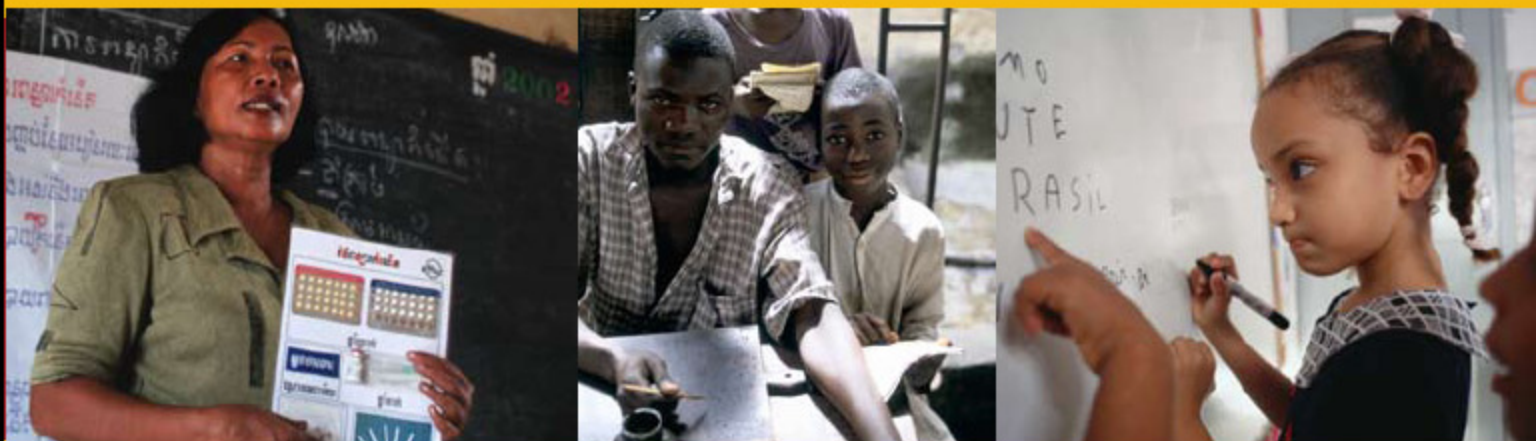# WHAT DO WE KNOW ABOUT SCHOOL-BASED MANAGEMENT?

THE WORLD BANK

# *What Do We Know About School-Based Management?*

# *What Do We Know About School-Based Management?*

November 2007

Education

Human Development Network

**THE WORLD BANK**
Washington, DC

# Contents

## *Box*

## *Tables*

For more than 100 years, the lack of a school management methodology has been the cause of countless complaints. But it has been only in the last 30 years that efforts have been made to find a solution to this problem. And what has resulted so far? Schools continue exactly the same as before.

**Jan Amos Comenius, 1632**

# *Preface*

School-based management has become a very popular movement over the past decade. Our school-based management work program emerged out of a need to define the concept more clearly, review the evidence, support impact assessments in various countries, and provide some initial feedback to teams preparing education projects. During first phase of the School-based Management work program, the team undertook a detailed stocktaking of the existing literature of School-based Management. At the same time we identified several examples of School-based Management reforms that we are now supporting through ongoing impact assessments. An online toolkit providing some general principles that can broadly be applied to the implementation of SBM reforms has been developed and can be accessed on http://www.worldbank.org/education/economicsed.

See companion piece: ***What Is School-Based Management?***

# *Acknowledgments*

## Introduction

Impact evaluations of school-based management (SBM) programs, or any other kind of program, are important because they can demonstrate whether or not the program has accomplished its objectives. Furthermore, these evaluations can identify ways to improve the design of the program. For example, they can establish whether the population that is benefiting from the program is the intended target population. These evaluations can also make successful interventions politically sustainable and can create a consensus on a plan for reforming an unsuccessful program. This report presents recent evidence from impact evaluations of SBM programs. Despite the clear benefits of conducting impact evaluations, very few have actually been carried out on SBM interventions around the world. This means that very little evidence exists both in terms of the number of cases studied and in terms of the strength of the empirical methodologies used to evaluate the interventions. This report presents the findings of the few rigorous and well-documented studies that exist and makes some recommendations about how to improve and increase the evidence base in the future.

The fundamental problem in estimating the impact of SBM programs is that communities and schools either self-select into the program—a problem known in the literature as selection bias—or the authorities choose the participants using some selection criteria, which can lead to program placement bias. Under these conditions, making a comparison between communities or schools that participate and those that do not participate confounds the effects of the program with the initial differences in characteristics between participants and nonparticipants. For example, SBM schools in homogeneous and highly participatory communities can be expected to perform differently than schools in heterogeneous and fragmented communities. Therefore, their final educational outcomes are likely to be the result not only of the SBM intervention but also of inherent differences in the characteristics of the communities.

Likewise, in any comparison over time between different communities that have implemented SBM programs, the impact of the program will be mixed up with any other changes that may have occurred during the period of the intervention, such as modifications in educational laws or an economic recession. Usually this comparison strategy of estimation is used in cases in which the reform applies to all schools, and at the same time, so that there is no group of schools to which the reform does not apply—in other words, there is no counterfactual. For example, the Chicago School Reform Act of 1988 was implemented in all city schools and at the same moment (Hess 1999 and Bryk et al. 1998). In this case, constructing a comparison group of schools to which the reform does not apply would be extremely difficult.

Not only do communities or schools sometimes self-select into SBM programs, but also students may reinforce the problem by self-selecting into SBM schools, which leads to sorting bias. Once the reform has become operational, it is feasible that families may try to enroll their children in SBM schools or, on the contrary, to enroll them in the non-SBM schools. Thus, this unknown sorting behavior may make it difficult to unravel the effects of the SBM program with the differences in the characteristics of the students. Presumably, students who choose

to enroll in schools with SBM are different from students in other schools. Therefore, any simple comparison between schools that are participating in the program and those that are not participating may pick up not only differences in educational outcomes that are due to the SBM program but also the differences in the characteristics of the students and their families in the two groups of schools.

The direction of the bias in simple comparisons between SBM schools and other schools is not clear. For example, schools in Nicaragua are self-selected to participate in the Autonomous School Program (ASP). The program was first implemented in 1991 and gives wide autonomy to schools (di Gropello 2006). The teachers in each school can vote to be part of the SBM. On the assumption that teachers who vote are more motivated and more active, it is likely that making a simple comparison between non-ASP schools and ASP schools will lead to upward biases that will make the impact of the program seem bigger than it actually is. (Annex 1 presents a brief summary of some evaluated SBM programs.)

On the contrary, in programs that target certain communities, typically low-income ones, making simple comparisons between those communities that do and do not participate in the program can lead to downward biases. For example, PROHECO (*Projecto Hondureño de Educacion Comunitaria* or the Honduran Project for Community Education) in targets rural areas affected by Hurricane Mitch (di Gropello 2006). In this case, the Ministry of Education chooses the communities in which the program operates. Presumably, making a simple comparison between PROHECO schools and other schools may lead to downward biases since the target schools are in poor rural areas.

In short, the key challenge in measuring the impact of SBM programs is to find a good control group of schools to compare with those that have benefited from the SBM program. In other words, the challenge is to find the right counterfactual. It is necessary to find ways to estimate the impact that can take into account the process of selection.

This report discusses the challenges associated with establishing impact evaluations in the SBM setting. It presents the most robust evaluations of different SBM programs around the world, classifying them by type of evaluation—randomization, regression discontinuity analysis, instrumental variables, difference in differences, and matching estimators. The report assesses the strength of the literature and discusses key aspects of evaluating SBM, such as how the intervention affects educational outcomes and how quickly the impact is likely to be seen. Also, it gives guidance on how best to implement an impact evaluation.

## School-Based Management Program Evaluations

In general terms, a good evaluation should include three important steps (Gertler et al. 2007):

1. *A clear definition of the intervention.* All interventions modify margins and incentives differently for different stakeholders. It is critical to define what is being modified in the program, the new set of incentives, and to whom the modifications apply.

2. *A description of how the intervention is expected to achieve the final desired outputs.* Understanding how the intervention will lead to the desired result is fundamental for the evaluation. In general terms, sound economic theory (see World Bank 2007) should guide the analysis of how the intervention will affect the desired outcomes.

3. *A definition of the identification strategy.* An identification strategy is the mechanism by which it is possible to attribute causal effects between an intervention (for example, the SBM program) and a set of outcome variables (for example, educational outcomes such as dropout rates or standardized test scores). In order to be able to attribute changes in outcome variables to the program, it is necessary to overcome the problems of self-selection.

In the case of SBM programs, these three steps that are essential to performing a rigorous impact evaluation are particularly challenging. Defining the intervention is very difficult because of the complexity of the

SBM concept. Likewise, how the intervention is likely to achieve the desired results will depend on the complexity of the specific intervention in question. Finally, it is difficult to identify causal effects because of the three sources of bias—self-selection of schools, the selection of schools by authorities, and the process by which students are enrolled in the SBM schools.

## Defining the Intervention

SBM programs can take on different forms in terms of who has the power to make decisions as well as the degree of decision-making devolved to the school level. While some programs transfer authority only to school principals or teachers, others encourage or mandate parental and community participation, often in the form of school committees. Most SBM programs transfer authority over one or more activities. These could be: budget allocation; the hiring and firing of teachers and other school staff; curriculum development; the procurement of textbooks and other educational material; infrastructure improvement; and the monitoring and evaluation of teacher performance and student learning outcomes. While we define SBM broadly to include community-based management and parental participation schemes, we do not explicitly include stand-alone or one-off school grants programs that are not meant to be permanent alterations in school management (World Bank 2007).

Based on this definition, the two key dimensions of the intervention are: first, to whom the power is transferred; and, second, what types of decisions they are authorized to make. It is important to identify both aspects in order to define the intervention.

In terms of the first dimension, SBM policies can transfer power to parents, communities, schools, or a combination of all of these. Within a school, the transfer can be to the principal or head of the school, the teachers, and, in some cases, even the students. For example, the 1988 Chicago reform transferred power to both schools and communities, while reforms in El Salvador (1991) and Honduras (1999) transferred power to local communities. However, there are also

cases (for example, Nicaragua in 1991) where the transfer of power has not been as clear, making the evaluation of that program more difficult. Di Gropello (2006) presents a general review of the Central American cases of SBM and Bryk et al. (1998) describes the process in Chicago.

On the second dimension—the type of decisions over which authority is devolved—the transfer of power can apply to a limited number of functions or to a wider range of functions. An example of a limited transfer would be a policy giving the school or community a specific amount of money for any infrastructure improvements that they may deem to be necessary, as in the Apoyo a la Gestión Escolar (AGES) school management support reform in Mexico (Gertler et al. 2006). The transfer of power can also involve several different aspects of the educational process, such as decisions about the hiring and firing of personnel, the curriculum, what pedagogic method should be used, and what type of infrastructure investments should be made. In Nicaragua, authority over almost all of the operational aspects of school management was devolved to the school level, ranging from the hiring of teachers to the maintenance of infrastructure (di Gropello 2006).

## Elements of Impact

How the intervention will produce the desired outcomes depends on which type of SBM program is adopted. The design of the intervention can be very complex involving several stakeholders and several inputs, or it can be a simple change in the allocation of a specific resource.

The branch of the SBM literature written by education experts (for instance, Bauer et al. 1998) suggests that the impact of SBM programs can be measured by three elements—"scope," "decision-making," and "trust." "Scope" refers to the clarity of goals set by the members of the school council or the extent of the influence that the school has over input decisions. "Decision-making" practices refer to the actual implementation practices of the school council. "Trust" refers to the interaction between the members of the community or council and parents.

This literature (for example, the original work of Bauer 1996, 1997, and Bauer et al. 1998) has created several instruments to measure these elements. However, the instruments and the scale of measurement are difficult to put into practice. For instance, several of the proposed measures are perceptions, which are subjective and difficult to compare. For this reason, this report suggests another course of action. Based on the economic theory behind SBM programs, it proposes a different set of indicators by which to measure internal changes in the SBM schools. When inputs inside the school change (what is referred to in the literature as "inside the black box"), educational outcomes can change as well. Table 1 presents these two different kinds of indicators for measuring the outcomes of SBM programs in schools.

The theory of SBM emphasizes that there are several ways in which this kind of intervention can change educational outcomes (Gertler et al. 2007 and Santibañez 2006). First, one of the main ideas behind SBM is that those at the local level (community members, parents, school staff, and students) have more information about the school than the central government does. This means that local people will make better, more appropriate choices for the school than the centrally based Ministry of Education or even the local education authority. In this sense, it is important to track changes inside the school in the following areas:

a. *Key decisions about personnel (teachers and administrative personnel) such as firing, hiring, rotation time, and teacher training.* It is important to know not only which aspects of these variables have been devolved to the school level and the frequency with which they are decided

**Table 1    Inside the Black Box: How to Measure the Impact of School-Based Management Programs**

| Dimension | Objective | Type of question | Examples of Questions / topics |
|---|---|---|---|
| **A. Education literature** | | | |
| Scope | Clarity of goals and the real influence of the board | Self-diagnosis; "site team" (e.g., the community, council, or board of the school) | Site team members (…) agree on what kinds of decisions the team may and may not make; or the site team has real influence on issues of importance |
| Decision-making | Actual implementation practices | Self-diagnosis; "site team" | Members work to implement decisions once they have been made; or members work to correct problems that arise during the implementation of team decisions |
| Trust | Interaction between members | Self-diagnosis; "site team" | All members of the site team have an equal opportunity to be involved in decisions; or site team members communicate openly and honestly during meetings |
| **B. Economic literature** | | | |
| Information at the local level | Changes in decisions | Key decisions about personnel (teachers and administrative) | Firing, hiring, rotation time, teacher training, among others: who makes these decisions |
| | | Key decisions about spending | Spending on infrastructure, training of teachers |
| | | Changes in educational process | Change in pedagogical methods; changes in allocation of time; absenteeism of teachers |
| | | Resource mobilizations | Amount of resources from community into the school |
| Accountability and monitoring | Involvement of parents and community in the school; better accountability and monitoring | Direct involvement of parents and community in the school | Power of board; type and number of meetings; decisions in meetings |
| | | Links between parental involvement and decisions at the school level | Do complaints / praise about teachers translate into decisions about the teacher |
| | | Changes in the accounting systems of the school | Implementation of EMIS; changes in account tracking system |
| | | Changes in the climate of the school | Changes in attitude of teachers and students about the school |

*Sources:* Education literature: Bauer et al. (1998). Economic literature: Gertler et al. (2007).

upon, but also who exactly makes the decisions. For instance, is it the community or parents who have the real power to hire and fire teachers?

b. *Key decisions about spending.* It is important to track changes in the magnitude of spending on infrastructure, administration, and training of personnel. Also, it is critical to determine who made those investment decisions.

c. Changes in the educational process. It is important to record any changes in pedagogic methods, such as the way in which teachers conduct their classes, and the extent to which students are encouraged to participate (passive versus active activities) in the classroom. SBM may change how teachers allocate their time between teaching, administrative tasks, and meetings with parents/community members. Also, SBM can change the rate of teacher absenteeism.

d. *Resource mobilization.* Greater community and parental involvement in school affairs can sometimes lead to the school receiving more private donations and grants on top of the money that the school receives from the national government or from local taxes.

The second way in which SBM can theoretically change educational outcomes is by promoting more involvement by the community and parents in the school, and by holding accountable and monitoring those making decisions about school management. Along these lines, it is important to look into the following items:

a. *Direct involvement of parents and community in the school.* It is important to ascertain what formal mechanism of interaction exists (for example, a school council) among community members, parents, and the school and to identify its members. Also, it is critical to find out how many meetings there have been between the community and the school, as well as the type of meeting (for example, meetings at which decisions were made or that were just for informational purposes).

b. *Links between parental involvement and decisions at the school level.* For example, it is important to know if systematic complaints or praise about a teacher by parents/community members ever translate into the firing or promotion of the teacher. It is also important to know if their suggestions about infrastructure problems lead to expenditures being made to solve those problems.

c. *Changes in accounting.* Community members and parents, by involving themselves in school affairs, can persuade the school to improve its Education Management Information System (EMIS), its systems for tracking students' academic progress, and its systems for tracking financial inputs. In turn, these changes can improve the administration of the school and, eventually, educational outcomes. For example, if having a better EMIS liberates teachers from administrative tasks, then they will have more time available to spend teaching.

d. *Changes in the school climate.* Community involvement can change the school climate either positively or negatively. It is important to gather information on the attitudes of teachers and students toward the school, for example, by asking direct questions about their level of satisfaction with the content of classes, among other issues.

One of the complexities that must be contended with in evaluating the impact of SBM programs is timing. In general terms, SBM reforms take a long time to produce their expected outcomes. In the first year or so of an SBM reform, there is an adjustment period during which changes in personnel occur and management changes (for example, the creation of a school council) are gradually put into operation. In the short run, these adjustments can have a negative impact on educational outcomes, but once the school adjusts to the changes, positive changes can be expected.

The speed of the effect depends as well on the type of outcomes being assessed. Some outcomes can be expected to change faster than others because the incentives

**B O X   1.**     *800 Models, 29 Evaluations, and 8 Years to Yield Results*

In a meta-analysis of the effectiveness of SBM models in the United States—or Comprehensive School Reform (CSR)—Borman et al. (2003) reviewed 232 studies with 1,111 independent observations that evaluated 29 CSR programs in the United States. From these observations, they computed the size of the effect that these models had had on student achievement. The authors regressed weighted effect size on the moderator variables to obtain the residuals from the regression and added the mean weighted effect size to each observation, thus calculating effect sizes that were statistically adjusted for all of the methodological variables. They found that the number of years of implementation of the CSR was a statistically significant predictor of the student achievement effect size.
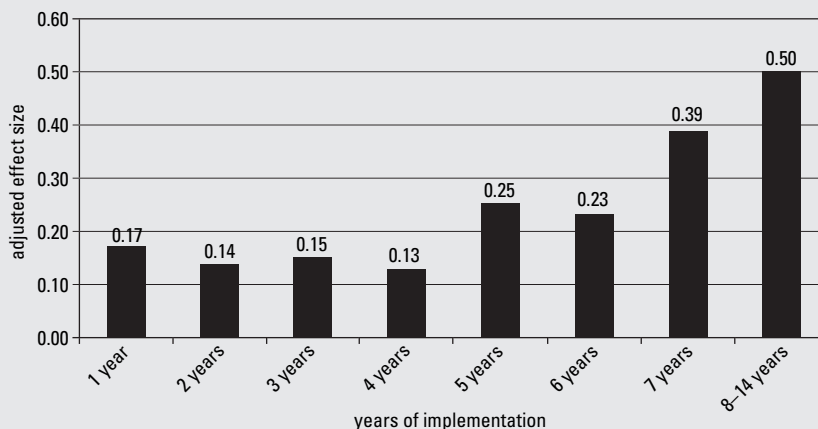


*Source:* Borman et al. 2003.

that drive them are easier to change. For instance, attendance rates, measured by the number of days on which a student is present at school, are easier and faster to change than enrollment rates. So, in the short run, an SBM intervention can have a positive impact on attendance, repetition, and failure rates, but outcomes such as dropout rates or test scores take longer to change.

In the United States, it has been argued that SBM needs about five years to bring about fundamental changes at the school level, and about eight years to yield changes in difficult to modify indicators such as test scores (Borman et al. 2003 and Cook 2007). Box 1 synthesizes the evidence of 800 models and 29 evaluations to test this hypothesis and concludes that the projects started to deliver results after an average of eight years. However, it is important to find robust evidence to back up this general assumption for each instance of SBM reform, especially in developing countries, given the wide range of different designs that is possible for SBM programs.

## Identification

As discussed in the introduction to this report, identifying or isolating the impact of SBM programs is difficult because of program placement bias, self-selection bias, or sorting bias in how communities, schools, and students are selected to participate in the program.

In the impact evaluation literature, the "gold standard" of identification strategies is the randomization of treatment (Shadish et al. 2002 and Duflo et al. 2006). However, in the absence of randomization, it is possible to estimate the true impact of the interventions using other techniques such as regression discontinuity analysis, instrumental variables, Heckman correction procedures, difference in differences estimators, and matching estimators. The first set of methods—regression discontinuity, instrumental variables, and

Heckman correction procedures—estimate the effects of a program either by using the entry rule to participate in the program or by modeling the program participation decision. The second set of methods—difference in difference and matching estimators—construct a comparable control group that has not benefited from the program.

Randomization and regression discontinuity analysis both provide estimates of the true effects of programs; in other words, their estimates are unbiased. In many cases, however, the design of the program does not allow for these types of analyses. In contrast, instrumental variables, difference in difference, and matching estimations can be used when the policy design is not an experiment or when there are no definite cut-off criteria. However, the validity of these methods depends on some assumptions that are, in some cases, difficult to meet.

The following sections will discuss each of these techniques with reference to the empirical literature of SBM programs. Adapted from Santibañez (2006), Table 2 presents general descriptions of the most rigorous evaluations of SBM programs that have been conducted since 1995. The objective of this section is to present the strengths of the literature rather than present an array of SBM cases. A brief review of the programs that have been evaluated since 1995 is presented in Annex 1. For a more complete description of the interventions, see di Gropello (2006), Paes de Barros and Mendonca (1998), and Gertler et al. (2006).

It is important to highlight two ideas from the outset before reviewing the empirical literature on SBM. First, only a very few rigorous studies of the impact of SBM exist. Santibañez (2006) consists of a literature review of the 53 evaluations carried out since 1995 of the impact of SBM programs on educational outcomes. This report deliberately discusses only those studies that made a clear attempt to correct problems of endogeneity, which reduces the original number of 53 to 13.

Second, despite the fact that, to our knowledge, these 13 studies are the best estimates available, some of them have serious limitations. For instance, five articles used instrumental variables approaches with questionable instruments. Four articles used matching estimation, some of them with limited or even with no baseline information. Only two of the articles that used difference in difference estimations verified the equality of trends between the control and treatment groups before the intervention. Nevertheless, these 13 articles represent the best effort to date to estimate the effects of SBM, albeit with limited data.

Another challenge in the review of the empirical literature is to evaluate the size effects of the impact due to the heterogeneous presentation of metrics and results in the different studies. Several studies only reported the estimated coefficient of impact and, therefore, it is very difficult to translate these into impact size because they depend on the specific measurement of both the independent and dependent variables. Others presented information on the percentage changes in some outcome variables due to the intervention. Again, the metric of the output variables is very different from one study to the other. Nonetheless, we report the size of effects for those studies that have a clear interpretation of the results; otherwise we indicate the direction and significance of the coefficient of impact.

*Randomization and Regression Discontinuity Designs.* Randomization and regression discontinuity designs (RDD) produce unbiased estimators of the impact of SBM programs. Unfortunately, no evaluations have been done since 1995 of the effects of SBM on educational outcomes using randomized or RDD evaluations.

Impact evaluation using randomization strategies is based on the idea that a lottery will *de facto* create similar treatment and control groups in terms of observable and unobservable characteristics. In this sense, the mean of observable variables and unobservable variables will be equal across groups. The only difference between the treatment and control groups is the intervention. Therefore, any differences in outcomes can be attributed solely to the program.

For example, in an extreme case of using randomization to assess changes in SBM schools, randomization would be performed at two levels. First, the schools that are to participate in the SBM program

**Table 2    Evaluations and Impacts: Evidence of School-Based Management from the More Rigorous Studies since 1995**

| Study | Country | Program | Date of program | Year of study | Estimation / identification strategy | Limitations | Results |
|---|---|---|---|---|---|---|---|
| **A. Randomization and regression discontinuity design** | | | | | | | |
| No available evidence | | | | | | | |
| **B. Instrumental variables and Heckman correction models** | | | | | | | |
| di Gropelo and Marshall (2005) | Honduras | PROHECO | 1999 | 2003 | Heckman correction model; exclusion restriction: presence of potable water and community services | Not a solid exclusion restriction | Small changes in dropout rates; no effects on test scores |
| Gunnarsson et al. (2004) * | Several countries | Several programs | Several years | 1997 | Instrumental variables: principal's attributes and legal structure | Not a solid instrument | No impact on test scores; positive impact on parental participation |
| Jimenez and Sawada (1999) | El Salvador | EDUCO | 1991 | 1996 | Heckman correction model; exclusion restriction: government prioritizing targeting formula | Not a solid exclusion restriction | Increases reading scores and decreases absenteeism |
| Jimenez and Sawada (2003) | El Salvador | EDUCO | 1991 | Panel: 1996 and 2003 | Heckman correction model; exclusion restriction: government prioritizing targeting formula | Not a solid exclusion restriction | Increases probability of students staying in school |
| King et al. (2003) * | Several countries | Several programs | Several years | Two points: 1995 and 1997 | Instrumental variables: principal's attributes and legal structure | Not a solid instrument | No effects on test scores |
| **C. Difference in difference and matching estimation** | | | | | | | |
| Gertler et al. (2006) | Mexico | AGE | 1996 | Panel at school level 1998–2002 | DD fixed effects; pre-intervention trends | Did not control for time variant unobservable effects | Positive impact on failure and repetition rates; no effects on dropout rate |
| King and Ozler (1998) | Nicaragua | ASP | 1991 / 1993 | Pseudo-panel: 1995 and 1997 | Matching, panel data | No pretrend validation | De jure autonomy: no impact. Real autonomy (hire and fire teachers): positive impact on standardized test scores |
| Lopez-Calva and Espinosa (2006) | Mexico | AGE | 1996 | 2003–2004 | Matching, cross-section | No baseline data | Positive impact on test scores |
| Murnane et al (2006) | Mexico | PEC | 2001 | Several sources: 2000–2004 | DD: more systematic check of equal trends between treatment and control groups | Not control for time variant unobservable effects | Positive impact on dropout rates; no effect on repetition |
| Paes de Barros and Mendonca (1998) | Brazil | Decentralization | 1982 | Panel, state level; 1981–1993 | DD; no preintervention trends | Aggregation of data; no pretrend validation | Positive impact on repetition and dropout rates; no impact on test scores |
| Parker (2005) | Nicaragua | ASP | 1991 / 1993 | 2002 | Matching, panel data | No pretrend validation | Positive impact on test scores |
| Sawada and Ragatz (2005) | El Salvador | EDUCO | 1991 | 1996 | Matching, cross-section | No baseline data | Positive impact on test scores |
| Skoufias and Shapiro (2006) | Mexico | PEC | 2001 | 2000–2003 | Matching estimation with DD; one year preintervention trend | No pretrend validation | Positive impact on dropout, failure, and repetition rates |

* School self-reported levels of autonomy
Source: Cited articles and Santibañez (2006).

are picked by chance and then students are randomly assigned to the SBM schools. The only difference in educational outcomes, such as dropout rates, between the SBM and the other schools can thus be attributed to the intervention since there was no self-selection.

Usually randomized experiments collect data for a minimum of two points in time. Data on the treatment and control communities, schools, and students are collected before the intervention (baseline information), and then data on the same indicators are collected after the implementation of the program. The baseline data set is important because it can be used to test that the randomization was correctly implemented and that the two groups (treatment and control) are similar in (at least) the observable characteristics—in essence the baseline validates the randomization. In the case of SBM, the outcome variables can be processes, like the ones described in Table 1, or educational variables like repetition rates, dropout rates, absentee rates, failure rates, and test scores.

The timing of the follow-up data collection is critical in SBM reforms. Collecting these data too soon after implementation will probably reflect only the adjustment period and may show the program's impact to be *negative*. However, after the adjustment period, SBM policies can be expected to start delivering *positive* results and, therefore, it is important to allow a sufficiently long period of time to pass before collecting follow-up data. Also, it is advisable to collect more than one round of follow-up data.

SBM policies are often very complex interventions. Even if it is possible to randomize and, thus, attribute any difference in educational outcomes to the SBM program, it is not possible to attribute the impact to any specific change of the many that may have been brought about by the program. For example, an SBM program may change both the decision-making process involved in hiring teachers and how teachers allocate their time. Even if, using a randomized experiment, we were to discover that the program had changed educational outcomes, it would be difficult to distinguish whether these improvements in educational outcomes were due to the change in hiring practices or those related to how teachers spend their time. For this reason, it is crucial to analyze all internal changes in the school to understand which specific changes at the school level are affecting educational outcomes.

It is possible to use RDD if the program identified its beneficiaries using an assignment variable. For example, in some states in Mexico, the Programa Escuelas de Calidad (PEC) program uses a poverty index that is also used by the conditional cash transfer program *Progresa / Oportunidades* (Skoufias and Shapiro 2006) to identify schools that qualify for the program benefits. Other states rank schools by the quality of their improvement plans**.** Regression discontinuity analyses can be used in such cases since they make use of the assignment variable and the observations with scores close to the cutoff point to establish eligibility for the program. If all schools with a score below a certain cutoff are enrolled in the program and those with a score above the cutoff are denied access to the program, then schools with scores just below the cutoff point (beneficiaries) may be very similar to those schools that are just above the cutoff point (the comparison group). In this case, it is possible to compare the outcome variables for those two groups and attribute the differences to the program, given that we expect the schools in the two groups to have very similar characteristics. Regression discontinuity analysis resembles a randomization since, from the point of view of the school, to be "just below" or "just above" the arbitrary cutoff point is almost like taking part in a lottery.

The difficulty with this approach, however, is the potentially limited number of observations around the cutoff point. Since RDD estimates the effects of the program using observations around the cutoff point, it requires a smooth assignment variable with a large number of observations on both sides of the cutoff value. If there are only a few observations, then the estimate of the impact will be very imprecise.

Furthermore, it is important to note that RDD is a local estimator; in other words, the estimation gives the evidence of the program's impact on individuals close to the cutoff point but says nothing about its impact on those individuals with low (or

high) scores. On the one hand, this characteristic of the RDD is a limitation since it is not possible to estimate the average effect of the program. On the other hand, this characteristic can be desirable since, in certain situations, the most relevant impact is that on the margin—the impact close to the cutoff point.

*Instrumental Variables and Heckman Correction Models.* Both instrumental variable estimation (IV) and Heckman correction models base their identification strategy on a variable that can explain the participation of communities and/or schools in the program (Angrist and Imbens 1995 and Heckman 1976).

The IV approach uses a variable with two characteristics—it can explain participation in the program but is uncorrelated with the outcome measures of interest. For example, the evaluator of a hypothetical training program that targets people born in a certain month of the year may want to determine the impact of the training program on the probability of its graduates becoming employed. In this case, given that the candidates' birth month is correlated with their entry into the program but is presumably not correlated with the probability of them being employed, the month of birth can be used as an instrumental variable.

The main problem with the instrumental variable approach is to find a valid instrument—in other words, a variable correlated with the decision to participate but not with the final outcome of interest. Most available variables that are correlated with participation are correlated with the outcome as well. Even if it is possible to find a variable that it is correlated with participation, it is impossible the test whether the variable is uncorrelated with the unobservable part of the outcome variable.

Two studies used IV to estimate the effects of SBM. More precisely, these two articles studied the effect that self-reported school autonomy has had on test scores. Gunnarsson et al. (2004) used regional test score data from 1997 in several Latin American countries, and King et al. (2003) complemented these data with results from the international standardized test Trends in International Mathematics and Science Study (TIMSS) in 1995. Both are cross-section, country-level estimations. The instrument that they used is the legal structure of the country (political stability, regulatory quality, and rule of law). This variable is presumably correlated with participation in the program. However, it is very feasible to argue that the variable is correlated with educational outcomes. As we described above, it is critical that the instrument is not correlated with the outcomes, and, therefore, the estimation strategies used in these two studies present serious problems. In any case, neither study found that SBM reforms, or more precisely self-reported school autonomy, had had any impact on test scores. According to Gunnarsson et al. (2004), scores in schools with the greatest autonomy are between 4 percent *higher* and 13 percent *lower* than in less autonomous schools.

The Heckman correction method is based on the estimation of two equations. First, it models the participation decision. For example, the dependent variable is an indicator of program participation as a function of variables likely to influence the decision to participate in the program. Second, it estimates the program's impact by regressing the outcome variable against the unexplained component of the participation equation—the residuals from the participation decision equation—and other variables (Heckman 1976).

In the Heckman correction model, there are two ways to identify the true impact of the program. The first method is to rely on assumptions about the distribution of the errors in the participation and outcome equations, but these assumptions are very unlikely to be valid. The second method is to use an "exclusion" variable, a variable that is in the participation equation but not in the impact equation, to estimate the impact. Clearly, this second method is very similar to finding an appropriate instrumental variable that can explain participation but not the final outcome and, thus, is as difficult to implement as an IV methodology.

Using the targeting formula as the identifying variable in a Heckman correction model, Jimenez and Sawada (1999) analyzed the case of EDUCO in El Salvador

(Annex 1 describes the program). They found that SBM had increased standardized test scores and reduced both student and teacher absenteeism. Also, they found that parents participated more in the EDUCO schools than in schools that were not in the program. Jimenez and Sawada (2003) used the same identification strategy, but this time with panel data for 1996 and 1998. They found that SBM had a positive impact on the probability of students staying in school. As with the previous cases, the validity of the instruments used in these studies is questionable. In short, it is very likely that the targeting formula of the program is correlated with educational outcomes, which would invalidate the instrument that both studies used.

Using a two-stage procedure, di Gropello and Marshall (2005) evaluated the impact of the Honduran PROHECO program. Their exclusion variables were community services and the presence of potable water. Once they corrected for selection, they found that SBM had no effect on either teachers' efforts or test scores. Once more, it is difficult to argue that the IV was not correlated with the outcome variable.

In short, out of the five studies using IV or Heckman procedures, only two showed that SBM had a positive impact on test scores, and only two of them found that it had had a positive impact on dropout rates and on the probability of staying in school.

*Difference in Differences (DD) and Matching Estimation (ME).* The richest evidence on SBM has been found using DD and ME. Some of the programs have extensive data sets that made it possible to use these two strategies to evaluate their impact. DD and ME methods generate a counterfactual using nonbeneficiaries with similar characteristics as the beneficiaries. In DD, the true effects of a program are identified by verifying the similarity of trends in observable characteristics between the treatment and control groups (Athey and Imbens 2006). In contrast, ME uses all of the observable baseline characteristics to find close matches in the control group for each treated observation (Rosenbaum and Rubin 1983 and Heckman et al. 1998).

DD is more demanding in terms of data than ME. In DD, it is necessary to have data for at least three moments in time—preintervention trends (that is, at least two data points before the intervention) and data capturing the changes that have occurred since the intervention was implemented. This amount of data is rarely available. Moreover, it is common to find studies that use data for only two moments in time, one observation before the intervention and one after for each participant. Results obtained in this way cannot be validated; in other words, it is impossible to say whether the estimated impact was due to the program or was a trend that already existed between the two groups prior to the implementation of the program.

Nonetheless, DD estimation has one important property. When it is estimated using fixed effects (for example, a dummy variable for each unit of observation and a dummy variable for each time period), DD controls for time invariant unobservable and observable differences between the control and treatment groups. In other words, the fixed effects estimation controls for differences between the two groups in both observable and unobservable characteristics that do not change over time.

Using ME in an impact evaluation requires rich and abundant baseline data. Furthermore, using ME requires that the process for selecting program participants be based only on observable characteristics. If some unobservable characteristic plays a role in the selection process, then the ME estimate will be biased. Moreover, due to data limitations, several impact evaluations using ME have been forced to use data to match the treatment group with a control group that was put together only after the program was already being implemented. This procedure creates problems when the observable characteristics used for selecting program participants also change because of the intervention (Rosenbaum and Rubin 1983).

Evidence of the impact of the Apoyo a la Gestión Escolar (AGES) *School Management Support* program in Mexico is presented in Gertler et al. (2006). The authors used the order in which schools entered into the

program to construct a DD estimator that controlled for fixed effects. They presented preintervention trends between the control and treatment groups and found no differences in educational outcomes prior to the intervention, thus validating the use of the DD strategy. They found that the program reduced repetition rates in 4 percent and failure rates in 4.2 percent of the treatment schools, but they did not find any impact on dropout rates. On the other hand, Lopez-Calva and Espinosa (2006), with data from 2003 to 2004 and using matching techniques, found that the AGES had had a positive impact on test scores. The main limitation of this study was the lack of baseline data.

To estimate the effect of decentralization of school autonomy in Brazil, Paes de Barros and Mendonca (1998) (see also Carnoy et al. 2004) constructed a panel data set at the state level between 1981 and 1993. They used a DD strategy with a fixed-effects model. The level of aggregation of the data (state level) meant that they were faced with the problem of having to evaluate the program's impact with only a limited number of observations. In any case, they found that SBM had had a positive impact on dropout rates (reductions of between 3.4 and 6.6 percent) and repetition rates (reductions of between 1.7 and 4.2 percent), but that it had had no effect on test scores.

Two articles evaluated Mexico's PEC program—a voluntary, urban-based program open to all public schools—using DD estimators. Shapiro and Skoufias (2006) and Murnane et al. (2006) used the same data source with the difference that Murnane et al. incorporated one more year of observations. Shapiro and Skoufias used a matching DD estimation. Murnane et al. argued that the counterfactual of Shapiro and Skoufias had different preintervention trends and, therefore, they created another counterfactual using a new group of schools that had just entered into the program. They found that SBM had reduced dropout and failure rates by 0.24 percentage points and repetition rates by 0.31 percentage points. In contrast, Murnane et al. found a positive effect only on dropout rates (of 0.27 percentage points).

Evidence of the impact of the EDUCO in El Salvador using ME is presented in Sawada and Ragatz (2005). One major limitation of this study is the lack of baseline data. The authors found that SBM increased the amount of time that teachers could spend on teaching, which in turn translated into a positive impact on test scores.

In summary, six studies used DD and ME. Three of them presented evidence that SBM had had a positive impact on test scores, and the majority of the studies presented evidence that SBM had had a positive impact on dropout, failure, and repetition rates.

## Assessment of the Literature

One first general conclusion on the evidence base of SBM since 1995 is that the sample of carefully documented, rigorous impact evaluations is very small, given the large number of known SBM programs that exist around the world. This situation is changing, but we know very little about the effects of SBM at this time. Moreover, the few rigorous studies reviewed here used empirical strategies that are open to question.

Nonetheless, these studies represent an important effort to quantify the impact of some SBM programs on educational outcomes. It can be argued that they have reduced the bias that is undoubtedly present in simple comparisons and, in this way, have made important advances in our understanding of the impact of SBM policies.

Despite the fact that it is very difficult to establish the sizes of the outcome variables of interest because of the different metrics used in the various studies, it is nevertheless possible to list some findings about the impact of SBM based on the more rigorous analyses:

1. Some studies found that SBM policies actually changed the dynamics of the school, either because parents got more involved or because teachers' actions changed (King and Ozler 1998; Jimenez and Sawada 1999; and Gunnarsson et al. 2004).

2. Several studies presented evidence that SBM had had a positive impact on repetition rates, failure rates, and, to a lesser

degree, dropout rates (di Gropello and Marshall 2005; Jimenez and Sawada 2003; Gertler et al. 2006; Paes de Barros and Mendonca 1998; and Skoufias and Shapiro 2006).

3. The studies that had access to standardized test scores presented mixed evidence (Jimenez and Sawada 2003; King and Ozler 1998; and Sawada and Ragatz 2005).

The general finding that SBM had had positive results on some variables—mainly, repetition, failure, and attendance rates—in contrast with the mixed results in test scores could be due to the timing and strength of the particular SBM reforms. Research in the United States suggests that an SBM reform has to have been in operation for about five years before any fundamental changes are seen at the school level, and only after eight years of implementation can changes be seen in more difficult to modify indicators such as test scores. Moreover, it is possible to argue that school learning is a cumulative process and that students need to have been exposed to the reform for a longer period of time to enjoy its potential benefits.

Three studies (Paes de Barros and Mendonca 1998; Lopez-Calva and Espinosa 2006; and Parker 2005) allowed more than eight years before measuring the effects of the intervention on test scores. Paes de Barros and Mendonca found that the reform in Brazil had produced no test score improvements after 11 years of implementation, but the other two studies showed that the reforms in Nicaragua and Mexico had positive effects on test scores after 11 and 8 years, respectively. Other studies measured SBM's impact on repetition and failure rates (intermediate indicators) closer to the initial implementation period and the authors of these studies found positive effects after only two years of implementation in the case of rural Mexico (Gertler et al. 2006) and after only two to three years in urban Mexico (Skoufias and Shapiro 2006).

The lack of cost-benefit analyses of SBM is also an important gap in the literature. Clearly, SBM is a very inexpensive initiative since it constitutes a change in the locus of decision-making and not necessarily the amount of resources in the system. If the few positive impact evaluations are true, then SBM can be regarded as a very cost-effective initiative.

## Guidance on How to Implement Impact Evaluations

Based on our review of the articles on SBM impact, it is clear that retrospective evaluations (or evaluations based on programs that are already implemented and have limited data) are extremely difficult to perform, especially when they use IV methods. It is preferable to carry out prospective evaluations on programs that have yet to be implemented so that baseline (preintervention) data can be collected in advance. Box 2 presents some cases where this strategy is being adopted.

There are three main ways to identify the causal effects of SBM programs. First, there are strategies in which a randomization of treatment is implemented, second, there are strategies in which the entry order into the program is randomized, and third, there are strategies that encourage participation.

Randomization at the school level is very difficult to observe in reality, so it is better to use some geographical criterion. Even if randomization at the geographical level is possible, reallocating students between schools will result in problems of selection. For this reason, it is critical to collect information on students who switch schools and to analyze differences in the characteristics of those students who stay in one type of a school and those who decide to attend a different type of a school.

When randomizing is performed at some higher geographical level than the school level, it is important to have detailed baseline information. For example, using a randomization when the units of observation are states can result in imbalances in the treatment and control groups because of the potential low number of observations and these states may have very distinct characteristics. Baseline data can indicate whether there are any differences (in observable characteristics) between the treatment and control groups, and then analysts can control for them in the estimation.

If pure randomization is not possible, then a strategy that randomizes entry time may be feasible. In this case, the order in which SBM is implemented in localities can be chosen by lottery. A simple example is the case in which the program is implemented first in one group of communities and then later in another group. The group that enters later is the control group for the initial participants. Ideally, the information would be collected at least three times— before the intervention, before the intervention in the second group, and at some point in time after both groups have received the intervention. The last data collection point makes it possible to detect the intensity of the effects and the speed of the impact. Indeed, observing differences between the two groups allows analysts to make inferences about the speed of the program's effects, since the first group will have been exposed to the program for longer than the group that entered later.

The last randomization strategy is to use an encouragement model. In short, active campaigns can be introduced to encourage a group of communities, chosen randomly, to participate in the program. These campaigns can include visits by program promoters, NGO representatives, or social workers to explain the program and describe the potential benefits of the intervention. The rest of the communities will have access to general information about the program, but their participation will not be actively solicited. In this case, the promotion campaign is used as an instrumental variable of participation. Since the campaign is not correlated with the educational outcomes of the school in the community but is eventually correlated with participation into the program, the instrument is a valid one. Duflo and Saez (2003) and Hirano et al. (2000) are examples of studies that have used this strategy.

In short, the ideal evaluation will use some form of randomization. However, if randomization is not an option, RDD and DD strategies are an alternative. First, an RDD procedure is suitable when the program is targeted using some continuous variable as the entry criterion. The estimation will then discover the true effect of the intervention without the need for randomization in the design of the program. This fact makes RDD a more flexible procedure, especially for evaluating programs that are already in place.

The second promising nonrandomized strategy uses a nonrandom phase-in approach. It is possible to use this source of variation to evaluate the effects of an SBM program. For example, Gertler et al. (2007) used this strategy. As stressed before, for this evaluation method to be technically sound, it is critical to ensure that the later treatment group has similar pretreatment observable characteristics as the group that initially enters the program. This necessitates the existence of good preintervention data as well as good postintervention data.

## Conclusions

In conclusion, the number of rigorous studies of the impact of SBM is very limited. We found only 13 studies that had made a clear attempt to correct problems of endogeneity. However, some of them have serious limitations. The lack of randomized experiments has led some researchers to carry out retrospective analyses, which are open to criticism about the validity of the instruments used. The studies that used differences between beneficiary and nonbeneficiary groups over time or that tried to match beneficiaries to a similar nonbeneficiary group were limited by a lack of data, either because the baseline data were not rich enough or because preprogram trend data did not exist.

Among the 13 most rigorous studies, some found that SBM policies actually changed the dynamics of the school, either because parents got more involved or because teachers' actions changed. Several studies presented evidence that SBM had had a positive impact on repetition rates, failure rates, and, to a lesser degree, dropout rates, but those studies that had access to standardized test scores presented mixed evidence about the impact of SBM on those scores.

## Annex 1. A Selection of Evaluated School-Based Management Programs

| Year of program | Country | Program description | Selection of schools / communities | Scope |
|---|---|---|---|---|
| 1982 | Brazil | Decentralization: direct transfer of funds to schools, election of principals, and creation of local school councils | Phased in | All schools |
| 1991 | El Salvador | EDUCO: Community associations are responsible for administering funds, hiring / firing teachers, and monitoring and maintaining infrastructure | Municipalities and national level, with the help of promoters, identified communities | Not all schools in the country participate |
| 1991 and 1993 | Nicaragua | ASP: In 1991, establishment of consultative councils; in 1993, transformed into management boards; wide scope of autonomous decisions | Teachers vote to enter the program | Not all schools in the country participate |
| 1996 | Mexico | AGES (Support for School Management): gives the parents' associations small amounts of money for civil works and infrastructure | National government target areas; phase-in program: first indigenous populations, lagging primary schools, disadvantaged rural areas | Targets schools in rural areas |
| 1999 | Honduras | PROHECO: School councils have autonomy over hiring and firing teachers, monitoring, managing funds, and maintaining infrastructure | National government targets rural schools affected by Hurricane Mitch; social promoters approach communities to raise awareness and help in the process | Not all schools in the country participate |
| 2001 | Mexico | PEC (Quality School Program): gives schools resources for implementing a school plan, in consultation with parents; part of the money goes to infrastructure and part to teacher quality | National government targets areas; voluntary, disadvantaged urban areas | Priority to disadvantaged rural areas |

*Sources:* di Gropello (2006), Paes de Barros and Mendonca (1998), and Gertler et al. (2006).

## References

Angrist, J. D., and G.W. Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association* 90(430): 431–42.

Athey, S., and G. W. Imbens. 2006. "Identification and Inference in Nonlinear Difference-In-Difference Models." *Econometrica* 74(2): 431–97.

Bauer, S. 1996. "Site-Based Management: A Design Perspective." Unpublished doctoral dissertation, Cornell University.

Bauer, S. 1997. "Designing Site-Based Systems, Deriving a Theory of Practice." *International Journal of Educational Reform* 7(2): 108–121.

Bauer, S. C., I. E. Bogotch, and H. S. Park. 1998. "Modeling Site-Based Decision-making: The Relationship between Inputs, Site Council Practices, and Outcomes." Paper presented at the Annual Meeting of the American Educational Research Association. San Diego, CA, April.

Borman, G. D., G. M. Hewes, L. T. Overman, and S. Brown. 2003. "Comprehensive School Reform and Achievement: A Meta-Analysis." *Review of Educational Research* 73(2): 125–230.

Bryk, A.S., Y. M. Thum, J. Q. Easton, and S. Luppescu. 1998. "Academic Productivity of Chicago Public Elementary Schools." A Technical Report Sponsored by The Consortium on Chicago School Research, March.

Carnoy, M., A. Gove, S. Loeb, J. Marshall, and M. Socias. 2004. "How Schools and Students Respond to School Improvement Programs: The Case of Brazil's PDE." Forthcoming in *Economics of Education Review.*

Cook, T. D. (2007). "School-Based Management in the United States." Background paper prepared for the SBM study for the World Bank.

di Gropello, E. 2006. "A Comparative Analysis of School-Based Management in Central America." World Bank Working Paper No. 72, World Bank, Washington D.C.

di Gropello, E., and J. H. Marshall. 2005. "Teacher Effort and Schooling Outcomes in Rural Honduras." In *Incentives to Improve Teaching*, Emiliana Vega. World Bank, Washington, DC.

Duflo, E., and E. Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics* 118: 815–42.

Duflo E., R. Glennerster, and M. Kremer. 2006. "Using Randomization in Development Economics Research: A Toolkit." NBER Technical Working Paper No. 333.

Gertler, P., H. Patrinos, and M. Rubio-Codina. 2007. "Methodological Issues in the Evaluation School-Based Management Reforms." Draft, World Bank, Washington, DC.

Gertler, P., H. Patrinos, and M. Rubio-Codina. 2006. "Empowering Parents to Improve Education: Evidence from Rural Mexico." World Bank Policy Research Working Paper 3935. Washington D.C.

Gunnarsson, V., P. F. Orazem, M. Sánchez, and A. Verdisco. 2004. "Does School Decentralization Raise Student Outcomes? Theory and Evidence on the Roles of School Autonomy and Community Participation." Iowa State University Working Paper.

Heckman, J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Economic and Social Measuremen*t 5(4): 475–92.

Heckman, J., H. Ichimura, and P. Todd. 1998. "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies* 65: 261–94.

Hess, A. 1999. "Understanding Achievement (and Other) Changes under Chicago School Reform." *Educational Evaluation and Policy Analysis* 21(1): 67–83.

Hirano K., G. W. Imbens, D. B. Rubin, and X. Zhou. 2000. "Assessing the Effect of an Influenza Vaccine in an Encouragement Design." *Biostatistics* 1: 69–88.

Jimenez, E., and Y. Sawada. 1999. "Do Community-managed Schools Work? An Evaluation of El Salvador's EDUCO Program." *World Bank Economic Review* 13(3): 415–41.

Jimenez, E., and Y. Sawada. 2003. "Does Community Management Help Keep Kids in Schools? Evidence using Panel Data from El Salvador's EDUCO Program." Discussion Paper CIRJE-F-236., World Bank, Washington, DC.

King, E., and B. Ozler. 1998. "What's Decentralization Got to Do with Learning? The Case of Nicaragua's School Autonomy Reform." Working Paper on Impact Evaluation of Education Reforms, World Bank, Washington, DC.

King, E., P. F. Orazem, and V. Gunnarsson. 2003. "Decentralization and Student Achievement: International Evidence on the Roles of School Autonomy and Community Participation." Paper presented to the Fourth Annual Global Development Conference on Globalization and Equity, Cairo, Egypt, January 19–21.

Lopez-Calva, L. F., and L. D. Espinosa. 2006. "Efectos Diferenciales de los Programas Compensatorios del CONAFE en el Aprovechamiento Escolar." In *Efectos del Impulso a la Participación de los Padres de Familia en la Escuela.* CONAFE, México.

Murnane, R. J., J. B. Willet, and S. Cardenas. 2006. "Did Participation of Schools in Programa Escuelas de Calidad (PEC) Influence Student Outcomes?" Working Paper, Harvard University Graduate School of Education, Cambridge, MA.

OECD (Organisation for Economic Co-operation and Development). 2004. "Raising the Quality of Educational Performance at School." Policy Brief. Paris: OECD (www.oecd.org/dataoecd/17/8/29472036.pdf).

Paes de Barros, R., and R. Mendonca. 1998. "The Impact of Three Institutional Innovations in Brazilian Education." In *Organization Matters: Agency Problems in Health and Education in Latin America,* W.D. Savedoff. Inter-American Development Bank, Washington, DC.

Parker, C.E. 2005. "Teacher Incentives and Student Achievement in Nicaraguan Autonomous Schools." In *Incentives to Improve Teaching,* E. Vegas. World Bank, Washington, D.C.

Rosenbaum, P. R., and D. B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1): 41–55.

Santibañez, L. 2006. "School-Based Management Effects on Educational Outcomes: A Literature Review and Assessment of the Evidence Base." Working Paper, World Bank, Washington, DC.

Sawada, Y., and A. B. Ragatz. 2005. "Decentralization of Education, Teacher Behavior and Outcomes." In *Incentives to Improve Teaching,* E. Vegas. World Bank, Washington, DC.

Shadish, W., T. D. Cook, and D. T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston: Houghton Mifflin Company.

Shapiro, J. S., and E. Skoufias. 2006. "Evaluating the Impact of Mexico's Quality Schools

Program: The Pitfalls of Using Nonexperimental Data." World Bank Policy Research Working Paper 4036. Washington, DC.

Skoufias, E., and J. Shapiro. 2006. "The Pitfalls of Evaluating a School Grants Program Using Nonexperimental Data." World Bank Policy

Research Working Paper No. 4036. Washington, DC.

World Bank. 2007. *What Is School-Based Management?* World Bank, Washington, DC.

**THE WORLD BANK**
1818 H Street, NW
Washington, DC 20433 USA
Telephone: 202-473-1000
Facsimile: 202-477-6391
Internet: www.worldbank.org/education
E-mail: eservice@worldbank.org