TOPICS IN DEVELOPMENT - PART I

REMEDYING EDUCATION: EVIDENCE FROM TWO RANDOMIZED EXPERIMENTS IN INDIA

ABSTRACT. This paper presents the results of two randomized experiments conducted in schools in urban India. A remedial education program hired young women to teach students lagging behind in basic literacy and numeracy skills. It increased average test scores of all children in treatment schools by 0.28 standard deviation, mostly due to large gains experienced by children at the bottom of the test-score distribution. A computer-assisted learning program focusing on math increased math scores by 0.47 standard deviation. One year after the programs were over, initial gains remained significant for targeted children, but they faded to about 0.10 standard deviation.

BG:.

(1) Remedial Education: The Balsakhi Program, where a young woman from the community works on basic skills with children who have reached grade 3 or 4 without having mastered them. These children are taken out of the regular classroom to work with this young woman for two hours per day (the school day is about four hours). The program has very low overhead and capital costs, which distinguishes itself from standard remedial education programs in the developed world.
(2) Computer-Assisted Learning: It is a computer-assisted learning program where children in grade 4 are offered two hours of shared computer time per week during which they play games that involve solving math problems whose level of difficulty responds to their ability to solve them.

EVALUATION DESIGN. *Balsakhi (Vadodara):* To ensure a balanced sample, assignment was stratified by language, pretest score, and gender. Ninety-eight of Vadodara's 122 government primary schools participated in year 1 of the study. Half the schools (Group A) were given a balsakhi to work with children in grade 3; the other half (Group B) were given balsakhis to work in grade 4. Table I describes the design and reports the sample size of the study. The program continued during the school year 2002-2003, etc (see table 1). The estimates of the program effect would be biased downwards if the schools reassigned resources from one grade to the other in response to the program. In practice, the way schools are organized in urban India (and, in particular, in Vadodara and Mumbai) makes this extremely unlikely: schools have a fixed number of classes (a group of students and a teacher) per grade.

Computer-Assisted Learning (Vadodara): The CAL Program was first implemented in almost half of the municipal primary schools in Vadodara in 2002-2003, focusing exclusively on children in grade 4. In a few schools, computers could not physically be installed either because of space constraints or lack of electricity to run the computers. These schools were excluded from the randomization. In the final sample for the study, fifty-five schools received the CAL Program (Group A1B1), and fifty-six served as the comparison group (Group A2B2). The program was continued in 2003-2004, after switching the treatment and comparison groups.

Mumbai 2001-2002 and 2002-2003: We selected one ward (the L Ward) to implement a design similar to the design in Vadodara. In total, seventy-seven schools were included in the study. half the schools were randomly selected to receive a balsakhi in grade 3 (Group C, see Table I), and half the schools were randomly selected to receive a balsakhi in grade 2 (Group D). In the second year of the study, the Mumbai program experienced some

administrative difficulties...Throughout the paper, the schools that were assigned balsakhis but did not get them are included in the "intention to treat" group. The regression analysis then adjusts the estimates for the fraction of the treatment group that was effectively treated by using the initial assignment as an instrument for treatment.

[Table 1]

Outcomes. Differential attrition between the treatment and comparison groups could potentially bias the results. For example, if weak children were less likely to drop out when they benefited from a balsakhi, this could bias the program effect downwards. To minimize attrition, the testing team returned to the schools multiple times, and children who still failed to appear were tracked down at home and, if found, were administered the same test. Furthermore, the pretest scores of children who left the sample were similar in treatment and comparison groups, suggesting that the factors leading to attrition were the same in both groups. These two facts together suggest that attrition is unlikely to bias the results we present below.

The randomization appears to have been successful: with the exception of the CAL Program in year 3 in V adodara, none of the differences between the treatment and comparison groups prior to the implementation of the program are statistically distinguishable from zero.

Only 19.5 percent of third grade children in V adodara and 33.7 percent in Mumbai pass the grade 1 competencies in math. The results are more encouraging in verbal competencies. The baseline achievement level is much higher in Mumbai, where students are less poor than in Vadodara, and schools have better facilities. Another outcome of interest is attendance and dropout rates.

[Table 2]

SHORT-TERM EFFECTS.

Balsakhi Program. The Balsakhi Program appears to be successful: in all years, for both subjects, in both cities, and for all subgroups, the difference in post test scores between treatment and comparison groups is positive and, in most instances, significant. Then regresse the change in a student's test score (post-test score minus pretest score) on the treatment status of the child's school-grade, controlling for the pretest score of child i in .grade g and school j:

$$Y_{igjPOST} - Y_{igjPRE} = \lambda + \delta D_{jg} + \theta_{yigjPRE} + \varepsilon_{igjPOST} \tag{0.1}$$

For all years and samples, except Mumbai in year 2, (1) is estimated with OLS. However, for Mumbai in year 2 (and when both cities are pooled), to account for the fact that not all schools actually received a balsakhi, (1) is estimated by two stage least squares, instrumenting for actual treatment status of the school-grade.

These estimates suggest a substantial treatment effect: the impact of the program on overall scores was 0.14 standard deviations overall in the first year, and 0.28 standard deviations in the second year, both very significant.

Comparing Mumbai and Vadodara, the effects are very similar for math in both years (0.19 in Vadodara vs. 0.16 in Mumbai in year 1, and 0.37 vs. 0.32 in year 2), but in Mumbai, the effects for language are weaker and insignificant in both years (0.09 and 0.07 in year 1 and year 2), while they are significant in both years in Vadodara. The lower impact of language in Mumbai is consistent with the fact observed above, that most children (83.7 percent) in Mumbai already had some basic reading skills and are therefore less in need ofa remedial program that targets the most basic competencies in language. In math, where more lag behind, the program was as effective as it was in Vadodara.

For both cities and both subjects, the effects are very similar in grade 3 and grade 4. Results are also very similar when the analysis is conducted separately for girls or boys .

Compared to the other educational interventions, this program thus appears to be quite effective in the short-run.

Compare to STAR, The Balsakhi Program improved test scores by 0.27 standard deviations in the second year by using alternative instructors for part of the day. Moreover, the balsakhis were paid less than one tenth the teacher's salary, making this a much more affordable policy option than reducing class size.

Computer-Assisted Learning. The math test scores are significantly greater in treatment schools than in comparison schools in both years. Table IV corrects for this initial difference by estimating (1), where the treatment is the participation of the school in the CAL program. The CAL program has a strong effect on math scores (0.35 standard deviations in the first year (year 2) and 0.47 standard deviations in the second year (year 3». It has no discernible impact on language scores (the point estimates are always very close to zero).

Panel B of Table IV compares the Balsakhi and the CAL effects and examines their interactions in year 2 (2002-2003) when they were implemented at the same time using a stratified design. When the two programs are considered in isolation, the CAL has a larger effect on math test scores than the Balsakhi Program (although this dif ference is not significant) and a smaller effect on overall test scores (although, again, the difference is not significant). he programs appear to have no interaction with each other: the coefficients on the interaction on the math and overall test score are negative and insignificant.

[Table IV]

Distributional Effects. The Balsakhi Program could have harmed children at the bottom of the distribution (by sending them to a less-qualified teacher) while benefiting children at the top of the distribution (by removing the laggards or trouble-makers from the classroom). While this could result in an improvement in the average test score, it should probably not be construed as a success of the program. It is therefore important to know who among the children were affected by the program. Table V (Panel A for Balsakhi, B for CAL) shows the results for the year 2002-2003 (year 2) broken into three groups to measure test score gains for children who scored in the top, middle, and bottom third in the pretest. For the Balsakhi Program, the effect is about twice as large for the bottom third than for the top third (0.47 standard deviations versus 0.23 standard deviations for the total score). The program therefore does seem to have been more beneficial to children who were initially lagging behind. Children in the bottom group were more than twice as likely to be sent to a balsakhi. For the CAL Program, the impact is also higher for the bottom third, but the difference is not as large.

[Table V]

LONGER-RuN IMPACT. we start by comparing the effect of being exposed one versus two years to the program: if the effects are durable, they should be cumulative.

It seems possible that the foundation laid in the first year of the program helped the children benefit from the second year of the program. The same, however, is not true for the two-year effect estimates in Vadodara where the two-year effect is slightly smaller than the one-year effect in the second year of the program (though it is larger than the first year's effect). Almost all of the gains because of the balsakhi in Vadodara in the first year accrued in the first half of the year. In fact, test scores significantly declined in the second half of the year for both treatment and control students, many of whom were traumatized and absent, even when the schools re-opened.

We then investigate whether the program effect lasts beyond the years during which the children were exposed. We were able to track a substantial fraction of these children. The size oft he effects falls substantially, and, indeed, for the Balsakhi Program, the average effect becomes insignificant. However, the effect for the bottom third of the children, who

were most likely to have spent time with the balsakhi and for whom the effect was initially the largest, remains significant

INSIDE THE Box: DIRECT AND INDIRECT EFFECTS. it conflates two effects: The program potentially had a direct impact on the children who were assigned to work with the balsakhi. It also could have had an indirect impact on the children who stayed behind in the classroom, both through a reduction in the number of students in the dass (a class-size effect) and by removing the weaker children from the room, which could change classroom dynamics (a peer effect). This suggests that the effect of the program may have been mainly due to children who were sent to the balsakhi, rather than to spillover effects on the other ones.

In Table VI, we present instrumental variables estimates of the direct and indirect impact of being in a balsakhi group, using the strategy described earlier.

The effect of the program appears concentrated on children who indeed worked with the balsakhi. The effect on the children sent to the balsakhi is large: they gain 0.6 standard deviations in overall test scores. The fact that the Balsakhi Program affects mostly children at the bottom of the test score distributions simply reflects the fact that the children at the bottom of the test score distribution are more likely to be assigned to the balsakhi group.

[Table VI]

CONCLUSION. Evaluations conducted in two cities over two years suggest that both are effective programs: the test scores of children whose schools benefited from the remedial education program improved by 0.14 standard deviations in the first year and by 0.28 in the second year. We also estimate that children who were directly affected by this program improved their test scores by 0.6 standard deviations in the second year, while children remaining in the regular classroom did not benefit. The computer-assisted learning program was also very effective, increasing math scores by 0.36 standard deviations the first year and by 0.54 standard deviations the second year.

Some may be puzzled by the effectiveness of these two programs and the lack of spillovers of the Balsakhi Program to the other children given that the balsakhis have less training than the formal teachers and that Computer-Assisted Learning Programs have not been shown to be effective in developed country settings.

Both programs, the Balsakhi Program in particular, are also remarkably cheap: Overall, the Balsakhi Program cost is approximately Rs. 107 ($2.25) per student per year, while the CAL Programs cost approximately Rs 722 ($15.18) per student per year, including the cost of computers and assuming a five-year depreciation cycle.

Nevertheless, these results suggest that it may be possible to dramatically increase the quality of education in urban India, an encouraging result since a large fraction of Indian children cannot read when they leave school.