures below investigate whether there has been any impact of the program school averaged measures of achievement. We first compare the mean of the performance indicators analyzed throughout this research. Figures 6 and 7 show there are no significant differences between these two groups.

Figure 6 illustrates the school average performance in all schools in Colima from 2005 to 2012. While private schools are the best performing schools in Colima, the selected schools for the experimental analysis perform higher than the average schools in Colima. In this graph, treatment and control schools were equal at the baseline. Moreover, this gap never was significant during the intervention or after the program.

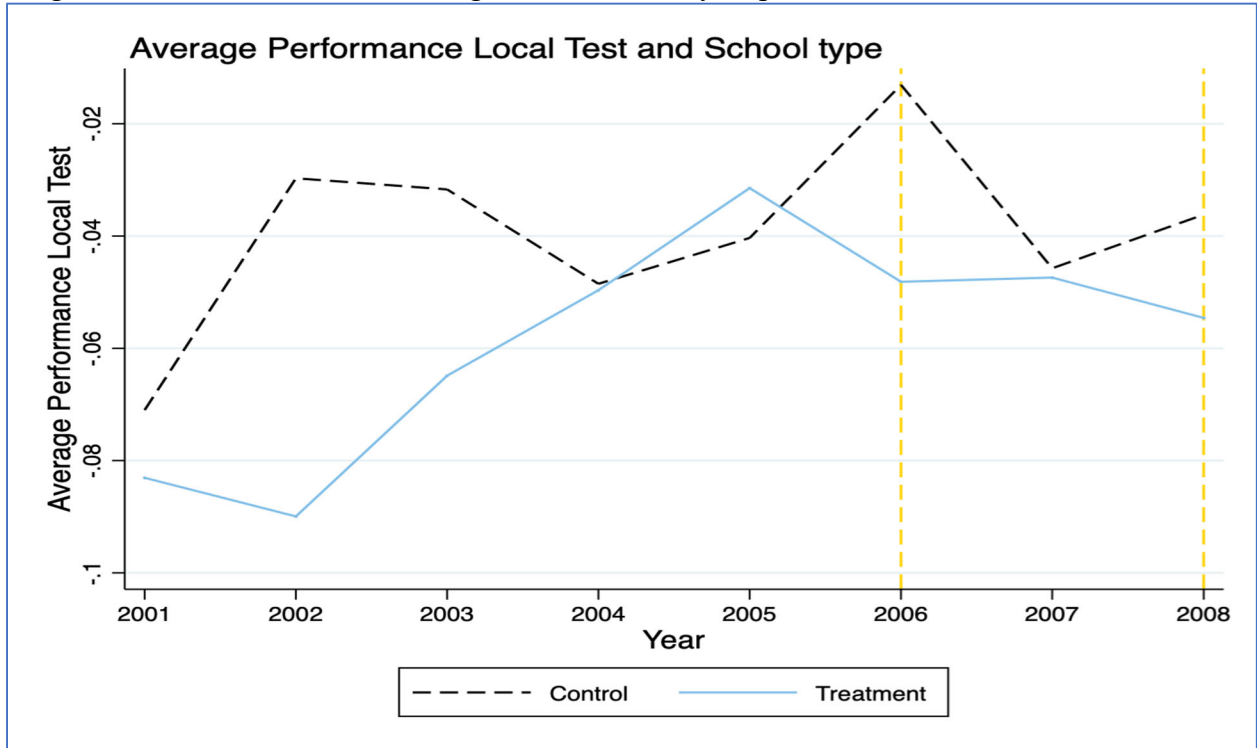
Figure 6: School Performance by Experimental Status, 2005-2012



Source: Authors' calculation using the ENLACE and administrative data from PEC.

Figure 7 illustrates the average school performance using the local standardized test in the experiment in Colima from 2001 to 2008. In this graph, treatment and control schools were equal at the baseline. Moreover, this gap never was significant during the intervention or after the program. The Annex presents more information on sample balance.

Figure 7: School Performance using the Local Test by Experimental Status, 2005-2012



Source: Authors' calculation using the local achievement test and administrative data from PEC.

Pitfalls of the Implementation

While the design of the program was done carefully, the PEC implementation had several pitfalls. In summary, three factors have compromised the reliability of this randomized trial. First, the fact that each school needed to apply each year distorts the incentives of participation to both groups. Since the application requires effort, coordination and time across school actors; treatment schools were asked to do this process again in additional procurement of the previous grant. The rules of the program include some stringent requirements that make treatment schools go through a bureaucratic process each year to get the additional funds. For the control group, the application represented a waste of their effort and time; to apply again was less attractive to these schools.

Second, some control schools had PEC during the second and third years of the evaluation, reflecting the lack of planning when control schools get the program. Finally, financial resources, from the federal authorities, were not disbursed for all the years of the evaluation even when the Ministry of Education was committed at the beginning, which affected the number of treatment schools that received the program.

After the first year of implementation, some control schools had PEC and some treatment schools did not receive PEC. These cross-overs challenge the exogeneity of the assignment when the internal validity of the experiment was compromised. In this case, causal inferences cannot be claimed for the second and third years of the evaluation. For the first year, the results are valid.

Table 2 shows the number of schools that received between 0 and 5 years of PEC by school year. In the baseline, 50 schools experience PEC at least one year (including treatment and control). After a year of the intervention, only 31 schools did not experience PEC or 53 percent of the control group never had any experience of PEC. The accumulative experience of PEC increased with the intervention among the treatment. The second year of the experiment, the years of PEC increased again. The final year of the intervention, 10 schools had five years of PEC experience, 11 schools had four years of PEC, 23 schools had three years, 24 schools had 2 years, 14 school had only one year of the program and finally 16 control schools never got PEC.

Table 2: Intensity of PEC Implementation

Intensity	2005	2006	2007	2008
0 PEC	48	31	16	16
1 PEC year	26	30	17	14
2 PEC years	24	25	32	24
3 PEC years	0	12	21	23
4 PEC years	0	0	12	11
5 PEC years	0	0	0	10
Total	98	98	98	98

Source: ENLACE and Evaluation roster

The fact that the evaluation was contaminated by the factors mentioned above makes the evaluation invalid after the first year. We use the number of years in PEC the last year of the evaluation and schools are identified since the baseline.

Table 3: PEC Schools by Years in PEC

Intensity	Baseline	2008	2012
0 PEC	48	16	2
1 PEC year	26	14	22
2 PEC years	24	24	13
3 PEC years	0	23	22
4 PEC years	0	11	17
5 PEC years	0	10	22
Total	98	98	98

Source: ENLACE and Evaluation roster

Table 4 illustrates the intensity of treatment analysis. Schools with zero years with PEC had an increase of 24 points (0.24 standard deviation). Schools with only one year in the program had an increase of 6 points; whereas, schools with three years of PEC increased by 9 points. Those schools that have 3 and 4 years in PEC had an increase of 13 and 21 points.

Table 4: Average School Achievement and Years in PEC, Evaluation Sample

Intensity	2005	2006	2007	2008
0 PEC	563	560	557	560
1 PEC year	491	501	496	504
2 PEC years	515	527	525	529
3 PEC years	499	508	508	504
4 PEC years	507	518	515	524
5 PEC years	500	513	510	511

Source: ENLACE and Evaluation roster

EMPIRICAL STRATEGY

We follow two alternative methodologies to identify the effects of PEC on student-level test scores: an experimental design exploiting the random assignment of schools to the treatment and control groups and a differences-in-difference (DD) approach exploiting post-treatment differentiated performance on test scores between PEC and non-PEC schools. In both methodologies, schools are the unit of intervention and the unit of analysis is the student.

Experimental Design

Following the experimental design of the impact evaluation of PEC-Colima, the randomization of the program establishes a causal linkage between PEC (treatment) and school performance (outcome). The counterfactual is a group of comparison, which simulates what would happen to the target group in the absence of the intervention. A randomized experimental evaluation reduces the possibility that observed changes in outcomes in the intervention group are due to exogenous factors. Therefore, random selections of schools receive treatment and the remaining schools serve as controls. This process ensures a robust counterfactual measure of the causal effect of the intervention. The intervention consisted in two steps; the first step was the application process to get the PEC program. The second stage was the randomization among those schools that applied to the program. The evaluation selected 50 schools as treatment and 50 as control. The first intervention was winning the grant. This included two schools with two teachers (one control and one treatment), which were taken out from the analysis. The second part of the intervention was the implementation of the school plan given the grant.

Using empirical design from a randomized experiment framework helps us to create homogeneous groups in the observed and unobserved characteristics at the initial stage of the analysis. Consequently, this reduces bias and allows the greatest reliability and validity of statistical estimates of the treatment effect. To test for theoretical prediction, this paper first estimates the differences between the control and treatment groups. This follows the functional form:

$$Score_{st} = \beta_1 + \beta_2(treatment_{st}) + \beta_3(SchoolInputs_{st}) + \beta_4(SES_{st}) + \epsilon_{st}$$

where score is the test score for school s in year t . The variable *Treatment* is a dummy variable in which 1 indicates that the group received the second component. The coefficient, β_2 , is of interest.

This coefficient assesses the average treatment effect on the dependent variable. For interpretation purposes, we standardized the dependent variable's coefficients to have a mean of 500 and standard deviation of 100.

School Input is the variable that includes school inputs by schools s in year t such as student-teacher ratio, percent of teachers with university or more, and percent of teachers with an incentive program (*Carrera Magisterial*). The effects of these information assets are given by coefficient β_3 . The variable *SES* is a vector of socio-economic characteristics of the locality in which the school is located.

Difference-in-Difference Approach

A difference-in difference (DD) approach is used to exploit the intensity of treatment on test scores between PEC and non-PEC schools. In a DD approach, it is assumed that factors (observable and unobservable) relate with time are captured through a year dummy variable. In addition, it assumes that the PEC program impacted school performance incorporating a dummy for schools that participated in PEC. DD assumes orthogonality between intensity of the treatment assignment and pre-treatment trend. Formally, $Score_{ist}$ is the test score of the i th student in school s in year t and $PEC_intensity_s$ is a categorical dummy taking the value of 0 to 7 for each year of participating in PEC. The empirical approach using OLS is:

$$Score_{ist} = \alpha_{0t} + \beta_1(PEC_intensity_s) + Year_t + \sum_{t=t^*}^T \beta_t PEC_Intensity_s * Year_t + \sum_{k=1}^K \beta_{kt} X_{st}^k + \epsilon_{ist}$$

where $Year$ is a year fixed effect, X_{st}^k are a series of school-level covariates and ϵ_{ist} is a random component. β_t represents the parameters of interest capturing the impact of the intensity of PEC each year between 2005 to 2012 on test scores. This identification strategy assumes that control and treatment schools' trends are parallel ex-ante the intervention.

RESULTS

This section describes the results of the experimental design and the DD approaches.

Experiment approach

A randomized experiment design is used as the empirical framework to analyze the impact of a formal SBM intervention with a cash grant on school average performance. Since the allocation of the treatment was properly done only the first year, these results are only for a year after the program so that one can claim causality of the results. This empirical approach uses treatment and control schools, as the program was not contaminated. The empirical approach included covariates. This should not affect either the coefficient of interest or its significance; covariates are often added to improve precision. Table 5 shows the results of the empirical estimation of the impact of PEC on school performance just for 2006. The first column shows the impact of the treatment on school average; this effect is not significant. The second column shows the results when the dependent variable is the average Spanish score; again, there was no impact. The same story emerges for the Math average score. When analyzed by grade, one observes the same results. In summary, there is

no effect of the program after one year of the program when the evaluation did not have any implementation problem.

Table 5: Regression Analysis of the impact of PEC on School Performance

	School Average	Spanish	Math	3 rd grade	4 th grade	5 th grade	6 th grade
Years PEC before	-8.6 (3.99)**	-8.6 (4.08)**	-6.7 (4.43)	-10.5 (4.49)**	-4.1 (5.54)	-7.5 (4.13)*	-12.6 (5.75)**
Treatment	-1.1 (7.17)	0.2 (7.31)	-4.4 (7.54)	-4.2 (7.41)	-2.4 (9.08)	-3.0 (7.75)	6.5 (10.02)
Treatment*2006	-2.0 (4.54)	-1.9 (4.37)	0.7 (4.44)	3.5 (7.72)	2.2 (7.95)	-1.3 (7.06)	-12.9 (8.32)
Constant	458.4 (14.75)***	457.2 (14.97)***	456.9 (15.45)***	476.2 (16.12)***	452.5 (19.16)***	455.5 (16.75)***	450.7 (17.74)***
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.0606	0.0522	0.0548	0.0618	0.0632	0.0615	0.0674
Observations	26,841	27,078	27,30	6,493	6,871	6,709	6,768
Clusters	98	98	98	98	98	98	98

School-Level clustered standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01

We use specifications as in Table 5 using all the years of the intervention and of the program; the results show that there is no impact of the program for the schools in the experiment. The results of these specifications suggest that a formal SBM intervention with cash grants did not have an impact on student learning. Since schools in the experimental evaluation sample were equal in expectation before the treatment, also suggests that both groups already had in place formal school autonomy practices and some also had experienced cash grants; then, the intervention just supported these practices but did not change them into better school performance. Finally, schools with more years in PEC had lower and significant school performance (close to a 0.09σ) at the baseline.

DD Approach – To Test for Heterogeneous Effects Such as Intensity of Treatment

To analyze the change in school average performance that can be attributed to the intervention, two sets of estimations were used. First, it estimated the effect of PEC program on school average test scores from 2005 and 2012; PEC schools are identified from administrative records. Second, it estimated the effect of the intensity of treatment on school performance; PEC schools identified as treatment in the evaluation sample. The second empirical approach assumes that PEC had different intensity among participating schools. This approach considers not only initial years in the program before the treatment but also four years after the program. In this case, the inclusion of dummies defines schools with different years of PEC. Table 6 shows the results of the DD estimation of the impact of PEC on school performance. The first column shows the impact of the PEC on school average; this effect was only significant for 2010 by 0.15σ , a year after the end of the impact evaluation. The second column shows the results when the program variable is a dummy for the treatment schools; again, there was no impact.

Table 6. Estimating the impact of PEC on School performance

	PEC Dummy	Treatment Dummy
PEC	2.81 (2.06)	-2.00 (7.33)
2006 PEC	-2.39 (10.25)	-0.25 (4.32)
2007 PEC	4.56 (9.54)	0.76 (4.67)
2008 PEC	-5.27 (10.73)	-0.63 (4.95)
2009 PEC	-11.33 (10.29)	-2.30 (5.49)
2010 PEC	15.15 (8.43)*	-3.95 (5.71)
2011 PEC	-16.39 (15.06)	-3.88 (6.37)
2012 PEC	-3.12 (8.61)	-10.60 (8.00)
Constant	445.33*** (19.12)	447.36*** (17.01)
Controls	Yes	Yes
R2	0.06	0.06
Clusters	90	92
Observations	110,11	111,618

School-Level clustered standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01

Previous insignificant results overall could hide important heterogeneous effects of the intervention. A SBM program may have a differentiated effect within the schools in the evaluation. Table 7 shows specifications using the intensity of the treatment in the experiment sample in 2006. The intensity of the treatment among very selected schools in Colima shows positive effects on

schools with 5 years in PEC; this result implies a medium-term capitalization of their SBM practices since PEC started in 2001.

Table 7: Impact of PEC by Intensity of Treatment on learning outcomes in 2006

	Average	Spanish	Math	Male	Female
PEC 1 year	9.26*	16.95**	1.97	9.5**	8.64
PEC 2 year	14.17**	21.15**	7.60	15.2**	14.05**
PEC 3 year	8.18	14.97**	1.57	4.5	12.03*
PEC 4 year	9.17	15.46*	3.52	13.3**	4.92
PEC 5 year	14.14***	22.85***	9.43***	14.9***	13.58**

Control included and year fixed effects. School-Level clustered standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; more results available upon request.

The second column shows the results when the dependent variable is the average Spanish score; the effects on Spanish scores in 2006 range from 0.14σ to 0.23σ . The table also shows heterogeneous effects for boys and girls. The estimated effects of the PEC for boys, which are conditional to number of years their school had PEC, are about 0.09σ to 0.15σ .

CONCLUSIONS

PEC is a federal government intervention that empowers school administrators, teachers and parents in individual schools to engage in joint strategic planning about how to improve their schools. It also provides monetary resources to enable school communities to implement their improvement plans, many times in the form of infrastructure works. PEC aims at increasing school autonomy among beneficiary schools and improving school quality. It is based on the idea that local agents (principals, teachers and parents) know better about the school needs and how to address them. Moreover, because they are the end users of the service, they internalize the consequences of their decision making, which leads to efficiency gains.

The experiment in Colima was planned to investigate whether PEC increases students' school performance and learning through involving all school agents (parents, teachers, principals and students) in the decision-making process. This results in a better school climate that ultimately favors learning.

Schools in Colima experienced an SBM program before PEC; basic SBM activities were in place in public schools while a formal SBM program with monetary sources started in Colima. After 5 years of PEC, schools selected for the impact evaluation of the program had experienced SBM practices. In addition, these schools were high achievers in terms of the national standardized test before the evaluation, suggesting more efficient use of resources among public schools.

Estimating the treatment on the treated effect, the results suggest that the differences between treatment and control schools were not significant. Test scores increased overall in Colima, for

both treatment and control schools; monetary grants did not boost the SBM mechanisms to change the learning path of the treatment schools or at that stage a more school autonomy was needed. There is a discussion of whether one year is long enough to see differences between treatment and control schools, but using the first-year results, short and medium impacts of PEC among schools cannot be disentangled. This evaluation does not compare schools without the program before the intervention. Within both groups, it is the variance in the years in the program that is being evaluated. On average, both groups have the same years in PEC but within groups, there are schools with short and medium exposure.

However, results also show that capitalizing the advantages of a formal SBM program takes time. Those schools with more time in the program had positive and significant results in their school performance in the first year of the evaluation. Between 2003 and 2012, Colima had high scores in PISA in comparison with the rest of the Mexican states; the PEC program may have helped to support these results.

This evaluation has some lessons for SBM programs: when we analyzed the implementation of the treatment, we found that problems with the design, implementation and state financial resources caused this experiment to alter the incentives of participation and the intensity of the treatment. We cannot clarify if federal spending crowds out local spending leading to insignificant changes in total expenditures. In addition, the contamination of the treatment had an important consequence for the treatment: we do not identify the causal effect of the program for the second and third years. Moreover, the design is crucial; the program had significant bureaucratic issues for a formal SBM program. For example, the program increased the amount of time of the participating schools devoted to red tape and administrative tasks. While the design of the evaluation identified two groups equal at the baseline throughout randomization; the participation incentives are not clear after one year of the program for both the control and treatment groups due to the red tape. That is, the requirement to apply each year to the program even when the school had a program in a previous year and all the requirements of the program creates incentives to not participate.

Moreover, the design of PEC may create a vicious circle on educational outcomes due to the participation strategy of the program and characteristics of those who decide to be part of the selection process. Schools that apply to the program may be different in unobservable and observable characteristics and presumably have more capacities than those who do not apply. These schools receive more benefits than those schools that lag which do not apply. The most deserving schools may lack of the capacity to participate, and it is precisely these schools that are the most disadvantaged.

References

- Abdulkadiroglu, A., J.D. Angrist, S.M. Dynarski, T.J. Kane and P.A. Pathak. 2011. "Accountability and flexibility in public schools: Evidence from Boston's charters and pilots." *Quarterly Journal of Economics* 126(2): 699-748.
- Alvarez, J., V. Garcia Moreno and H.A. Patrinos. 2007. "Institutional effects as determinants of learning outcomes: exploring state variations in Mexico." World Bank Policy Research Working Paper Series No. 4286.
- Bando, R. 2010. The effect of School-Based Management on Parental Behavior and the Quality of Education in Mexico. Ph.D. dissertation, University of California, Berkeley.
- Bloom, N., R. Lemos, R. Sadun and J. Van Reenen. 2015. "Does management matter in schools?" *Economic Journal* 125(584): 647-674.
- Bruns, B., D. Filmer and H.A. Patrinos. 2011. *Making schools work: New evidence on accountability reforms*. World Bank.
- Cabrera Hernández, F.J. and M.E. Pérez Campuzano. 2018. "Autonomía de Gestión para la Calidad y Equidad Educativa: Una Evaluación del Programa Escuelas de Calidad (PEC)." *Revista Mexicana de Análisis Político y Administración Pública* 7(2): 153-174.
- DiGropello, E., and J. Marshall. 2005. "Teacher Effort and Schooling Outcomes in Rural Honduras." In E. Vegas, ed., *Incentives to improve teaching: Lessons from Latin America*. Washington, D.C.: World Bank.
- Duflo, E., P. Dupas M. and Kremer. 2015. "School governance, teacher incentives, and pupil-teacher ratios: Experimental evidence from Kenyan primary schools." *Journal of Public Economics* 123: 92-110.
- Gertler, P., H.A. Patrinos and M. Rubio-Codina. 2008a. "Empowering Parents to Improve Education: Evidence from Rural Mexico." World Bank Policy Research Working Paper Series No. 3935.
- Gertler, P., H.A. Patrinos and M. Rubio-Codina. 2008b. "Impact Evaluation of a School- Based Management Program in Colima, Mexico. First Report Experimental Design and Implementation." World Bank, Washington D.C.
- Hanushek, E.A., S. Link and L. Woessmann. 2013. "Does school autonomy make sense everywhere? Panel estimates from PISA." *Journal of Development Economics* 104 (2013): 212-232.
- Jimenez, E. and Y. Sawada. 1999. "Do Community-Managed Schools Work? An Evaluation of El Salvador's EDUCO Program." *World Bank Economic Review* 13: 415-441.

- Jimenez, E. and Y. Sawada. 2003. "Does Community Management Help Keep Kids in School? Evidence Using Panel Data from El Salvador's EDUCO Program." CIRJE F-236 Discussion Paper, University of Tokyo.
- Khattri, N, C. Ling and S. Jha. 2012. "The effects of school-based management in the Philippines: an initial assessment using administrative data." *Journal of Development Effectiveness* 4(2): 277-295.
- King, E. and B. Ozler. 1998. What's Decentralization Got to do with Learning? The Case of Nicaragua's School Autonomy Reform. World Bank Working Paper Series on Impact Evaluation of Education Reforms No. 9.
- King, E., L. Rawlings and B. Ozler. 1999. "Nicaragua's School Autonomy Reform: Fact or Fiction?" World Bank Working Paper Series on Impact Evaluation of Education Reforms No. 19.
- Loera, A., 2005. Cambios en las escuelas que participan en el PEC, Chihuahua, México: Heurística Educativa.
- Loera, A., Cazares, O., 2005 Análisis de cambios en logros académicos y eficacia social de las escuelas de la muestra cualitativa 2001-2005. Contraste con grupos de control, Chihuahua, México: Heurística Educativa.
- López-Calva, L.F., Espinosa, L., 2006. Impactos diferenciales de los programas compensatorios del CONAFE en el aprovechamiento escolar. In: *Efectos del Impulso a la Participación de los Padres de Familia en las Escuelas*. CONAFE (SEP): Mexico City.
- Murnane, R., J. Willet and S. Cardenas. 2006. "Has PEC contributed to improvement in Mexican Public Education?" In F. Reimers, ed., *Aprender Más y Mejor*. Secretaría Educación Pública (SEP): Mexico City.
- Parker, C., 2005. "Teacher Incentives and Student Achievement in Nicaraguan Autonomous Schools." In E. Vegas, ed., *Incentives to improve teaching: Lessons from Latin America*. Washington, D.C.: World Bank.
- Patrinos, H.A. "School-based management." *Making schools work: New evidence on accountability reforms* (2011): 87-140.
- Santibañez, L., R. Abreu-Lastra and J.L. O'Donoghue. 2014. "School-Based Management Effects: Resources or Governance Change? Evidence from Mexico." *Economics of Education Review* 39: 97-109.
- Schmelkes, S. 2001. School Autonomy and Assessment in Mexico. *Prospects* 31(4): 575-586.

- Shapiro, J. and E. Skoufias. 2005. "Evaluating the Impact of Mexico's Quality School Program: The Pitfalls of Using Nonexperimental Data." World Bank Policy Research Working Paper Series No. 4036.
- Summers, A. and A. Johnson. 1996. "The Effects of School-Based Management Plans." In E. Hanushek, D. Jorgenson, eds., *Improving America's Schools: The Role of Incentives*. Washington, D.C.: National Academy Press.
- World Bank. 2007. "What do we know about School-Based Management?" World Bank, Washington D.C.
- World Bank. 2006. "Mexico: Making Education More Effective by Compensating for Disadvantages, Introducing School-based Management, and Enhancing Accountability: A Policy Note" (Report No. 35650-MX). Latin America and the Caribbean, Human Development.
- World Bank. 2006. "Mexico: Decentralization, Poverty, and Development in Mexico." Colombia and Mexico Country Management Unit, Poverty Reduction and Economic Management Unit, Latin America and the Caribbean Region.

ANNEX

Roster of Schools: Randomized Sample

The random assignment of schools to the treatment and control groups was successful in generating a balanced sample (see Gertler et al 2008b, 2009). The numbers of schools in the treatment and control groups are evenly distributed across the 10 municipalities. In the previous report, it shows the validity of the randomization by comparing the means of a series of indicators of school size and school quality (performance) between schools in the treatment group (PEC beneficiary schools as off PEC VI) and schools in the control group at baseline; that is, in the pre-intervention period or school year 2005-06.

Sample Balance

To verify that the randomization was done properly, we analyze some indicators of the sample of schools in the baseline to check that both groups are equal in expectation, as the treatment and control groups were equally likely to be selected at the baseline through the randomization process. The roster of schools is based on the randomization done in Public General Schools in Colima (for more information, see Gertler et al. 2009).

Annex Table 1 shows school inputs at the baseline. Treatment and control schools do not have any statistical difference in the number of students per teacher, the percent of teachers with university or more, or percent of teachers with a teacher's incentive program (Carrera Magisterial). Finally, the average level of socioeconomic status – the *indice de marginalidad* – is similar between these two groups.

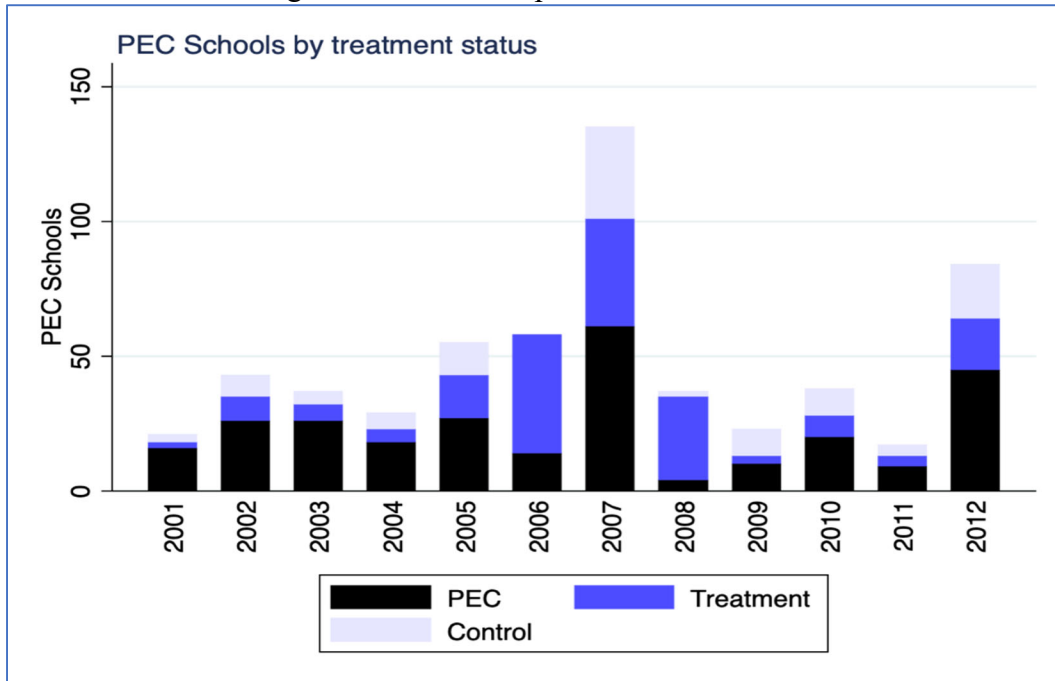
Annex Table 1: Selected Covariates at the baseline, evaluation sample

	Non-PEC	PEC	Diff.	S.E.
Student-Teacher ratio	25.7	27.0	-1.4	1.016
% Teachers with B.A.	0.7	0.7	-0.1	0.048
% Teachers with C.M.	0.7	0.6	0.1	0.065
Average <i>marginalidad</i>	4.4	4.3	0.1	0.197

Source: School Census 911 and Evaluation roster

Information of the PEC schools was provided by the unit at SEP in charge of PEC. This unit consolidated a roster of beneficiary schools and all the documentation of each state and each year of the program. Annex Figure 1 shows the total number of PEC schools by treatment status, it depicts that the experiment was not implemented completely: (1) the first year of the RCT, there are schools that received PEC outside the experiment and (2) control schools received PEC in 2007, 2008 and 2009. While for the first year of the experiment, it is possible to compare treatment and control schools, for the following two years of the experiment, the internal validity is questioned by “cross-over” participation in both directions. Annex Figure 2 shows that schools in the evaluation sample had PEC before the treatment.

Annex Figure 1: PEC Participation in Colima, 2005-2012



Source: Authors' calculation using the administrative data from PEC.

Annex Figure 2: PEC Participation by Treatment Status, 2005-2012

