

SIEPR

policy brief

Stanford Institute for Economic Policy Research

on the web: <http://siepr.stanford.edu>

The Implementation and Evaluation of School Decentralization Programs in Developing Economies: Evidence from India

By Anjini Kochar

The Growing Popularity of School Decentralization in Developing Economies

Low learning levels in many developing economies are most frequently blamed on the lack of accountability of teachers and school administrators in the centralized schooling systems that characterize these economies. To address this, governments are increasingly decentralizing oversight of schools, requiring schools to form school development and management committees (SDMCs). These committees, comprising parents, teachers, and members of the local community, are vested with the responsibility of overseeing teachers and ensuring school performance. This represents a significant change in the institutional arrangements that govern education. In most developing countries, prior to this change, poor rural communities, characterized by low levels of adult literacy, were only minimally involved in their children's

schooling and had little interaction with teachers and school administrators, most of whom had significantly higher levels of schooling than parents. Changing institutional arrangements requires changing norms of behavior within communities, providing parents with the information and managerial skills necessary for being effective partners in the schooling process, and putting in place arrangements that make teachers accountable to parents.

Because of the magnitude of the change in institutional arrangements, most governments have accompanied the shift to school decentralization with training programs for school management committees. However, the considerable change required has also made many skeptical about whether such training can succeed in improving the quality of SDMCs and in enhancing learning. In recent research (Gowda et al. 2014) we

continued on inside...

About The Author

Anjini Kochar, is both a senior research scholar and India Program director at the Stanford Center for International Development (SCID). Prior to her position at SCID she was an Assistant Professor in the Economics Department at Stanford University. Kochar has a Ph.D. in Economics and an M.A. in International Relations from the University of Chicago and a B.A. from Bryn Mawr College. Her research interests include the microeconomics of development, poverty, education and health issues in developing economies.



SIEPR *policy brief*

evaluated an intensive SDMC training program implemented by a non-governmental organization, Prajayatna, in the state of Karnataka, south India, one very much along the lines of government programs implemented in other developing countries. The program started with a community-level meeting that provided information to parents about the school, government policies, and the importance of parental involvement in schools. This was followed by monthly management training of SDMC members. These training meetings also served as a means of promoting accountability of teachers to the parent community: teachers were required to report and discuss students' learning levels.

Three years after its implementation, we found that the program did improve SDMC quality and that this, in turn, had modest effects on learning. As discussed below, our finding is in contrast to much of the current empirical evidence on the topic. This is a consequence of methodological and measurement differences, with the latter reflecting differences in the test scores used by researchers and those to which schools and teacher are held accountable.

Evaluation Results From Previous Studies

Available empirical evidence suggests that greater community control of schools has been difficult to achieve. Programs intended to enhance community participation in schools by providing parents information about schools have been shown to be relatively unsuccessful in countries such as India (Banerjee,

Banerji, Duflo, Glennerster and Khemani 2010). Similarly, programs that train and fund school management committees over a 3-4 year period also appear to have had little effect, both on community involvement and on learning levels (Blimpo and Evans 2011, Pradhan et al. 2011). Greater success in improving learning appears to come from programs with a narrower focus on providing a specific input or teaching tool, rather than changing institutional structures. Reflecting the weak results from broadly designed medium- to long-term programs in contrast to evidence of success from narrowly defined short-term programs, an emerging consensus is that providing a specific pathway to engage communities in schools without attempting the broader objective of changing underlying institutional arrangements may be more successful (Banerji et al 2010, Blimpo and Evans 2011).

These studies rely on randomized controlled trials (RCTs) to establish the causal effects of the program on learning outcomes. RCTs have become increasingly popular in the program evaluation literature because of their simplicity. When schools or individuals either self-select or are selected for participation in programs, the sample of participants generally differs in observed traits from non-participants, but also in unobserved characteristics that may directly influence outcomes. A simple comparison of mean outcomes across participants and non-participants thus confounds the effect of the program with these differences in attributes. In contrast, when schools are randomly assigned to treatment and control samples, mean

learning and other outcomes across these samples show no significant difference in a baseline study. Consequently, any improvement in learning at the time of the end-line survey can be attributed to the program.

To compare our results to those in the literature, we initially used this same methodology to evaluate the effect of the program on SDMC quality¹ and test scores in language and mathematics, randomly selecting clusters for treatment and control samples and comparing mean outcomes across these samples. Learning tests, developed by an independent testing agency, evaluated students against state-specified standards for their grade. As in similar programs in other countries (Blimpo and Evans 2011, Pradhan et al 2011), we found only minimal effects of the program on both SDMC quality and test scores.

These results may be misleading for two reasons. First, the complexity of programs whose goal is institutional change makes delays in implementation unavoidable. As we discuss below, this makes it difficult to interpret the results from RCTs. A second issue relates to the data used to assess teacher accountability. If accountability mechanisms work, they will likely be reflected only in the scores that teachers are evaluated on. These scores may differ from survey test scores, even when the latter are designed to test state-specified standards.

Implementation Failures

Our first set of results reflects an oft-voiced concern with the

¹ Measured through SDMC member responses to a detailed survey that evaluated their knowledge of SDMC managerial roles and responsibilities and the extent to which they were being implemented.



use of RCTs for establishing causal effects of programs with weak implementation. This is generally the case when the objective is institutional change. Such change requires more than a one-time intervention; prolonged and sustained intervention over a period of years is needed. With a relatively long time span, however, delays in implementation are likely, both as a consequence of endogenous factors such as the inability to assemble the necessary quorum of members during the busy harvest season to weather and other exogenous factors that may upset training schedules. That is, delays need not signal “good” or “bad” institutions; all institutions may well complete the program in its entirety. However, they imply that at any given point of time, such as at the end-line survey, schools may differ in the extent of their implementation of the project. In the project we study, the average number of meetings conducted over a three-year period was only 11 out of a targeted 27. However, the number of meetings ranged across schools from 2 to 22.

Such variation implies that RCTs, at best, measure only the effect of the program at the average level of implementation at the time of the end-line study. From a policy point of view, it is more important to assess whether successfully implemented programs provide an effective means of improving learning. This question can be answered by exploiting the variance in the total number of training meetings across schools. However, in doing so, changes in the SDMC over the course of the project period make it necessary to distinguish meetings conducted with the SDMC in place at the end-line

survey from those conducted with the SDMC at base-line; meetings conducted with an earlier committee will likely have a far weaker effect on the quality of the current SDMC. Changes in SDMCs during the project are common because committees serve only a 2- or 3-year term. In contrast, the duration of programs intended to change institutional arrangements is understandably longer (4 years for the program we study).

Evaluating the success of the program by comparing SDMC quality across treatment and control schools at the time of the end-line survey combines the effect of training with old and current committees. If training provided to the old committee has little effect on the quality of the new committee, then the use of RCTs in this context yields estimates of the effectiveness of the program on SDMC quality that are biased downwards. This helps explain the insignificant results from this and other studies that base evaluation on RCTs.

Utilizing the variation in the timing of SDMC elections in our sample of schools, we provide an alternative set of results that estimates the effect of training conducted with the current SDMC. We find a significant effect of such training on SDMC quality. Breaking down our measure of SDMC quality into different components, we find that the program was particularly successful in enhancing committee members’ knowledge of students’ learning levels and promoting teacher accountability.

Effect on Learning

Given this result, one should expect a positive effect of high-quality SDMCs on learning. As before, the randomization cannot

be used to infer the effect of SDMCs on learning because it fails to capture the differential effect of training on old and current committees. We again exploit the fact that SDMC quality improves over time, comparing outcomes for newly appointed committees with those that have been in place for longer periods of time. We find a small, statistically significant, effect of SDMC quality on language test scores, but little effect on mathematics scores.

Accountability and Assessment

We suggest that the relatively weak effect on learning is a consequence of teachers being evaluated by measures that differ significantly from our survey tests. As is well recognized by the literature on the topic, accountability systems give schools incentives to improve scores only along the dimensions on which they are evaluated; they need not result in improvements in general skills or skill sets other than those for which schools are accountable. In Texas schools, for example, Klein et al. (2000) demonstrated that accountability resulted in improvement in the test scores that schools were evaluated on (the Texas Assessment of Academic Skills, TAAS), but not in scores on other comparable tests (National Assessment of Educational Progress, NAEP). This difference in achievement in “high stakes” versus “low stakes” tests has also been noted by other researchers (Figlio and Rouse 2006). And, while it may well be the case that different tests measure different skill sets, there is also the possibility that teachers may simply inflate scores for those students who matter the most

in order to meet accountability standards (Jacob and Levitt 2003).

The link between accountability and assessment in developing economies has not received the same attention that it has in developed economies. For example, even while the Indian Government emphasizes the importance of holding teachers and schools accountable for improving learning, the Right to Education Act of 2009 eliminates any centralized testing of students during elementary school, leaving assessment up to schools and teachers (through a system referred to as “continuous and comprehensive” evaluation). Local communities can then evaluate learning improvements only on the basis of information provided by teachers. This may differ from the information contained in the survey tests implemented by researchers who are justifiably hesitant to evaluate learning based on school-specific scores.

School and Survey Tests

Distribution functions for “school scores”, from school tests designed, implemented and corrected by teachers, and “survey scores” from tests that we implemented and corrected are graphed in figures 1a (language) and b (mathematics). The figures reveal that the distribution of school test scores lies to the right of survey tests: school test scores significantly exceed survey test scores, with the mean school score on language being 66.7 compared to the mean survey score of 37.1. Mean school and survey scores for mathematics are 63.4 and 37.1 respectively. To evaluate whether this could reflect a greater degree of difficulty of our survey tests relative to school tests, we also graph the distribution function of

survey test scores for this same set of students (currently in grade 5), but only from questions that evaluate knowledge of grade 3 subject matter. School test scores lie to the right of this distribution, as well, suggesting that school tests evaluate grade 5 students at a level below that intended for grade 3 students.

Both tests may still identify the same set of “good” and “weak” students. To assess this, we plot the distribution of school test scores for the set of students who placed in the bottom and top quartiles of the survey test (figures 2a and b). These histograms reveal that, despite the positive correlation between

these two measures, a significant proportion of students from the lowest quartile of the survey test distribution is in the top two quartiles of the school test distribution. This difference in the skill set identified by school and by survey test scores suggests differences in their determinants, and hence in their responsiveness to different policy instruments.

Program Effects on School Scores

Supporting the hypothesis that the marginal effect of the program on survey test scores reflects the fact that teachers are

continued on flap...

Figure 1: Distributions for School and Survey Test Scores

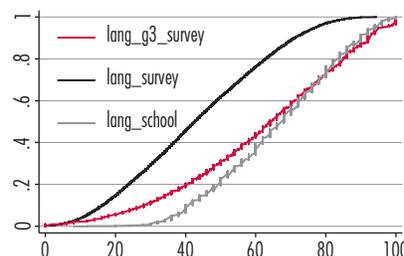


Figure 1A: Language

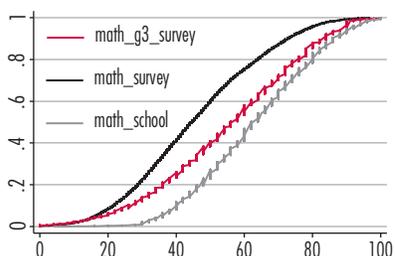


Figure 1B: Mathematics

Figure 2: Mathematics School Scores, for students who placed in the lowest and highest quartile of the survey test

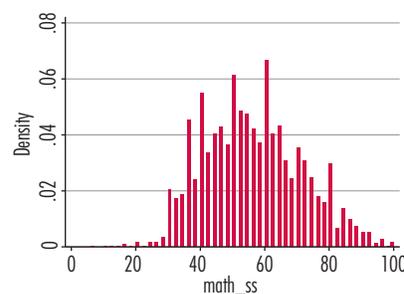


Figure 2A: Lowest quartile of survey test

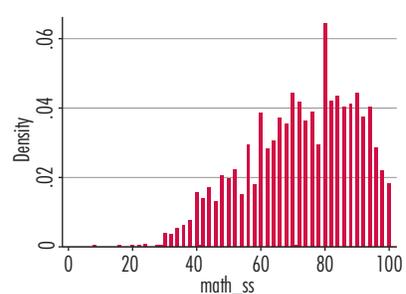


Figure 2B: Highest quartile of survey test

not accountable for these scores, we find that the program does increase the school scores that teachers are accountable for. School scores for both language and mathematics are higher in treatment schools in comparison to those in control. Moreover, we find that, as in similar research conducted on U.S. systems of teacher accountability, school scores rise primarily for the weakest students who are the focus of attention in meetings between the SDMC and teachers.

The significant effect of the program on school scores, but not on survey scores, suggests that the two tests assess different skills. We believe that our survey test better measures learning: We find that the survey test scores respond to factors known to enhance learning (specifically, to improvements in student-teacher ratios), but school scores do not. We also show that parents are more influenced to make decisions that affect their children's schooling outcomes by school scores than by survey scores. Thus, the failure to test students in schools through a set of centrally administered and verifiable tests may be a significant determinant of low levels of learning in developing economies such as India.

Policy Implications

Our research highlights the difficulties in changing institutional

arrangements, but also suggests that such changes are possible and that, despite the high illiteracy and poverty rates that characterize the rural areas of developing economies, local school management committees can be effective partners in schooling. Good policies can, however, help make local institutions more effective. For example, policies that diminish the costs of transition from one committee to another, such as a staggered replacement of SDMC members, can improve SDMC quality.

Our study also emphasizes that effective accountability requires attention to assessment. Promoting institutions to enhance teacher accountability will not improve learning unless the outcomes that teachers are accountable for are well defined and unless their correlation with learning outcomes is well established. In the current system in India, without any centralized testing through the elementary schooling years it is quite possible to have improvements in accountability without any improvement in learning.

References

Banerjee, Abhijit V., Rukmini Banerji, Esther Duflo, Rachel Glennerster and Stuti Khemani. 2010. "Pitfalls of participatory Programs: Evidence from a Randomized Evaluation in

Education in India." *American Economic Journal: Economic Policy* 2:1: 1-30.

Blimpo, Moussa P. and David K. Evans. 2011. "School-Based Management and Educational Outcomes: Lessons from a Randomized Field Experiment." Stanford University: Manuscript.

Figlio, D. and C.E. Rouse. 2006. "Do Accountability and Voucher Threats Improve Low-Performing Schools?" *Journal of Public Economics* 90(1-2): 239-255.

Gowda, Kallan, Anjini Kochar, C.S. Nagabhushana and N. Raghunathan. 2014. "Evaluating Institutional Change through Randomized Controlled Trials: Evidence from School Decentralization in India." Stanford University: Manuscript.

Jacob, B.A. and S.D. Levitt. 2003. "Rotten Applies: An Investigation of the Prevalence and Predictions of Teacher Cheating." *Quarterly Journal of Economics* 118(3): 843-878.

Klein, S., L. Hamilton, D. McCaffrey and B. Stecher. 2000. "What do Test Scores in Texas Tell Us? *Education Policy Analysis Archives* 9(49).

Pradhan, Menno, Daniel Suryadarma, Amanda Beatty, Maisy Wong, Armida Alishjabana, Arya Gaduh and Rima Prama Artha. 2011. "Improving Educational Quality Through Enhancing Community Participation." Washington, D.C.: The World Bank.

SIEPR *policy brief*

Stanford University
366 Galvez Street
Stanford, CA 94305
MC 6015

A publication of the
Stanford Institute for
Economic Policy Research

Non-Profit Org.
U.S. Postage
PAID
Palo Alto, CA
Permit No. 28

SIEPR

About SIEPR

The Stanford Institute for Economic Policy Research (SIEPR) conducts research on important economic policy issues facing the United States and other countries. SIEPR's goal is to inform policymakers and to influence their decisions with long-term policy solutions.

Policy Briefs

SIEPR policy briefs are meant to inform and summarize important research by SIEPR faculty. Selecting a different economic topic each month, SIEPR will bring you up-to-date information and analysis on the issues involved.

SIEPR policy briefs reflect the views of the author. SIEPR is a non-partisan institute and does not take a stand on any issue.

For Additional Copies

Please see SIEPR website at
<http://SIEPR.stanford.edu>.