Public Disclosure Authorized

Public Disclosure Authorized

WPS5248

Policy Research Working Paper 5248

The Effects of School-based Management in the Philippines

An Initial Assessment Using Administrative Data

Nidhi Khattri
Cristina Ling
Shreyasi Jha

The World Bank
Independent Evaluation Group
East Asia Education Sector Unit
&
World Bank Institute
March 2010

Abstract

This paper estimates the effect of school-based management on student performance in the Philippines using the administrative dataset of all public schools in 23 school districts over a 3-year period, 2003–2005. The authors test whether schools that received early school-based management interventions (training in school-based management and direct funding for school-based reforms) attained higher average test scores than those that did not receive such inputs. The analysis uses school-level overall composite test scores (comprising all subject areas tested) and test scores in three separate subject areas: English, math, and science. Their preferred estimator, difference-in-difference with propensity score matching, shows that the average treatment effect of participation in school-based management was higher by 1.5 percentage points for overall composite scores, 1.2 percentage points for math scores, 1.4 percentage points for English scores, and 1.8 percentage points for science scores. These results suggest that the introduction of school-based management had a statistically significant, albeit small, overall positive effect on average school-level test scores in 23 school districts in the Philippines. The paper provides a first glimpse of the potential for school-based management in an East Asian context based on available administrative data. The authors suggest that the next order of research is to answer policy-related questions regarding the reforms: what aspects of the reform lead to desired results; are there differential effects across subpopulations; and what are the potential downsides to the reforms? The Philippines is embarking on a nation-wide implementation of school-based management and the authors recommend that mechanisms for rigorous evaluations be advanced simultaneously. Such evaluations should not only provide more accurate estimates of the effectiveness of the reforms, but also help answer policy-related questions regarding design and implementation of those reforms in different socio-cultural contexts.

This paper—a joint product of East Asia Education Sector Unit, Independent Evaluation Group; and the World Bank Institute—of a broader Bank-wide effort to promote rigorous and analytical approaches to monitoring and evaluation. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at nkhattri@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.

Produced by the Research Support Team

The Effects of School-based Management in the Philippines: An Initial Assessment Using Administrative Data

Nidhi Khattri[1]

Cristina Ling

Shreyasi Jha

Keywords: school-based management, decentralization, evaluation, propensity score matching, administrative data

[1] Nidhi Khattri, Senior Evaluation Officer, Independent Evaluation Group, The World Bank, Washington D.C. email: nkhattri@worldbank.org; Cristina Ling, Evaluation Officer, World Bank Institute Evaluation Group, The World Bank, Washington D.C. email: cling@worldbank.org; Shreyasi Jha, Adjunct Professor, The Wilf Family Department of Politics, New York University, New York. Email: sj46@nyu.edu.

We would like to thank the Department of Education of the Philippines and Professor H. Nuqui, University of the Philippines, for providing the data and information on SBM. The authors also gratefully acknowledge the support of Dingyong Hou (Senior Operations Officer, EAPHD, and Task Team Leader for the Third Elementary Education Project), Lynette Dela Cruz Perez (Senior Operations Officer, EASHE), Christopher Thomas (Manager, EAPHD), and Eduardo Velez Bustillo (EASHE). This work was partially supported by the Education Program Development Fund. We are indebted our reviewers for providing helpful comments and suggestions on previous drafts. All remaining errors are ours alone. The opinions and conclusions expressed in this paper do not necessarily reflect those of the Executive Directors and officials of the World Bank or of its member governments.

1. Introduction

Decentralization is a key feature of institutional reform throughout the world.[2] The main argument underpinning decentralization policies is that they empower people to be part of the local decision-making process –they improve government performance by alleviating information asymmetries and costs by bringing decision-making closer to the people concerned. However, decentralization can also worsen the provision of public goods in the presence of externalities, lack of technical capabilities by local governments, or capture of lower-level administration by local elites.[3] In the context of the education sector, decentralization typically includes one or more of the following features: decentralized revenue generation, curriculum design, school administration, and teacher hiring and management. Decision-making authority for these types of functions is devolved to regional/municipal governments or to schools themselves. The policy of allowing schools autonomy in decisions in these areas is referred to as school-based management (SBM), school based governance, or school self management[4],[5]. Responsibility and decision-making over different types of school operations are transferred to individuals at the school level, who in turn must conform within a set of centrally or state-level determined policies.

The popularity of SBM is evidenced by the large number of development agencies promoting it as a key component of the decentralization reforms and the growing number of countries that have adopted aspects of this approach. SBM reforms began in the 1970s in Australia. Since then, a wide range of countries have experimented with or introduced SBM in all regions of the world, including Hong Kong (China), Indonesia, El Salvador, Nicaragua, Kenya, Kyrgyz Republic, Nepal, Paraguay and Mexico. [6]

Nevertheless, the impact of SBM on education quality, including student outcomes, remains a contentious issue, with some researchers arguing that SBM leads to enhanced educational outcomes,[7] while others contending that SBM leads to the deterioration of educational quality especially among the weakest schools.[8] The range of SBM approaches and the contexts in which they are implemented makes the debate about SBM quality an intricate one. The evaluation of SBM is complicated by the diversity of approaches to and elements of decentralization that collectively constitute "SBM" and by the institutional and socio-cultural contexts in which they are implemented. Nonetheless, some studies in recent years have found that SBM reforms are associated with improved education outcomes and processes (e.g.,

[2] King and Ozler, 1998.
[3] Galiani and Schargrodsky, 2001
[4] De Grauwe, 2004
[5] Gertler, Patrinos, and Codina, 2006
[6] See "What is SBM?" World Bank (2007)
[7] e.g., Gertler, Patrinos, and Codina, 2006
[8] Bardhan, 2002.

Skoufias & Shapiro, 2006; Sawada & Ragatz, 2005; Gunnarsson et al., 2004; Eskeland & Filmer, 2002). However, rigorous evidence base for the effectiveness of SBM in boosting student performance is thin. A recent review of the empirical literature on SBM since 1995 indicates that only 14 studies utilized rigorous methods to assess the impact of SBM, and only six reported positive impacts on students' test scores (Barrera-Osorio et al., 2009) . Eleven studies are country-specific from Latin America, one from Kenya, and two exploit data from multiple countries. No empirical evidence is available from East Asia.

This paper contributes to the small but growing empirical literature on the effectiveness of SBM by extending the research to East Asia. The paper provides an initial analysis of the potential of school-based management for improving educational outcomes in the Philippines, using aggregated school-level test scores and administrative data. The data are quite limited and do not allow for a thorough analysis of the processes and approaches through which SBM reforms affect outcomes. Nonetheless, the paper attempts to use existing data to answer an initial question: did the introduction of SBM lead, on average, to enhanced educational outcomes? It also demonstrates how administrative data can be mined for exploratory assessments of potentially larger programs.

The remainder of the paper is organized as follows: the next section provides a brief overview of the SBM program in the Philippines; section 3 outlines the methodological approaches used for the analysis; section 4 describes the data; section 5 discusses the empirical results; and section 6 provides a summary of our conclusions, discusses the limitations of the study, and highlights the implications for future, more rigorous SBM evaluations in the Philippines.

2. SBM Program in the Philippines

SBM was implemented in between 2003 and 2005 in 23 districts participating in the Third Elementary Education Project (TEEP) supported by the World Bank.[9] The project provided funding for school infrastructure, training, curriculum development, and textbooks. SBM was introduced as an integrating framework for obtaining school-level project inputs and building school capacity for education planning and program implementation beginning in school year 2003-04. Schools participating in SBM were required to design a five-year School Improvement Plan (SIP) in partnership with parents and the community using data such as student achievement and students' learning needs assessments, with the school principal or head teacher leading the process. Based on the SIP, schools developed an Annual

[9] Although some other school districts in the Philippines had ongoing SBM initiatives, they are not included in this analysis as insufficient information is available about the nature and timing of those initiatives.

Implementation Plan (AIP) at the beginning of the school year and a report card to be shared with the community at the end of the school year. Project inputs for infrastructure, training, textbooks, and so forth, were partially based on the SIP. Principals and head teachers received training in leading the development and implementation of the SIP and the AIPs in collaboration with teachers and key members of the larger community. SBM schools also received funds for maintenance and operating expenses directly in cash rather than in kind, as had been the case previously. These cash funds could be used by the schools based on their AIP, The cash allocation was based on a formula that provided each school with a flat amount of funds plus a prorated figure based on the number of student and teachers as well as other criteria, such as percentage of indigenous student population in the school. Schools not participating in the SBM received no SBM-related training and no cash funds, and they were not required to develop SIPs and AIPs.

The SBM was designed to improve student outcomes through two main venues: by empowering the school community to identify education priorities and to allocate the school maintenance and operating budgets to those priorities (such as curriculum enrichment programs); and by enhancing transparency and accountability through the annual implementation plans and school report cards. However, the SBM program articulated no explicit assumptions regarding the timeframe within which improvements in student achievement were expected to take place. Systematic data on the level of uptake and implementation of the key features of the reforms are also not available.

The SBM training, funds, and requirements were rolled out in three batches and eventually covered almost all (84 percent) of the 8,613 schools in the 23 project districts. The first batch comprised 1,666 schools in 2003-2004, largely because they were perceived to be more capable, although no explicit assignment mechanism was designed.[10] The next batch of 2,700 schools was targeted for SBM rollout in 2004-2005, and another batch of 1,529 was included in 2005-2006.

3. Analytic Approach

We assess the effects of SBM on student performance using average school-level test scores from all schools and school-level indicators between 2003 and 2005 from the 23 TEEP districts in the Philippines. We base our analysis on a pipeline comparison strategy. Since we have student achievement data over a three year period, but schools were inducted into the program in three batches over time, we identify the

[10] No written materials were available on the identification strategy. Project officials interviewed as part of this paper mentioned that schools were chosen on the basis of the perceived strength of their capabilities.

treatment group as the *first* batch of schools that received SBM funds in 2003-2004. The treatment group consists of the schools that did not receive SBM funds and training in 2003-2004. Therefore, for this study, we operationalized SBM batch one schools as the *treatment group* (n=1666) and all other schools (batches that received the SBM intervention later) as the *control group* (n=3501).[11] In our analysis, the treatment group (batch one) therefore had exposure to SBM for a period of two years, 2003-04 and 2004-05.

Because selection into SBM was not voluntary, there is likely to be placement bias in our control and treatment groups. We therefore use two non-experimental evaluation techniques to estimate program impact: difference-in-difference estimators and difference-in-difference with propensity score matching estimators. Using nonparametric kernel matching techniques, we correct for potential sources of bias.

Specifically, we tested whether the composite test scores and test scores individually in three subject areas (math, science and English) were higher after SBM was introduced for schools in the treatment group compared to those in the control group.

4. Data and Statistical Specification

We use school administrative data from all 5,167 schools in the 23 TEEP districts to examine the effect of SBM program on student test scores. The dataset includes information on school personnel, students, facilities, and enrollment and dropouts for all schools in the district over a 3-year period, 2002-2003 to 2004-2005. The data were collected by the Philippines Department of Education as part of the management of the Third Elementary Education Project (TEEP) of which SBM was one component.

Student achievement is measured using National Achievement Tests conducted in 2002-03 (year one – pre-intervention) in Grade 4, 2003-04 (year two) in Grade 5, and 2004-05 (year three, post-intervention) in Grade 6 in English, mathematics, science. However, only the school-level Mean Percentile Score (MPS) data are available in the dataset, limiting the analysis to the school level. We utilize the composite MPS computation (based on tests in all subjects), as well as the English, mathematics, and science MPS for 2002-03 (pre-intervention) and 2004-05 (post-intervention).

[11] The number of schools in the analysis is not the same as the number of schools included in the SBM program due to missing information. The difference is 728 schools.

Our method of analysis consists of two quasi-experimental evaluation techniques. We have data on test scores over three years for all schools. We also have time-invariant data on school characteristics. Once we determined the control and treatment groups as described in the preceding section, the preferred estimator to use in this context was the *difference-in-differences (DID)* estimator. However, in order for the DID estimator to provide credible estimates of the program impact, it should be free of any biases inherited by pre-existing (pre-intervention) differences between treatment and control groups. Because of problems incurred with using the DID estimator due to way in which SBM program was rolled out (i.e., no clear assignment mechanism), we also estimate and discuss DID in conjunction with *propensity score matching (PSM)* to estimate average treatment effects.

(i) Difference-in-differences (DID) estimation

The ideal experiment for identifying the effects of SBM on student test scores would have been to randomly assign schools to control and treatment groups. The SBM effect could then have been estimated by comparing schools that received SBM with their peers in the comparison group. Unfortunately, such a design is not feasible in this case because of the nature of the SBM program rollout. Therefore, we assigned the set of schools that received early intervention in 2003 as the treatment group. The control group comprised all other schools that did not receive SBM support in 2003 – i.e., schools that received the intervention later. The SBM effect on student test scores can be evaluated as follows:

SBM effect = $(TestScores_{SBM.POST} - Testscores_{SBM,PRE}) - (Testscores_{NONSBM,POST} - TestScores_{NONSBM.PRE})$

--------------------(eq. 1)

Converting equation (1) above to an estimating model, we get:

$$Y_{s,t} = \beta_0 + \beta_1 Year_t + \beta_2 SBM_s + \beta_3 (Year_t * SBM_s) + \beta_4 Z_{ts} + \mu_p + \varepsilon_{ts}$$

-------------------(eq. 2)

where Y is the outcome measure. We include school-level composite test scores as well as test scores in math, science, and English as outcome measures. Year is a dichotomous variable that equals 1 for post-intervention and equals 0 for pre-intervention. SBM is a dichotomous variable that is equal to 1 indicating that a school received SBM support during all of the treatment years (*t* =2002-2004). Z is a vector of school characteristics. μ_p is division-specific fixed effects intended to capture division-specific aggregate fixed effects correlated with schooling outcomes (demographic trends or changes in government, for example). $\varepsilon_{t,s}$ is the school level error term that includes all the unobserved school characteristics that we assumed are uncorrelated with the explanatory variables for the time being.[12]

The vector of school characteristics includes total enrollment, total dropout, student-teacher ratio, textbooks per student, teacher manual per teacher, school type, school head type, whether the school was an elementary leader school (a school considered to be a leader among others in the district) and whether the school had received SIIF funds.[13] School type consists of four categories: complete combination and multi-grade schools; complete large and medium school; complete small schools; and primary schools. School head type consisted of three categories: head principal, head teacher, and teacher in charge. Table 1 summarizes the descriptive statistics for all schools and by treatment status of the school.

β_0 is the intercept showing the average test scores of the SBM and non-SBM schools; β_1 shows the change between 2002-03 and 2004-05; β_2 is the difference in the treatment and control group; and β_3 is the DID estimator capturing the differences between the control and treatment groups, before and after the intervention. More specifically, it measures the SBM effect as the change in average school-level test scores among two groups before and after the 2003 SBM intervention. These groups included: the "first batch," or the schools that received SBM first (treatment group) to schools that received the intervention later and those that did not receive the intervention at all (comparison group).

In order to obtain the DID estimates, we created a panel dataset from our original dataset of 5,167 schools. Since we had test scores for all three years, we simply replicated this data for every year to obtain a panel dataset with 15,501 observations. In this dataset, each school is observed in each of the three years (2002-03 to 2004-05).

In order to use the DID estimates in this context, we need to be certain that the difference in post-intervention outcomes between SBM and non-SBM schools would have been identical in the absence of the intervention. However, this assumption is impossible to test because we do not actually observe the counterfactual. We can nonetheless test whether pre-intervention and "mid-term" educational outcomes

[12] The intervention might alter the number of children enrolling in a particular school. If as a consequence the distribution of students' skills changes in treatment schools (with respect to control schools), then the program impact estimates are likely to be biased. We will explore the existence of this bias in section five when discuss the caveats. The characteristics of the average student in the school are also included in the error term because of lack of data on individual students' characteristics.

[13] School Improvement and Innovation Facility that granted funds to schools based on proposals for school improvement programs.

under study were similar between the treatment group and the proposed control group. If the pre-intervention and “mid-term” (past baseline and probably without implemented intervention) outcome measures were not significantly different between treatment and control schools, there is no compelling reason to believe they would be significantly different in the post-intervention periods had the SBM program not been put in place. However, if we observe changes between pre-intervention and mid-term, we can consider the possibility that the two groups would have diverged with respect to their outcomes even in the absence of the SBM intervention. We can test this assumption for the proposed treatment and control groups by running the following equation on the pre-SBM data (i.e. 2002-2003):

$$Y_{st} = \beta_0 + \beta_1 (Year_{2002=0/2003=1}) + \beta_2 (Year_{2002=0/2003=1} * SBM_{1/0}) + \mu_p + \varepsilon_{ts}$$

-------------------(eq. 3)

Y is the outcome measure including: school level composite test scores and test scores individually in math, science, and English. *Year* is a dummy variable that takes value = 1, if year =2003-04 and =0 if year=2002-03 (both pre-intervention and implementation years). SBM is a dichotomous variable that is equal to 1 if the school *s* received SBM support (*SBM*=1). As in eq. (2), μ_p is intended to capture division-specific aggregate fixed effects and $\varepsilon_{t,s}$ is the error term. We are interested in two issues: first, was there a pre-existing difference between SBM and non-SBM schools before the intervention; and two, did the difference increase over time (between 2002-03 and 2003-04). The coefficient for the interaction term can be interpreted as follows: if the coefficient β_2=0, then the pre-SBM outcomes for schools that would eventually receive SBM funds are not significantly different from those in the control group (i.e., schools that did not receive SBM funds in 2003).

We run equation (3) separately for years 2002-03 and 2003-04. When year=2002-03 (pre-intervention), we do not have a coefficient for β_1. Table 2 reports the results from equation (3). For 2002, we report coefficient β_2 which shows the difference between SBM and non-SBM schools at the baseline (2002-03). For 2003-04 (mid-term), we report β_2 which shows the increase in test scores of SBM schools compared to non-SBM schools in 2003-04 compared to the baseline (2002-03). For both years Model A is a simple OLS and Model B includes division fixed effects. For overall test scores, Model B in Table 2 shows that β_2 =2.0, implying that overall test scores for schools that would eventually receive SBM funds in 2003 were about 2 percentage points higher than non-SBM schools in 2002-03. In 2003-04 (mid-term), the difference in overall test scores between SBM and non-SBM schools was about 3.0, implying that overall test scores for SBM schools 3 percentage points higher than non-SBM schools in 2003 compared to 2002.

Table 2 shows there are significant differences in the pre-intervention period in test scores (math, science, English and overall) across schools that would eventually receive SBM and those that did not receive SBM in 2003. Based on these results, it appears that program placement bias is a serious concern in this context, and that difference-in-difference estimates may not give us unbiased estimates of the program effect. Therefore, in addition to difference-in-difference estimates, we also obtain propensity score matching estimates as a means of achieving unbiased identification of the SBM effect.

(ii) Propensity score matching (PSM) estimation

One of the potentially most serious problems in any evaluation study is the occurrence of placement bias. In this case, the problem arises because we would like to know students' test scores with and without the SBM intervention. The results in Table 2 indicate that placement bias is a serious concern in this case because SBM and non-SBM schools are substantially different even in the absence of the treatment. The matching approach is one possible solution to the selection problem. This method originated from the statistical literature and shows a close link to the experimental context[14]. The basic idea is to identify within a large group of non-participants those individuals who are similar to the participants in all relevant pre-treatment characteristics X. The assumption is that once this is done, differences in outcomes between the matched control group and the treatment participants can be attributed to the program.

Since conditioning on all relevant covariates is limited in case of a high dimensional vector X (`curse of dimensionality'), Rosenbaum and Rubin (1983) suggest the use of so-called balancing scores $b(X)$, i.e., functions of the relevant observed covariates X such that the conditional distribution of X given $b(X)$ is independent of assignment into treatment. One possible balancing score is the propensity score, i.e., the probability of participating in a program given observed characteristics X. Matching procedures based on this balancing score are known as propensity score matching (PSM) and we will use PSM matching in conjunction with DID as the preferred estimator in this paper. PSM proposes that all pre-treatment characteristics of a unit of observation is summarized into a single index (the propensity scores) and the units of observations are then matched on their propensity scores.

According to Skoufias and Shapiro (2006), the propensity score is defined as the conditional probability of receiving a treatment (receiving SBM) given pre-treatment characteristics *X, i.e.*

[14] Propensity Score Matching (PSM) was proposed by Rosenbaum and Rubin (1983).

$$p(X) \equiv \Pr(D = 1 | X) = E(D | X)$$

------------------------------(eq. 4)

where D is the indicator for receiving treatment (=1 if SBM, =0 if non-SBM). As shown by Rosenbaum and Rubin (1983), if the propensity score $p(X)$ is known, then the PSM estimator for ATT (`average treatment effect on the treated') can be estimated as follows:

$$T_{ATT}^{PSM} = E_{P(X/D=1)}\{E[Y_i / D = 1, p(X)] - E[Y_0 / D = 0, p(X)]\}$$

------------------------------(eq.5)

The potential outcomes are defined as $Yi(Di)$ for each individual i, where $i = 1\ldots.N$ and N denotes the total population. To put eq. 5 in words, the PSM estimator is simply the mean difference in outcomes over the common support, appropriately weighted by the propensity score distribution of participants.

When estimating the propensity scores, two choices must be made. The first one concerns the model to be used for the estimation, and the second one requires identifying the variables to be included in the model. We use a logit model to estimate propensity scores. The variables in the model were selected to satisfy the *conditional independence assumption (CIA).* The matching strategy builds on the CIA, requiring that the outcome variable(s) must be independent of treatment, conditional on the propensity score. Hence, implementing matching requires choosing a set of variables X that credibly satisfy this condition. Only variables that influence simultaneously the participation decision and the outcome variable should be included. We included five explanatory variables in the logit model: student-teacher ratio; size of school (total enrollment in 2002 base year); school type (incomplete vs. complete schools & combination/multigrade schools vs. mono schools); type of school head (principal vs. teachers and teacher-in-charge).

The results from the logit model are presented in Table 3. We impose the common support area condition to ensure that any combination of characteristics observed in the treatment schools can also be observed among the control group.[15] We then apply the nonparametric kernel matching technique to match schools based on their propensity scores within school divisions. We matched schools at the division level because the SBM intervention took place within school divisions. Kernel matching uses weighted averages of all individual cases in the control group to construct the counterfactual outcome. Thus, one major advantage of these approaches is the lower variance which is achieved because more information is used.

[15] Given that common support area condition is met, PSM approach eliminates a substantial amount of selection bias that would alter conclusions in interpretation (Heckman, Hidehiko, Smith and Todd, 1996).

Since we do not condition on all covariates but on the propensity score, it has to be checked if the matching procedure is able to balance the distribution of the relevant variables in both the control and treatment group. Several procedures have been suggested to check the matching quality. The basic idea behind these approaches is to compare the situation before and after matching, and verify whether any differences remain after conditioning on the propensity score. If differences exist, matching on the score was not (completely) successful and remedial measures must be taken. We use a stratification test proposed by Dehejia and Wahba (2002) that divides observations into strata based on the estimated propensity score, such that no statistically significant difference remains between the estimated propensity score means of the treatment and control groups. We then use t-tests within each stratum to test if the distribution of *X*-variables is the same between both groups. Table 4 presents the absolute t-scores for the X-variables across SBM and non-SBM school by quintiles. We notice that across variables there are no significant differences between SBM and non-SBM schools on most variables except total enrollment.[16,17]

Finally, given the availability of data over three years we use the differences-in-difference estimator (developed by Heckman et al. in 1997, 1998a,b) to calculate the average treatment effect *(ATE).* This estimator compares the before-after test scores of SBM schools with the corresponding before and after changes among non-SBM schools, conditional on covariates X. This builds on the simple differences-in-difference statistic by controlling for covariates X and estimating the differences using nonparametric methods. The advantages of this estimator are similar to the advantages of the DID regression – it eliminates any unobserved factors that vary between observations, although not over time. This matching estimator allows selection on unobservables as long as the unobservable factors do not change between observations over time. Given the aforementioned advantages of the simple regression-based DID estimator, the DID with matching is our preferred estimator of program impact.

5. Results

Table 5 presents the simple means of school level test scores across SBM and non-SBM schools, before and after the intervention in 2003, without controlling for other explanatory variables. These figures do not control for any selection bias. Given that there is a strong possibility of selection bias, these estimates are helpful in showing trends and observing overall differences. However, we cannot make any claims of

---

[16] Some of the cells do not show t-scores because there were no observations for that particular variable in that cell.
[17] We tried several combinations of variables, but due to a limited dataset and number of available variables, kernel matching using the variables noted above provided the closest matches.

causality based on these comparisons. We examined test scores along four categories: math, English, science and overall (composite of all test scores).

Comparing SBM schools over time, we find that SBM schools scored higher in 2004 (year 3) in all areas including: math by 16.57 percentage points, science by 11.74 percentage points, English scores by 18.57 percentage points and in overall scores by 16.16 percentage points. We find similar increases in test scores for non-SBM schools, although the increases are not as large as they are for SBM schools. Math scores increased by 15.46 percentage points, science scores increased by 10.37 percentage points, English scores increased by 16.66 percentage points and overall scores increased by 14.65 percentage points.

Cross-sectionally, we find that SBM and non-SBM schools performed relatively similarly in 2002 (pre-intervention) (see Table 5). In 2005, the SBM schools performed better than did non-SBM schools. For example, math scores between SBM and non-SBM schools increased to 1.58 percentage points, science scores were 2.03 percentage points higher, English scores were 2.20 percentage points higher and overall scores are 1.98 percentage points higher for SBM schools compared to non-SBM schools.

A simple difference-in-difference comparison (without controlling for any factors) between SBM and non-SBM schools shows that test scores increased more rapidly among SBM schools. For instance, math scores increased by 1.12 percentage points for SBM schools, science scores increased by 1.36 percentage points, English scores by 1.91 percentage points and overall scores by 1.51 percentage points. Although these results point toward larger increase in test scores, we cannot make any causal inferences based on these estimates.

Table 6 presents the estimates of the average treatment effect derived from three multivariate analysis techniques. Since we are primarily interested in testing whether the introduction of SBM had an effect on test scores, we present only the coefficients associated with the variable measuring program impact. The detailed regression results are presented in the appendices.

The first column presents the estimates of the average treatment effects obtained from running an ordinary least squares (OLS) model of equation (2) without any division level fixed effects. We find that there was a significant increase in overall test score (1.15 percentage points), as well as an increase in test scores in specific subject areas: science (1.37 percentage points) and English (1.14 percentage points); the increase in math was not significant.

Next, we ran equation (2) including division level fixed effects and find more or less similar results. We find that there was a statistically significant increase in overall combined test scores (1.06 percentage points) as well as an increase in science (1.28 percentage points) and English scores (1.05 percentage points). As with the OLS model, the increase in math scores for SBM schools compared to non-SBM schools was not significant.

Finally, we present the results from DID estimators based on kernel matching. Generally, we find that the size of the impact was higher relative to DID without matching. Using kernel propensity score matching, we find that participating in the SBM program led to an increase of 1.45 percentage points in overall test score, 1.82 percentage points in science scores, and 1.32 percentage points in English scores. At 1.88 percentage points, the relative increase in math scores is also significant at the 10 percent level.

Summing up the findings, we can conclude that SBM in the Philippines had a statistically significant impact on overall student test scores and test scores in English and science. The effect sizes were small to moderate: .10 for overall scores, .13 for science, .09 for English, and .07 for math.[18] These are not inconsistent with Borman's (2002) analysis, which shows effect sizes of .17 to .14 after one and two years of implementation, respectively, and increasing with years of continued implementation. If we use the coefficients based on OLS and Fixed-Effects models, the effect sizes would be estimated to be lower.

6. Conclusions, Limitations, and Implications for the Philippines

This study of the SBM program in the Philippines shows that school-averaged student performance on national tests improved between 2002-03 and 2004-05 and that the level of improvement was higher for schools involved in SBM for two years compared with schools that had not yet received the intervention or received the intervention later. School-averaged student performance improved in math, science, and English and on the composite score. Improvement for schools that received SBM early was significantly higher in science and English and on composite test scores.

The Philippines has embarked on a nation-wide effort to introduce and implement SBM. While this paper provides an early indication of the usefulness of SBM in a few districts in the country and its promise for other parts of the country, it also highlights an opportunity for introducing a SBM program rollout design that would permit a more rigorous analysis of the contribution of SBM to student outcomes in the

[18] See Cohen (1988) for a discussion on the magnitude of effect size. Effect size is calculated by dividing the coefficient for interaction term, SBMxPostintervention, by the standard deviation of the 2002-03 baseline score.

Philippines. We discuss the limitations of the study and provide some recommendations for a program of SBM evaluation in the Philippines.

First although the study found significant differences between SBM and non-SBM schools on school-level outcomes, the possibility of unmeasured differences influencing outcomes exists. The findings in impact evaluations are critically dependent on the choice of the comparison group. In the context of this study, although a comparison group was identified, whether the group represented a good counterfactual is more difficult to argue. SBM funds were disbursed in batches and since we had data for all schools, we chose the first batch of SBM schools as the treatment school and created a counterfactual using available administrative data.

Although our model controls for a number of important school-level factors such as student-teacher ratio, school size, school type, and school head type, it is quite possible that other unobservable variables affected the outcomes. For example, the model does not capture student-level data (e.g., socio-economic background), which literature shows to have a significant relationship to student achievement. School personnel quality may similarly play a role. Furthermore, although we rely on the assumption that propensity-score matching would have accounted for some unmeasured differences, important differences cannot be ruled out.

While quasi-experimental evaluation methods are an improvement over simple multivariate techniques, they are not a substitute for more rigorous experimental designs. Taking an experimental approach where the treatment group is randomly selected would allow for a direct test of the effects of the intervention. For instance, if those who received SBM in the first batch were randomly selected with control schools that meet the same criteria, a simple differences-in-difference method would have sufficed. It is understandable that in real world conditions, however, an experimental design can be challenging to implement. While it is difficult to identify control groups that will never receive the program benefits, it is possible to implement a pipeline approach and roll out the program in batches, using randomization or a clear assignment mechanism over time, to enable "control groups" to be built in. This would provide program leaders with an opportunity by which to assess the effectiveness of the intervention among beneficiaries and non-beneficiaries who are similar in every other respect. Such a method would be superior to the one we have presented here, where we rely on statistical methods to create a control group for our analysis.

Second, the time frame captured in this study is very small, only two years of SBM preparation and implementation. In order to fully assess the impact of SBM, the assumptions regarding how the reforms play out over time and eventually affect student achievement will need to be articulated and examined explicitly. The possibility of a "hawthorne" effect cannot be eliminated, and, potentially, an evaluation with a longer time-frame would permit the identification of longer-term effects. Behrman and King (2008) highlight the risks of a poorly timed evaluation, ranging from finding partial or no impacts, when they in fact would take a longer time to materialize, to the risk of scaling up a poor program, by waiting too long to evaluate. Articulating clear assumptions regarding the reforms would assist with implementing evaluations in a timely fashion.

Third, the study does not examine the processes through which SBM affected school performance. Barrera-Osorio et al.'s (2009) review of the SBM literature conceptualizes the range of SBM approaches along two dimensions: *who* has the power to make decisions and the *degree* of decision-making devolved to the school level. The review notes that "With so many possible combinations of these two dimensions, almost every SBM reform is unique." (pg. 4). While some programs transfer authority to principals or teachers only, others encourage or mandate parental and community participation, often through school committees. Furthermore, the contexts in which the authority is devolved are likely to play a role in the eventual effects. For example, engaging parents and the wider community in school matters may not be easy in certain contexts – studies indicate that in communities with social and political tensions, the school board may be used as an instrument by the elite to orchestrate greater inequities.[19] Evidence from New Zealand and Australia demonstrates the under-representation of minority groups in the composition of school boards[20].

Several previous papers have examined some components of, and the processes through which, SBM affects student achievement. For example, Jimenez and Sawada (1998) examined 1996 data on Community Managed Schools (EDUCO), a program created to expand coverage in rural El Salvador, where communities received significant authority over schools, including in financial and staffing areas. They found that enhanced community and parental involvement in EDUCO schools improved students' language skills and diminished student absences, which may have long-term effects on achievement. Sawada and Ragatz (2005) also evaluated the EDUCO program and found that EDUCO transferred few administrative processes to local levels but gave local actors greater perceived influence in hiring and firing teachers. Gunnarsson et al.'s (2004) model of education decentralization suggests that principals,

[19] Caldwell, 2005.
[20] Ibid.

teachers, and parents are more informed than are national decision-makers in the areas that should be a priority for improving learning quality. Thus, by empowering local authorities to make spending decisions, resources will be allocated more strategically for better quality. Eskeland and Filmer (2002) similarly find that the autonomy of teachers, principals, and parents to make organizational and pedagogical decisions and the participation of parents in schools significantly increased primary school test scores in Argentina. Available reports on TEEP indicate that several aspects of SBM may have been conducive to improved student performance: improved school management through intensive principal and head-teacher training, identification of school-level needs and allocation of resources to those specific needs, greater community attention to schools' and students'concerns.

Fourth, the study does not examine the distribution of effects across different types of schools, both in terms of specified outcomes – student achievement – as well as unintended effects, such as toll on the principal's and teachers' time in community engagement. For example, Berlinski and Galiani (2007) find that Argentina's decentralization of secondary schools significantly increased test scores overall but decreased scores for schools in poor areas and in provinces with pre-decentralization fiscal deficits. On the other hand, King and Ozler's (1998) study of School Autonomy, a Nicaraguan SBM program, found that autonomous schools, most of which were catering to deprived areas, yielded results comparable to those of other schools. Unintended effects, or processes that eventually undermine desired effects, may exist as well. The Implementation Completion Report for the TEEP project indicates that, indeed, community engagement demanded that principals and teachers spend considerable time on community relations in addition to their administrative and pedagogical responsibilities, a commitment that several were beginning to find burdensome.

Thus, answers to questions regarding the conditions under which the different SBM models work and which implementation processes are effective are critically important from a policy and program design and implementation perspective. A new evaluation would present an opportunity to study such issues through focusing on such issues as well.

Despite these limitations, this study intends to contribute to the growing body of studies that examine school based management as a tool for improving student outcomes. This study provides an initial case study on the Philippines, an area where SBM has not yet been evaluated. Thus, this analysis suggests there are possible benefits from SBM that may be applicable in the country as a whole.

We end this paper with some words of advice for future evaluations of the ongoing SBM program in the Philippines. Barrera-Osorio et al.'s (2009) review of SBM studies concludes that retrospective evaluations of such program are extremely difficult to implement and recommends using prospective methods. The SBM program in the Philippines is an ongoing exercise, and the analysis in our paper suggests that it was successful with respect to its objectives of improving student achievement, although the effects were small. However, as the Department of Education in the Philippines moves forward with SBM and goes to scale across the country, it is important to collect data simultaneously in a systematic manner to enable scientific evaluations. The current data has several limitations which make it difficult to espouse conclusive statements about the success of the program and how it might generalize beyond the 23 districts to the country as a whole. A new evaluation approach could enable the Department of Education to study both the implementation processes (what really makes SBM work) and assess impact in a more rigorous fashion.

REFERENCES

Bardhan, P. (2002). *Decentralization of Governance and Development*. Journal of Economic Perspectives 16(4): 185–205.

Barrera-Osorio, F., T. Fasih, & H.A. Patrinos (2009). *Decentralized decision-making in school. The theory and evidence on school-based management.* The World Bank. Washington, D.C.

Behrman, J. R. and E. M. King, 2008, *Program Impact and Variation in Duration ofExposure*, in Samia Amin, Jishnu Das, and Markus Goldstein, eds, *Are You Being Served: New Tools for Measuring Service Delivery*, Washington, DC: World Bank, 147-172.

Berlinski, S. & Galiani, S. (2007). *The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment*. Labour Economics, Elsevier, vol. 14(3), pages 665-680, June.

Borman, G.D., Hewes, G. M., Overman L.T., & Brown, S. (2002). *Comprehensive school reform and student achievement: A meta analysis.* Report no. 59, University of North Carolina. Greensboro.

Caldwell, B. J. (2005). *School-based Management*. Education Policy Series. The International Institute for Educational Planning and The International Academy of Education, Paris and Brussels.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (second ed.). Lawrence Erlbaum Associates.

De Grauwe, A. (2004) *School-based management (SBM): does it improve quality*? Background Paper for All Global Monitoring Report 2005.

Dehejia, R. H. and Wahba, S. (2002.) *Propensity Score-Matching Methods for Nonexperimental Causal Studies*. The Review of Economics and Statistics, Vol. 84, No. 1, Pages 151-161.

Eskeland, G. and Filmer, D. (2002) *Does Decentralization Improve Learning? Autonomy and Parental Participation in Argentine Schools.* The World Bank. Washington D.C.

Galiani, S. and Schargrodsky, E. (2001) *Evaluating the Impact of School Decentralization on Education Quality.* Economica. Vol. 2(2): 275-314.

Gertler, P., Patrinos, H. and Rubio-Codina, M. (2006). *Empowering Parents to Improve Education: Evidence from Rural Mexico.* Impact Evaluation Series No. 4, World Bank Policy Research Working Paper No. 3935. The World Bank. Washington D.C.

Gunnarsson, Louise V., Orazem, P., Sanchez, M., and Verdisco, A. (2004). *Does School Decentralization Raise Student Outcomes? Theory and Evidence on the Roles of School Autonomy and Community Participation*. Working Paper 04005, Iowa Stata University Department of Economics.

Heckman, J., Hidehiko I., and Todd, P. (1997). *Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme.* Review of Economic Studies 64(4): 605-54.

Heckman, J., Hidehiko I., Jeffrey S., and Todd, P. (1998a.) *Characterizing Selection Bias Using Experimental Data.* Econometrica 66(5):1017-1098.

Heckman, J., Hidehiko I., Jeffrey S., and Todd, P. (1996) *Sources of Selection Bias in Evaluation Social Programs: An Interpretation of Conventional Measures and Evidence on the Effectiveness of Matching as a Program Evaluation Method.* Proceedings of the National Academy of Sciences. Vol. 93 No. 23 13416-13420.

Heckman, J., Hidehiko I., and Todd, P. (1998b.) *Matching as an Econometric Evaluation Estimator.* Review of Economic Studies 65(2): 261-294.

Jimenez, E. and Sawada, Y. (1998). *Do community-managed schools work? An evaluation of El Salvador's EDUCO program*. Working paper No. 8, series on Impact Evaluation of Education Reforms, Development Economics Research Group, The World Bank.

King, E. and Ozler, B. (1998). *What decentralization got to do with learning? The case of Nicaragua's school autonomy reform.* Working paper No. 9, Series on Impact Evaluation of Education Reforms, Development Economics Research .Group, The World Bank

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

Sawada, Y. and Ragatz, A. (2005). *Decentralization of Education, Teacher Behavior and Outcomes: The Case of El Salvador's EDUCO Program*. In Emiliana Vegas, Ed., *Incentives to Improve Teaching: Lessons from Latin America*. Washington, DC. The World Bank.

Skoufias, E. and Shapiro, J. (2006). *Evaluating the Impact of Mexico's Quality School Program: The Pitfalls of Using Nonexperimental data.* Impact Evaluation Series No. 4, World Bank Policy Research Working Paper No. 4036. The World Bank. Washington D.C.

World Bank. (2007). *What is School Based Management? An Introduction.* Education Human Development Network. The World Bank, Washington D.C.

Table 1: Descriptive Statistics, by All Schools, SBM, and Non-SBM

	All Schools			SBM Schools			Non-SBM School		
Variable	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**
Dependent Variables, Year= 2002									
Math	5167	46.14	17.32	1666	46.45	17.25	3501	46	17.35
English	5167	43.09	15.11	1666	43.29	14.79	3501	43	15.25
Science	5167	44.69	13.53	1666	45.14	13.17	3501	44.48	13.7
Total	5167	44.65	14.06	1666	44.97	14.01	3501	44.5	14.09
Dependent Variables, Year= 2003									
Math	5167	53.98	17.15	1666	54.95	16.99	3501	53.51	17.2
English	5167	50.01	15.11	1666	51.02	14.9	3501	49.54	15.19
Science	5167	49.86	13.35	1666	50.81	13.25	3501	49.41	13.38
Total	5167	51.28	14.26	1666	52.26	14.18	3501	50.82	14.28
Dependent Variables, Year= 2004									
Math	5167	61.97	17.1	1666	63.06	16.54	3501	61.46	17.34
English	5167	55.51	14.02	1666	56.9	13.56	3501	54.85	14.18
Science	5167	60.38	14.88	1666	61.88	14.49	3501	59.66	15.01
Total	5167	59.79	13.83	1666	61.15	13.38	3501	59.14	14
Other School Characteristics									
Total enrollment	5167	266.06	219.64	1666	348.9	6.96	3501	226.86	2.86
Total Dropouts	5167	8.97	13.26	1666	11.78	16.19	3501	7.63	11.38
Student teacher ratio	5167	32.45	14.63	1666	31.64	11.35	3501	32.83	15.94
Textbooks per student	5167	3.58	2.58	1666	3.29	1.81	3501	3.71	2.86
Manuals per teacher	5167	4.05	2.98	1666	3.81	2.49	3501	4.17	3.18
Head teacher	5167	0.25	0.43	1666	0.29	0.45	3501	0.22	0.42
Head principal	5167	0.21	0.41	1666	0.33	0.47	3501	0.15	0.35
Complete combination and multigrade sch	5167	0.28	0.45	1666	0.16	0.37	3501	0.34	0.47
Complete Large and Medium Schools	5167	0.24	0.43	1666	0.39	0.49	3501	0.17	0.38
Complete Small Schools	5167	0.44	0.2	1666	0.44	0.13	3501	0.45	0.23
Primary Schools	5167	0.03	0.17	1666	0.01	0.09	3501	0.04	0.19
Elementary Leader Schools	5167	0.07	0.25	1666	0.15	0.36	3501	0.03	0.16
School received SIIF funds	5167	0.34	0.47	1666	0.48	0.5	3501	0.27	0.44

Table 2: Differences in Scores between SBM and Non-SBM Schools during Pre-Intervention (2002-2003) and Implementation (2003-2004) Periods

	Math Score		Science Score		English Score		Total Score	
	1A	1B	2A	2B	3A	3B	4A	4B
Diff in 2002	0.463	2.112	0.289	1.845	0.669	2.148	0.469	2.022
	[0.90]	[4.35]***	[0.64]	[4.35]***	[1.67]*	[5.63]***	[1.11]	[5.13]***
Diff in 2003	1.432	3.081	1.469	3.025	1.387	2.865	1.429	2.982
	[2.79]***	[6.35]***	[3.27]***	[7.13]***	[3.47]***	[7.51]***	[3.39]***	[7.57]***
Division FE		Incd		Incd		Incd		Incd
N	10334	10334	10334	10334	10334	10334	10334	10334

Note: Absolute value of t statistics in parentheses.

Note: A denotes OLS; B denotes Fixed Effects

*Significant at 10%; **Significant at 5%; ***Significant at 1%

Table 3: Estimation of Propensity Scores

Number of observations	=	5167
LR Chi2(5)	=	439.44
Prob>chi2	=	0
Pseudo R2	=	0.0676
Log likelihood	=	-3028.71

Dependent variable: SBM

Variables	Coefficient	Std. Error	z
Student-teacher ratio	-0.011	0.002	-4.68
School size	0.001	0.000	9.60
Complete schools	0.471	0.253	1.86
Monograde schools	0.544	0.083	6.55
School head principals	0.370	0.086	4.29
Intercept	-1.848	0.271	-6.82

The common support option was selected. The region of common support is .119, .943
The final number of blocks is 10. This number of blocks ensures that the mean propensity score is not different for treated and controls in each block.

Table 4: P-Score Quintiles, by SBM and Non-SBM Schools

	Quintile1 (n=1030)			Quintile2 (n=1030)			Quintile3(n=1029)			Quintile4 (n=1030)			Quintile5 (n=1029)		
	Non-SBM	SBM	t-score	Non-SBM	SBM	t-score	Non-SBM	SBM	t-score	Non-SBM	SBM	t-score	Non-SBM	SBM	t-score
Propensity Score	0.174	0.174	0.24	0.23	0.24	3.27***	0.31	0.31	1.42	0.35	0.35	1.42	0.52	0.54	3.24***
Head_Principal	0.002	0.001	0.65	0.021	0.048	2.26**	0.01	0.003	1.27	0.14	0.11	1.42	0.88	0.84	1.57
Complete Schools	0.92	0.93	0.56	0.006	0.013	0.66				1	1		1	1	
Monograde Schools	0.08	0.06	1.19	0.49	0.55	1.49				0.99	1	0.78			
Student teacher ratio	36.78	37.24	0.36	32.21	33.04	0.59	28.95	30.46	2.34	31.5	31.4	0.16	31.6	30.2	2.28**
Total Enrollment 2002	114.5	121.3	1.64*	151.38	165.2	2.32**	193.4	202.6	2.34**	280.4	290.7	1.92**	514.7	574.5	3.14***

*Significant at 10%; **Significant at 5%; ***Significant at 1%

Table 5: Unconditional Test Score Means, By SBM/Non-SBM and Year

	2002	2004	Diff	t-stat
Math Scores	46.14	61.96	15.82	46.71***
SBM	46.45	63.03	16.57	28.28***
non-SBM	45.99	61.45	15.46	37.29***
Diff	0.46	1.58	1.12	
t-stat	0.89	3.11***		
Science Scores	44.68	60.36	15.68	56.01***
SBM	45.14	56.87	11.74	34.82***
non-SBM	44.47	54.84	10.37	44.21***
Diff	0.67	2.03	1.36	
t-stat	1.66*	4.99***		
English Scores	43.09	55.50	12.41	43.29***
SBM	43.28	61.86	18.57	27.63***
non-SBM	42.99	59.65	16.66	33.66***
Diff	0.29	2.20	1.91	
t-stat	0.65	4.88***		
Overall Scores	44.64	59.78	15.13	55.14***
SBM	44.96	61.12	16.16	34.02***
non-SBM	44.49	59.14	14.65	43.63***
Diff	0.47	1.98	1.51	
t-stat	1.12	4.82***		

Table 6: Treatment Effects Based on OLS, DID (with fixed effects), and DID with Kernel-PSM

	Scores	OLS	Fixed Effects	Kernel-PSM
1	Math Scores	0.851	0.743	1.88*
2	English Scores	1.141**	1.046**	1.361**
3	Science Scores	1.368***	1.283***	1.818***
4	Overall Scores	1.151**	1.058**	1.45**

*Significant at 10%; **Significant at 5%; ***Significant at 1%

Appendix A: OLS Regression Estimates of SBM Effects on School-Averaged Test Scores

Dependent Variable	Math	English	Science	Overall
Post-intervention (2004 – year 3)	12.357	9.188	13.14	12.076
	[33.25]***	[28.85]***	[44.30]***	[40.02]***
SBM Batch1 School	-0.381	-0.568	-0.75	-0.541
	[0.98]	[1.70]*	[2.41]**	[1.71]*
SBM School x Post-intervention	0.851	1.141	1.368	1.151
	[1.32]	[2.07]**	[2.66]***	[2.20]**
Does the school have any dropouts?	3.443	2.901	2.222	2.883
	[9.37]***	[9.21]***	[7.58]***	[9.66]***
No of dropouts, given that the school had >0 dropouts	-0.336	-0.273	-0.264	-0.294
	[11.94]***	[11.01]***	[11.44]***	[12.53]***
Size of School (total enrollment)	0	0	0	0
	[.01]	[0.55]	[0.04]	[0.23]
Student Teacher Ratio	0.027	-0.003	-0.004	0.007
	[2.73]***	[0.28]	[0.49]	[0.83]
Textbook per student	0.742	0.589	0.641	0.656
	[8.99]***	[8.30]***	[9.69]***	[9.76]***
Teacher Manual per teacher	-0.522	-0.367	-0.417	-0.44
	[7.28]***	[5.95]***	[7.26]***	[7.53]***
Head Teacher	2.872	2.713	2.885	2.77
	[7.43]***	[8.18]***	[9.35]***	[8.82]***
Principal	1.299	2.057	2.186	1.896
	[2.59]***	[4.72]***	[5.39]***	[4.59]***
Complete Combination and Multigrade School	3.737	2.207	2.187	2.677
	[4.38]***	[3.02]***	[3.21]***	[3.87]***
Complete Mono Large and Medium School	7.312	5.652	6.384	6.303
	[7.77]***	[6.40]***	[7.77]***	[7.54]***
Complete Mono Small School	5.304	3.365	3.908	4.097
	[6.18]***	[4.53]***	[5.64]***	[5.82]***
Elementary Leader School	1.679	1.289	2.152	1.597
	[2.70]***	[2.41]**	[4.33]***	[3.16]***
Received SIIF Grant	0.756	0.903	1.187	0.955
	[2.34]**	[3.26]***	[4.59]***	[3.63]***
Constant	41.259	40.164	41.032	40.906
	[44.63]***	[50.62]***	[55.53]***	[54.42]***
Observations	14823	14823	14823	14823
R-squared	0.13	0.11	0.21	0.18

Appendix B: Difference-in-Difference with Division Level Fixed Estimates of SBM Effects on School-Averaged Test Scores

Dependent variable	(1) Math	(2) English	(3) Science	(4) Overall
Post-intervention (2004 – year 3)	12.484	9.328	13.23	12.186
	[36.30]***	[31.69]***	[48.16]***	[44.30]***
SBM Batch1 School	1.071	0.63	0.49	0.706
	[2.85]***	[1.96]**	[1.63]	[2.35]**
SBM School x Post-intervention	0.743	1.046	1.283	1.058
	[1.25]	[2.05]**	[2.70]***	[2.22]**
Size of School (total enrollment)	0	0	0	0
	[0.99]	[0.54]	[0.18]	[0.11]
Does the school have any dropouts?	3.78	3.322	2.442	3.171
	[10.88]***	[11.17]***	[8.80]***	[11.41]***
No of dropouts, given that the school had >0 dropouts	-0.208	-0.164	-0.162	-0.183
	[7.64]***	[7.04]***	[7.45]***	[8.41]***
Student Teacher Ratio	0.017	-0.002	-0.004	0.003
	[1.70]*	[0.29]	[0.46]	[0.38]
Textbook per student	0.229	0.159	0.229	0.198
	[2.72]***	[2.20]**	[3.40]***	[2.94]***
Teacher Manual per teacher	-0.222	-0.139	-0.206	-0.184
	[3.18]***	[2.33]**	[3.70]***	[3.30]***
Head Teacher	2.213	2.171	2.261	2.177
	[6.13]***	[7.02]***	[7.83]***	[7.53]***
Principal	1.832	2.281	2.304	2.193
	[3.81]***	[5.54]***	[6.00]***	[5.70]***
Complete Combination and Multigrade School	3.341	2.098	2.184	2.486
	[4.22]***	[3.09]***	[3.45]***	[3.92]***
Complete Mono Large and Medium School	6.937	5.254	6.001	5.902
	[7.23]***	[6.39]***	[7.83]***	[7.69]***
Complete Mono Small School	5.89	3.975	4.541	4.685
	[7.30]***	[5.76]***	[7.05]***	[7.26]***
Elementary Leader School	2.582	1.983	2.79	2.378
	[4.46]***	[4.00]***	[6.03]***	[5.13]***
Received SIIF Grant	0.145	0.287	0.534	0.311
	[0.47]	[1.09]	[2.17]**	[1.26]
Constant	41.426	39.704	40.672	40.764
	[47.15]***	[52.79]***	[57.95]***	[58.01]***
Observations	14823	14823	14823	14823
Division fixed effects	Incld	Incld	Incld	Incld
R-squared	0.14	0.12	0.23	0.2

Appendix C: DID with (Kernel) PSM of SBM Effects on School-Averaged Test

Dependent variable	(1) Math	(2) English	(3) Science	(4) Overall
Post-intervention (2004- year 3)	11.152	8.675	12.182	11.251
	[26.46]***	[24.12]***	[36.27]***	[33.01]***
SBM Batch1 School	-0.668	-0.77	-0.83	-0.737
	[1.54]	[2.07]**	[2.39]**	[2.10]**
SBM School x Post-intervention	1.188	1.361	1.818	1.45
	[1.66]*	[2.22]**	[3.18]***	[2.50]**
Size of School (total enrollment)	0	0	0	0
	[0.14]	[0.55]	[0.16]	[0.20]
Does the school have any dropouts?	3.338	2.897	2.345	2.917
	[8.03]***	[8.17]***	[7.08]***	[8.68]***
No of dropouts, given that the school had >0 dropouts	-0.313	-0.254	-0.239	-0.271
	[10.32]***	[9.80]***	[9.87]***	[11.03]***
Student Teacher Ratio	0.012	-0.017	-0.02	-0.007
	[0.93]	[1.62]	[1.96]**	[0.69]
Textbook per student	0.947	0.712	0.77	0.808
	[9.14]***	[8.05]***	[9.32]***	[9.64]***
Teacher Manual per teacher	-0.557	-0.349	-0.392	-0.446
	[6.65]***	[4.88]***	[5.87]***	[6.58]***
Head Teacher	2.626	2.233	2.64	2.473
	[5.93]***	[5.91]***	[7.48]***	[6.91]***
Principal	0.762	1.628	1.975	1.515
	[1.32]	[3.31]***	[4.30]***	[3.25]***
Complete Combination and Multigrade School	3.744	2.437	2.3	2.871
	[3.90]***	[2.97]***	[3.01]***	[3.70]***
Complete Mono Large and Medium School	7.07	5.712	6.055	6.189
	[6.11]***	[5.78]***	[6.56]***	[6.61]***
Complete Mono Small School	5.104	3.517	3.716	4.062
	[5.22]***	[4.21]***	[4.77]***	[5.14]***
Elementary Leader School	1.47	0.794	1.814	1.245
	[2.12]**	[1.34]	[3.28]***	[2.22]**
Received SIIF Grant	1.678	1.777	1.998	1.84
	[4.69]***	[5.83]***	[7.02]***	[6.36]***
Constant	41.001	39.425	40.452	40.294
	[38.29]***	[43.15]***	[47.42]***	[46.54]***
Observations	11349	11349	11349	11349
R-squared	0.12	0.11	0.2	0.17

Absolute value of t statistics in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%