Remedying Education: Evidence from Two Randomized Experiments in India

Abhijit Banerjee, Shawn Cole, Esther Duflo and Leigh Linden*

September 17, 2003

Abstract

This paper presents the results of a two-year randomized evaluation of a large scale remedial education program, conducted in Mumbai and Vadodara, India. The remedial education program hires young women from the community to teach basic literacy and numeracy skills to children who reach standard three or four without having mastered these competencies. The program, implemented by a NGO in collaboration with the government, is extremely cheap (it cost 5 dollars per child per year), and is easily replicable: It has been implemented in 20 Indian cities, and reached tens of thousands of children. We find the program to be very effective: On average, it increased learning by 0.15 standard deviations in the first year, and 0.25 in the second year. The gains are the largest for children at the bottom of the distribution: Children in the bottom third gain 0.2 standard deviations in the first year, and 0.32 in the second year. In math, they gain 0.51 standard deviation in the second year. The results are similar in the two grade levels, and in the two cities. At the margin, extending this program would be up to 12-16 times more cost effective than hiring new teachers.

*PRELIMINARY AND INCOMPLETE. This project is a collaborative exercise involving many people. Foremost, we are deeply indebted to the Pratham team, who made the evaluation possible and put up with endless changes in plans and new requests: Pratima Bandekar, Rukmini Banerji, Lekha Bhatt, Madhav Chavan, Shekhar Hardikar, and Rajashree Kabare. We thank Jim Berry and Marc Shotland, for their excellent work coordinating the fieldwork in Vadodara, and Mukesh Prajapati for assistance in Mumbai. They were helped by Nandit Bhat and many others. For financial support, we thank the ICICI corporation, the World Bank, and the MacArthur Foundation Network on the Costs of Inequality.

1 Introduction

There has been a lot of interest recently in the question of how to effectively deliver education to the poor in developing countries and a corresponding burgeoning of high quality research on the subject. A lot of the research focuses on the effects of reducing the cost of schooling, with the view that the important goal is to get the children into school. Examples of this kind of work include Banerjee, Jacob and Kremer (2002) on school meals in India, Duflo (2001) on school construction in Indonesia, Glewwe, Kremer and Moulin (1997) on school uniforms in Kenya, Spohr (1999) on compulsory schooling laws in Taiwan and Vermeersch (2002) on school meals for preschoolers in Kenya. The primary metric by which success is judged in these studies is attendance, and in each of these cases a significant impact was found.

Are students also learning measurably more as a result of these interventions? There is no obvious reason why they would. The influx of new students probably makes learning harder for the children who were already in school, simply because there are more demands on existing resources.[1] And while the newcomers will presumably learn more, just by the fact that they are now attending school, it is not clear that there is anyone with whom we could compare them.

At the other extreme are interventions that focus directly on improving test scores for students who are already in school. These are interventions where students are explicitly rewarded for doing well on tests: Angrist, et. al., (2002) study a program in Colombia that offers private school vouchers to students who keep their scores above a certain level. A recent study by Kremer, Miguel and Thornton (2002) looks at the impact of offering scholarships to students in Kenya who do well on a standardized test. Both studies find an impact on test scores, though in such cases the existence of an impact is perhaps less interesting than whether the gains are commensurate with the money spent.

Perhaps the most interesting case is the one in between: Interventions that purport to improve the quality of the learning experience, but for which no evidence exists that they actually do improve learning. Examples include increasing the teacher-student ratio (Banerjee, Jacob and Kremer, 2002), subsidized textbooks (Glewwe, Kremer and Moulin, 1997), free flip-charts (Glewwe, Kremer, Moulin and Zitzewitz, 1997); and then the interventions that improve the

[1] Indeed this is what Banerjee, Jacob and Kremer (2002) find for mid-day meals, and Glewwe, Kremer and Moulin (1997) find for a program that offered both free textbooks and free school uniforms.

health of school children (for example, deworming, as in Kremer and Miguel, 2002), incentives for teachers (Glewwe, Kremer and Moulin, 2002), and blackboards and other school inputs (Chin, 2001), etc. By improving school quality, these programs can increase attendance. One ought also to expect an improvement in test scores among those who were already in school. Nevertheless, it is notable that relatively few of the studies from developing countries report a positive impact on test scores for those who were already in school.[2] Moreover, one cannot rule out the idea that there is no impact on children's educational achievement *a priori*, because the quality of teaching in many schools leaves much to be desired. Or it could even be the case that the children do not learn because they do not want to: The returns are just not high enough.

This paper reports on a randomized evaluation of an intervention in urban India focused on improving the learning environment in public schools. The intervention is motivated by the belief that children often drop out because they fall behind and feel lost in class. The program, which is run Pratham, a Mumbai-based Non-Governmental Organization, provides remedial education, in small groups, to children that are lagging behind. To keep costs low and ensure a good instructor-student relationship, the program hires young women (the "Balsakhis") who have the equivalent of a high school degree from the local slum communities in which the schools are located.

The evaluation of the remedial education (Balsakhi) program offered an opportunity to implement an evaluation design that is often recommended but rarely, if ever, utilised.[3] First, it was a randomized evaluation. We can therefore be relatively confident of the absence of confounding factors.[4] Second, the program we study was run on a very large-scale (over 15,000

[2]The one exception we of we are aware is the study of a program that provides incentives for teachers in Kenya that is reported in Glewwe, Kremer and Moulin (2002b), though even in this case the authors seem to be somewhat disappointed by the lack of a more robust impact. Chin (2001) finds that Operation Blackboard in India did increase school completion rates for girls, which implies that there must have been an increase in test scores, but she cannot tell whether those who would have completed school in any case learn more as a result of the intervention. Vermeersch (2002) also finds an impact on test scores of a school meals program in schools where the teachers were trained, but she too cannot distinguish between those who were already in school and the newcomers.

[3]We are currently in the middle of a two-year evaluation of a Computer Assisted Learning program in Baroda whose research design is very similar to the one described here. However, we are unaware of any other evaluation of an educational program that meets the criteria listed below.

[4]In Baroda, the Computer Aided Learning (CAL) program was implemented in the second year of the Balaskhi

students were included in the study), and had already clearly demonstrated the ability to scale up in other cities, as the description below will make clear. In other words, there is no risk that what we are evaluating cannot be reproduced elsewhere. Third, we simultaneously carried out randomized evaluations of the program in two different cities, each of which had its own management team. This reinforces our confidence in the external validity of these results. Finally, we conducted the study over two years, using several tests, making it less likely that the results are a consequence of the newness of the program, or the effect of implementing an evaluation.

Finally, though we find no effect on attendance, we find that the program has a substantial positive effect on children's academic achievement. This impact is remarkably stable across years and cities, especially when we take into account the instability of the environment—there was a major riot and a catastrophic earthquake while the program was running. Moreover, the weaker students, who are the primary target of the program, gained the most. This study demonstrates both the efficacy of the remedial education program, and more generally, the feasibility of dramatically impacting test scores at very low cost. We also make an attempt to distinguish the direct effect of the program (on children who worked with the balsakhi) and the indirect (on those who did not). The estimates suggest that the reducing class size by hiring a balsakhi is at least twice as effective as reducing class size by keeping children with regular teachers.

2 The Programs

2.1 Remedial Education: The Balsakhi Program

Pratham was established in Mumbai in 1994, with support from UNICEF, and has since expanded to several other cities in India. Pratham now reaches over 220,000 children in 34 cities in India, and employs about 10,000 individuals. Pratham works closely with the government: Most of its programs are conducted in the municipal schools, and Pratham also provides technical

study, for fourth standard children only. However, the experimental design randomly assigned the CALprogram to half of the balsakhi-treatment and half of the balsakhi-comparison schools, thus allowing us to estimate unbiased balsakhi-only effects. The results of the CAL program will be reported once the study is complete.

assistance to the government.

One of Pratham's core programs is a remedial education program, called the Balsakhi program. This program, in place in many municipal schools, provides a teacher (usually a young woman, recruited from the local community, who has herself finished secondary school) to work with children identified as falling behind their peers. While the exact details vary according to local conditions, the typical instructor meets with a group of approximately 15-20 children in the morning for two hours, and with another group of the same size in the afternoon. Instruction focuses on the core competencies the children should have learned in the second and third standards, primarily basic numeracy and literacy skills. The instructors are provided with a standardized curriculum that was developed by Pratham. They receive two weeks of training at the beginning of the year and ongoing reinforcement while school is in session. The program has been implemented in twenty Indian cities, reaching tens of thousands of students. It was started in Mumbai in 1994, and expanded to Vadodara in 1999.

According to Pratham, the main benefit of the program is to provide individualized, non-threatening attention to children who are lagging behind in the classroom and are not capable of following the standard curriculum. Children may feel more comfortable with women from their own communities than teachers, who are often from different backgrounds. As the balsakhi's class size is relatively small, she may tailor the curriculum to the children's specific needs. Furthermore, because Pratham's program takes children out of the classroom, it may even benefit children who were not directly targeted by the intervention. Removing children from the classroom for two hours means the effective student-teacher ratio in the main classroom drops, and the teacher may be able to focus on more advanced material. Finally, if the balsakhis are indeed effective, even when the children are returned to the main classroom, the teacher may not need to keep re-teaching remedial material.

An important characteristic of this program is the ease which with it can be scaled up. Because Pratham relies on local personnel, trained for a short period of time, the program is very low-cost (each teacher is paid 500-750 rupees, or 10-15 dollars, per month) and is easily replicated. There is rapid turnover among the balsakhis (each of them staying for an average of one year, typically until they get married or get another job), indicating that the success of the program does not depend on a handful of very determined and enthusiastic individuals.

Finally, since the balsakhis use whatever space is available (free classrooms, or even hallways when necessary), the program has very low overhead and capital costs.

3 Evaluation Design

3.1 Sample: Vadodara

In 2000, when Pratham decided to expand their remedial education (balsakhi) program to cover the entire city of Vadodara, they decided to take advantage of the expansion to evaluate the effectiveness of the program in the remaining 98 eligible schools in the city. In November, 2000, they administered an academic test (designed by the Pratham team) to all children in the third standard. They then hired and trained balsakhis, which were sent to half of the schools in Vadodara. Assignment was random, with schools stratified by medium of instruction, gender, and pupil-teacher ratios. Unfortunately, the school year was disrupted by an earthquake in Gujarat, and children received only a few weeks of instruction between November and March. This year of the program is best understood as a pilot program.[5]

In July, 2001, the group of schools that had received a balsakhi in the previous year of the program received the balsakhi in the fourth standard, and the remaining schools received a balsakhi in the third standard. Children in the standard that did not receive the balsakhi in a given grade form the comparison group for children who did receive the balsakhi.

The program was continued during the school year 2002-2003, with the addition of the 25 remaining primary schools. Schools where the balsakhi was assigned in standard three in the year 2001-2002 were now assigned a balsakhi in standard four, so that in year 2, standard 4 children in the treatment group benefitted from two years of the balsakhi program. Schools where the balsakhi was assigned in standard four in the year 1 received balsakhi assistance for standard three in year 2. The new schools were randomly assigned to either group with equal probability in the same way that the original schools were assigned. The number of schools and divisions in the two groups are given in Table 1.

[5] Throughout the paper, we will refer to the year 2001-2002 as "year 1" and 2002-2003 as "year 2."

3.2 Sample: Mumbai

To ensure the results from the Vadodara study would be generalizable, the Balsakhi program in Mumbai was also evaluated, in 2001-2002 and 2002-2003. Mumbai was Pratham's birthplace, and Pratham is currently operating various programs throughout the city. We selected one ward (the L-ward) to implement a design similar to the design in Vadodara, including all Gujarati, Hindi, and Marathi schools. In total, 62 schools are included in the study. Schools were stratified according to their scores in a pre-test, as well as by the medium of instruction. Half the schools were randomly selected to receive a balsakhi in standard two, and half the schools were randomly selected to receive a balsakhi in standard three. In 2001-2002, data were collected only for standard three children, while in 2002-2003, data were collected for standards three and four. As in Vadodara, children kept their treatment assignment status as the moved from standard two to three (or three to four).

In the second year of the study, the Mumbai program experienced some administrative difficulties. A decision to require balsakhis to pass a competency test resulted in the firing of many balsakhis. Hiring new recruits was complicated by the fact that the administrative staff in L-Ward turned over between year 1 and year 2, and the new staff lacked community contacts necessary for recruitment. Finally, the principals of a couple of schools, hearing that the study was being conducted by a group of Americans, refused balsakhis. Thus, only two thirds of the schools assigned Balsakhis actually received them. (Schools could not refuse testing, because Pratham had obtained written permission for testing from the city administration). Throughout the paper, the schools that were assigned balsakhis but did not get them are included in the treatment group.

3.3 Outcomes

The main outcome of interest is whether the interventions resulted in any improvement in cognitive skills.

In the Vadodara pilot year, children were given a pretest in November, 2000, and post-test in March, 2001. In the first full year, the Vadodara pretest was at the beginning of the school year (August 2001), and the post test in March 2002. In the second full year, children were tested at the beginning of the school year (August 2002), in November 2002, and again in March, 2003. In the first year in Mumbai, children were tested in October, 2001 and March, 2002; in the second year tests were given in August, 2002, and February 2003.

In Vadodara, the same test is used for standard three and four children, so that the scores can be directly compared across grades. Scores on the pre- and post-test can also be directly compared, as the format of the questions and the competencies tested remain the same. The exam comprises two parts: A math section and a language section. In Vadodara, both parts focused on competencies that the Vadodara Municipal Corporation (VMC) prescribe for children in standards one through four. On the math exam, for example, tasks ranged from basic number recognition, counting, and ordering of single digit numbers to ordering of two digit numbers, addition of single and two digit numbers, and basic word problems. Tests were similar in Mumbai. In the first year, tests focused on competencies in standards one through three, while in the second year they included standards one through four. In the second year, the same test was used for third and fourth standard children.

The "pilot" year of the program (2000-2001) allowed Pratham to make significant progress in developing a testing instrument (the initial test was too difficult) and effective testing procedures to prevent cheating and exam anxiety. The test was administered in both cities by Pratham, with the authorization of the municipal corporation. At least three Pratham employees were present in the classroom during each test to minimize cheating.[6] To minimize attrition, Pratham returns to the schools multiple times, and children who still failed to appear and who could be tracked down were administered a make-up test outside of school.

In Vadodara, the school year 2001-2002 was disturbed by massive inter-communal riots following an attack on a train carrying Hindus. Although a post-test was conducted in March (after the riots had receded), attrition was high. We thus use the October 2001 mid-test as the "post-test" for year one in Vadodara.[7] The year one pretests were in August (Vadodara) and September (Mumbai) 2001, while the Mumbai year one post-test was in March 2001. In year 2, the pretests were in August 2002 (both cities), and the post-tests in February (Bombay) and

[6] In Mumbai, since administration of the pre-test was less than satisfactory at the first attempt, we conducted a second pre-test, which we use as the basis for the analysis.

[7] The results of the first year of the program do not significantly change if we use the mid- or post-test: There was no further improvement (or deterioration) of the performance in the treatment schools relative to the control schools between the mid-test and the post-test.

March (Vadodara) 2003.

Another outcome of interest is attendance and school dropout rates, which are collected weekly by Pratham employees, who made randomly timed weekly appearances in classrooms to take attendance. (Data from the official rolls was also collected, but administrators have incentives to inflate the attendance data). [8]

Finally, in the second year of the program, in both cities, data were collected on which specific children were sent to the Balsakhi. (Balsakhis work with, on average, about 20 children per school).

3.4 Statistical Framework

3.4.1 Effect of the Balsakhi Program on test scores

Given the randomized allocation of both programs, we expect the 2001 pre-test results in the treatment schools to be similar between those in the control. The results of the 2002 pre-test may be different in the treatment and control schools in standard four in Vadodara, as well as standard three and four in Mumbai, since they may reflect long-lasting benefits of the previous year's program for the children who were in the same school in the previous year. In both cities, the experimental design (in which each school was both in the treatment and comparison group, with one standard in each group) is such that even if a "good school" were in the treatment group for a given standard, the other standard of that "good school" would be in the comparison group, ensuring that the averages across the standard are likely to be very similar.

Denoting y_{igjk} the test score of child i in grade g in school j in test k (k is either "PRE" or "POST"), we start by comparing test scores in the treatment and comparison schools, in each city and standard.

We start by checking that there is no difference between treatment and control schools before the program was run:

$$y_{igjPRE} = \alpha + \beta D_{jg} + \epsilon_{igjPRE}, \tag{1}$$

where D_{jg} is a dummy indicating whether school j is in the treatment group in that particular year in standard g, and ϵ_{igjPRE} the error term.

[8] We report results from Mumbai in this draft. In Mumbai 2001-2002, teachers in some schools often refused to let the research assistants count the number of children present, resulting in biased data.

This regression is run separately in each standard, year and city. It is run separately for the math exam, the verbal exam, and the total score on the exam. The standard errors are clustered at the school level.[9]

We then run the same regression in the post-period ($k = POST$):

$$y_{igjPOST} = \alpha + \beta D_{jg} + \epsilon_{igjPOST}. \tag{2}$$

This provides a first estimate of the effect of being assigned to the treatment group. For all cities and year, except for Mumbai in year 2, this is also an estimate of the average effect of being a student in a school that was assigned a balsakhi. However, in Mumbai in year 2, because not all schools received a balsakhi (and not all classes within schools where treated), to obtain the average effect of receiving a balsakhi, we use the assignment to the treatment group (D_{jg}) as an instrument for whether or not the class of a specific child actually received the balsakhi (B_{jg}). In practice, we are estimate the following equation:

$$y_{igjPOSTt} = \alpha + \beta B_{jg} + \epsilon_{igjPOST}. \tag{3}$$

The first stage is the equation for whether or not a child's class was actually assigned a balsakhi:

$$B_{jg} = \alpha_1 + \delta_1 D_{jg} + \eta_{igjPOST}. \tag{4}$$

Because tests scores are very strongly auto-correlated, the precision of the estimate is increased by controlling for the child's test score in the pre-test.

We do so using the following difference in difference specification:

$$y_{igjk} = \lambda + \delta B_{jg} + \theta POST_k + \gamma(B_{ig} * POST) + \epsilon_{igjk}, , \tag{5}$$

where $POST_k$ is a dummy indicating whether the test is the post test. For all years and samples except Mumbai in year 2 $B_{jg} = D_{jg}$, and equation 5 is estimated with OLS. However, for

[9] If we instead use a nested random effects model (with a classroom effect nested within a school effect), the point estimates are very similar, and the the standard errors are smaller. Clustering is a more conservative approach.

Mumbai in year two (and when both cities are pooled), equation 5 is estimated with instrumental variables, with D_{jg} (the initial assignment to the treatment group), $POST$, and $D_{jg} * POST$ used as instruments.

We also present an alternative way to estimate the treatment effect in Mumbai, as a specification check. Since every school was supposed to receive a balsakhi in either standard 3 or standard 4, we keep in the sample only the schools that did receive a balsakhi. This means that a school will not be in the comparison school for one standard if the other standard did not receive a balsakhi. In this reduced sample, B_{jg} is equal to D_{jg}, and equation 5 is estimated by OLS. The assumption underlying this specification is that the characteristics that make the school more likely to have a balsakhi have the same influence on the test scores of children in standard 3 and standard 4.

To gain more insight about the impact of the program, we also present estimates of specifications similar to equations 3 and 5 using for y_{igjk} a binary variable indicating whether the child correctly answered the questions indicating competencies for standard 1, 2 and 3, respectively. Finally, we estimate the impact of being in the program for 2 years (for children who were in the treatment group in standard 3 in year 1, and whom the balsakhi has followed in year 2 when they moved to standard 4), by estimating equation 5 using the pre-test of year 1 as the pre-test, and the post test of year 2 as the post test.

3.4.2 Disentangling Class size and balsakhi effect

Estimating equations 3 and 5 generates estimates of the average impact of the program on all children who whose standard-school received a balsakhi. The program may impact the children in a treated school in two ways: directly, for children who were assigned to work with the balsakhi, or indirectly, because the weakest children are removed from the classroom for part of the day. This indirect effect can potentially work through two mechanisms: through a reduced number of students in the class (class size effect), and through the higher average quality of their classmates (tracking effect).

To separate the balsakhi and the indirect effects, an ideal experiment would have identified the children who would work with the balsakhi in all schools, *before* randomly assigning treatment and comparison groups (and to not allow substitution after the initial allocation). The balsakhi effect could then be estimated by comparing children at risk of working with the balsakhi in the treatment and the comparison group. The indirect effect would have been estimated by comparing the children who were not at risk of working with the balsakhi in the treatment and the comparison group. Unfortunately, this design was not practical in this setting.

We do know, however, that the assignment to the balsakhi group was based in part on pre-test score, and that a maximum of twenty children per school in Vadodara, and twenty per class in Mumbai were assigned to a balsakhi. In schools in the treatment group, we start by predicting assignment to the balsakhi as a function of the number of students (in the school in Vadodara, in the class in Mumbai), the sum of the math and verbal score at the pre-test, and a variable indicating whether the child is among the bottom 20 children in his group.

$$P_{ijg} = \pi_1 + \pi_2 S_{ijg} + \pi_3 y_{ijgPRE} + \pi_4 R_{ijg} + \pi_5 Z_{ijg} + \omega_{ijg} \tag{6}$$

where S_{ijg} is the number of student in the class or the school, y_{ijgPRE} is the score of the child at the pre-test, R_{ijg} is the rank of the child in the class (starting from the bottom), and Z_{ijg} is a dummy indicating whether the child is among the bottom 20 children in the class. We will show that, even after controlling linearly for the class rank, the dummy Z_{ijg} predicts whether or not the child was assigned to the balsakhi.

Denoting X_{ijg} the vector $[S_{ijg} \quad y_{ijgPRE} \quad R_{ijg}]$, the following equation (which interacts the variables in equation 6 with a dummy for whether the child is in the balsakhi group) predicts assignment to the balsakhi in the whole sample.

$$P_{ijg} = \alpha + \gamma D_{ijg} + \beta \left(Z_{ijg} * D_{ijg}\right) + \mu Z_{ijg} + X'_{ijg}\kappa + \lambda \left(X'_{ijg} * D_{ijg}\right) + \epsilon_{ijg} \tag{7}$$

We can then regress the post test scores on the same variables, and examine whether being one of the bottom 20 children is associated with a bigger effect for those in the balsakhi group:

$$y_{ijgPOST} - y_{ijgPRE} = \alpha + \beta Z_{ijg} * D_{ijg} + \gamma D_{ijg} + \mu Z_{ijg} + X'_{ijg}\kappa + X'_{ijg} * D_{ijg}\lambda + \epsilon ijg \tag{8}$$

Equation 8 and 7 form the first stage and the reduced form of a instrumental variables estimation of the following equation:

$$y_{ijgPOST} - y_{ijgPRE} = \alpha + \beta P_{ijg} + \gamma B_{ijg} + \mu Z_{ijg} + X'_{ijg}\kappa + X'_{ijg} * D_{ijg}\lambda + \epsilon_{ijg} \quad (9)$$

where D_{ijg} and $Z_{ijg} * D_{ijg}$ are the excluded instruments. The identification assumption underlying this estimation strategy is that the only reason why the treatment effect varies with the variable Z_{ijg} is because Z_{ijg} makes it more likely that the child is sent to the balsakhi group. However, the effect of the treatment is allowed to vary with class size, the test score, and the rank of the child. We also estimate an alternative specification which controls for a fourth-order polynomial in the rank of the child. In this equation, the effect of being assigned to the balsakhi group is given by $\beta + \gamma$, and the effect of being in a balsakhi school, but not assigned to the balsakhi group, is given by γ.

3.5 Disentangling tracking and class size effects

The effect of the program on children who were not assigned to the balsakhi can be further attributed to either a class size effect or a tracking effect: by removing the 20 weakest children in the class, the program might allow the teacher to spend more time on more advanced topics, or less time reviewing or disciplining children who do not follow the class.

To gain some insight on whether the two effects can be separated, we estimate an equation where we interact the treatment dummy with the average pre-test score of children who were sent to the balsakhi: if the test scores of children who go to the Balskahi are lower, the tracking benefits to the remaining children should be bigger. To instrument for the test score of the children who went to the Balsakhi, we use the average test score of the bottom twenty children in the class.

The reduced form equation is thus:

$$y_{ijgPOST} = \alpha + \beta Z_{ijg} * D_{ijg} + \gamma D_{ijg} + \delta D_{ijg} * \overline{T20}_{jg} + \mu Z_{ijg} + X'_{ijg}\kappa + X'_{ijg} * D_{ijg}\lambda + \epsilon ijg, \quad (10)$$

where $\overline{T20}_{jg}$ is the average test score among the bottom twenty children of school j.

The structural equation is:

$$y_{ijgPOST} = \alpha + \beta P_{ijg} + \gamma B_{ijg} + \delta B_{ijg} * \overline{TB}_{jg} + \mu Z_{ijg} + X'_{ijg}\kappa + X'_{ijg} * D_{ijg}\lambda + \epsilon ijg \quad (11)$$

where $\overline{TB}_{jg}$ is the average score among children who are assigned to the balsakhi.

4 Results

4.1 Descriptive Statistics: Level of Competencies and Pre-intervention Differences

Tables 2 through 5 present the descriptive statistics of the test scores for all samples used in this analysis (year 1 and 2 in Vadodara and Mumbai). The scores are presented in three different ways: percentage of points scored, normalized scores (relative to the distribution of the pre-test score in the comparison group in each city and year)[10], and as the percentage of children correctly answering the questions in the test relating to the competencies in each standard.

The randomization appears to have been successful: Neither in Mumbai nor in Vadodara are there any large or systematic differences between the pre-test score and the post-test score. None of the differences between the groups prior to the implementation of the program are significant.

The raw scores, and the percentage of children correctly answering the questions relating to the curriculum in each standard give an idea of how little these children actually know. In standard three in Vadodara in the second year, for example, the average student in math scores about 16%, both in the control and treatment groups. Since one math question is multiple-choice, on average a student who knows nothing will score 1.8% points. If a student can consistently order two numbers and add two single digit numbers, she earns the additional 14% needed to achieve the average third standard performance. Only 5.4% of third standard children in Vadodara pass the standard 1 competencies in maths in standard 3 in Vadodara (and 14% in Mumbai). Standard one competencies cover number recognition, counting and one digit addition and subtraction.

The results are more encouraging in verbal competencies: 50% of the standard 3 children pass the standard 1 competencies in Vadodara (reading a single word, choosing the right spelling

[10] Scores are normalized for each standard, year, and city, such that the mean and standard deviation of the comparision group is zero and one, respectively. (We subtract the mean of the control group in the pre-test, and divide by the standard deviation.) This allows for comparison across samples, as well with results from other studies.

among different possible spelling for a word), and 65% do so in Mumbai.

4.2 Attrition

Table 6 presents the levels of attrition in Mumbai and Vadodara. We present attrition that occurred between the pre-test and post-test for both cities in both years, as well as the two-year attrition (in Mumbai, for standard 4 only), broken down by treatment status. Attrition was generally very low, except for Vadodara in year 1. This is likely attributable to the civil unrest. The post-test was conducted after the riots, and the research team attempted to track down all of the children who did not appear for the exam. Attrition rates are not different in the comparison group than in the treatment group: In year 1 in Vadodara, attrition was 19% in the balsakhi treatment group, and 18% in the comparison group. In year 2, attrition was 4% in the balsakhi in both the treatment and the comparison group. In Mumbai year 1, attrition was 7% in the treatment group, and 7.5% in the comparison group, while in year 2 it was 7.7% in the treatment group and 7.3% in the comparison group.

The fact that there was no differential attrition rate in the treatment and control groups suggests that the estimate of the treatment effects will not be biased, unless different types of people drop out from the sample in the treatment and the control groups (Angrist, 1995). This does not seem to occur in our study: The second row in each panel presents the difference between the score at the pre-test of children who were not present at the post-test, by treatment status. The third column of each sample group present the differences-in-differences in the treatment and comparison groups. Children who will eventually leave the sample tend to be at the bottom of the distribution of the pre-test scores. However, the difference is very similar in the treatment and control groups, in both years and for both programs. In Mumbai in the second year, the difference-in-difference is almost statistically significant (p-value .067), with the attritors in the treatment group seeming to perform relatively better than those in the comparison group; this could be because the program encouraged weaker children to stay in school, making it more likely that they could be tracked down for the post-test. If anything, this would bias the estimates of the treatment effect downward.

Finally, both the attrition and the difference in test scores are also similar among the bottom 20 children in each school, the group of children who were the most likely to be assigned to a balsakhi (these results are not reported).

4.3 Effects of the Balsakhi Program

Effect on attendance

Part of the goal of the program was to make it easier for parents to play a role in their children's education, by serving as an intermediary between parents and the school environment. One could therefore have expected an impact of the program on attendance. In practice, there does not seem to be any: table 11 show two estimates of the program on attendance: a simple difference (comparing average attendance in the treatment and the comparison schools) and a difference in difference estimate (where we compare the change in attendance between the month of September-October and January-February in the treatment and the comparison group). Using either specification (and either the official roster data or data collected by our assistant) we do not find any impact of the program on children's attendance.

Overall effect on test score

Tables 2, 3, 4, and 5 present the first estimates of the effect of the balsakhi program, as simple differences between the post-test scores in the treatment and control groups. In all years and standards, for both tests, and in both cities, and for all subgroups, the difference in post-test scores between treatment and control groups is positive. In the first year in Vadodara (table 2), the difference in post-test score between treatment and control groups was 0.17 standard deviations in standard three for math, .15 in standard 3 for language, and .14 and .06 in standard 4, for math and language respectively.[11] The results in Mumbai (table 4) are remarkably similar, with the math and language test scores improving by 0.15 and 0.16 standard deviation, respectively.

In the second year of the program, the effects are larger: In Vadodara (table 3), the difference in total test scores is .44 for math and 0.25 for language in standard three, and .33 and 0.30 in standard four, for math and language respectively. In Mumbai in year two (table 5), the IV estimate of the impact of the program on test scores differences in are .25 and .09 in standard 3 (for math and language respectively) an .43 and .17 in standard 4 (for math and language

[11] Throughout the paper, test results, differences, and differences-in-differences are presented in terms of standard deviations, unless otherwise specified.

respectively). In year two in Vadodara, all of the differences between treatment and control groups are statistically significant, while for Mumbai, the standard four results are significant.

Because test scores have a strong persistent component, the precision of these estimates can be improved significantly, however, by turning to a differences-in-differences specification (equation 5). Since the randomization appeared to be successful, and attrition was low in both the treatment and comparison groups, the point estimates should be similar in the simple differences and the differences-in-differences specification. Table 7 presents differences-in-differences estimates of the effects of the balsakhi program, in various years, cities, standards, and sub-groups. For Mumbai in year 2, we estimate the treatment effect in two ways: first, we instrument for the dummy indicating whether or not the school received a balsakhi with a dummy for whether the school was assigned to the treatment group; second, we include only schools that got a balsakhi in at least one standard in the sample. The estimates using either specification are very similar.

As expected, the point estimates suggest a substantial treatment effect, and the standard errors are lower than the simple differences. Pulling both cities and standard together (in the first two rows of table 7), the impact of the program was 0.14 standard deviation overall in the first year, and 0.25 standard deviation in the second year. Both estimates for total score are significant at the 99% confidence level.

The impact is bigger in the second year, and bigger for math than for language in both years (0.19 standard deviations versus 0.069 in the first year, and 0.32 versus 0.15 standard deviations in the second year; all but first-year verbal scores are significant at the 99% level.) For both years and both subjects, the effect is larger in Vadodara than in Mumbai (with a total-score effect of 0.16 standard deviation versus 0.15 in the first year (standard 3 only), and 0.31 versus 0.15 in the second year (both standards)). The difference is the strongest for language, where there is a significant impact in both years for Vadodara (0.11 and 0.23 standard deviation respectively), but no significant impact in either year in Mumbai (0.06 standard deviation for standard 3 in year 1, and 0.032 standard deviation in year 2). For both cities and both subjects, the effects are very similar in standard 3 and standard 4. We also computed all those estimates for both genders separately, and found the impact to be very similar (results not reported).

In the last panel of the table, we display our estimate of the impact of the program for two years (for children who were in a treatment school in standard 3, and stayed in the treatment school). First, it appears that the effect of the first year does not seem to persist over the summer: at the pre-test in year 2, children who were in a treatment class in year 1 do not seem to know more than children who weren't. However, the effect of two years of treatment (from year1 pre-test score to year 2 post-test score) is substantially larger than that of either individual year (0.52 standard deviation in math, for example, versus 0.38 for one year): it seems likely that the foundation laid in the first year of the program helped the children benefit from the second year of the program.

Compared to the other educational interventions, this program thus appear to be quite effective. The Tennessee STAR experiment, for example, where class size was reduced by 7 to 8 children (from 22 to about 15), improved test scores by about 0.21 standard deviation. This program improved test scores by 0.25 standard deviations in the second year, by reducing effective class size from 35 to 20 children on average (averaging over the balsakhi and the non-balsakhi group) for part of the day, but doing so by hiring an assistant paid a fraction of the teacher's salary.

Effect on specific competencies

The test was designed to cover competencies from standard 1 to 3. In Mumbai in year 2, it also covered some standard 4 competencies. Table 8 offers more details on the level at which the program was effective. Estimates in this table suggest that, for math, the biggest effect was on the competencies of standard 1: in Vadodara for example, the program increased the fraction of children who mastered the competencies of the first standard in math by 5.0% in the first year, and 6.4% in the second year. In Mumbai the effect was 5.8% and 8.8% respectively. The effect on the fraction of children demonstrating knowledge of standard 2 and 3 competencies is much smaller. In language, the most important effect seems to be to help children master the competencies of standard 2 This may not be surprising, since many children seemed to have already mastered the competencies of standard 1. The effect of the program may thus be the strongest on the easiest competencies not already mastered by many pupils. These results correspond well with the stated role of the program, which was to work with children on basic competencies.

Distributional effects

The balsakhi program was primarily intended to help children at the lower end of the ability distribution, by providing especially targeted instruction. However, as we already mentioned, it could still help the higher scoring children, either because they are assigned to the balsakhi, or because they benefit from smaller classes when their classmates are with the balsakhi. In table 9, we directly ask whom the program benefitted, by breaking down the sample into thirds in the distribution of initial test scores. The effect of the program is indeed the strongest for children in the bottom of the pre-test score distribution. Taking both cities together, the impact in the first year on overall test scores was 0.21 in the bottom third, 0.12 in the middle third, and 0.08 in the top third. In the second year, the impact was respectively 0.32, 0.18 and -0.003. Children in the bottom of the distribution were also much more likely to see the balsakhi: 22% from the bottom third, compared to 16% from the middle third and 6% from the top third were sent to the balsakhi. By increasing scores of the lowest-scoring children more than their peers, the program has an equalizing effect on pupil achievement. It is worth noting that this occurs even in the Indian education system, which, like many in developing countries, places particular emphasis on the children who perform well The effect of the program on the bottom third of the distribution are impressive: in the second year, the math test scores increased by 0.46 standard deviation in Vadodara, and 0.51 in Mumbai. The language test score increased by 0.31 standard deviation in Vadodara, and 0.13 in Mumbai.

4.4 Inside the box: Balsakhi, class size, and tracking effects

These results lead to our next question: to what extent is the program effect due to a direct effect of the balsakhi teacher (affecting only the children who got assigned to the balsakhi group) and to an indirect effect, affecting children who were *not* assigned to the balsakhi group. The fact that both the program impact and the probability of being assigned to a balsakhi declines with a child's position in the test score distribution suggest that the impact of the program may have been larger for those who were actually assigned to the balsakhi. However, an alternative plausible explanation for this patter is that the direct (or indirect) effects of the test score are lower for children with higher test scores. This question therefore necessitates further investigation.

As we explained above, we propose using a dummy for whether a child belongs to the bottom 20 children of a class as an instrument for whether he is assigned to the balsakhi group.

Columns 1 to 4 in table 10 show that in both Mumbai and in Vadodara, a dummy for whether a child belongs to the bottom 20 in his class predicts his assignment to the balsakhi, even after controlling for her rank, her score at the pre-test, and the number of students in her class (which are all negatively and significantly associated with assignment to the balsakhi). Not surprisingly, because some schools in Bombay were not assigned a balsakhi, all coefficients are smaller. In column 4, we include a fourth-order polynomial of rank (interacted with treatment); the dummy for child rank below 20 remains a significant predictor of assignment to the balsakhi group. In columns 5 to 8, we present the reduced form estimates for test score gain. The coefficient of the interaction between the dummy for belonging to the bottom 20 children in the class and belonging to a treatment school is significant in all of these columns, which indicates that, conditional on being in a school assigned to the treatment group, the treatment effect is actually bigger if the child is more likely to be assigned to the balsakhi. The treatment effect also appears to be declining with initial test score and the number of students in the class, though, controlling for the rank of the child, this effect is not very large.

In panel B of table 10, we present instrumental variables estimates of the direct and indirect impact of being in a balsakhi group. The direct effect is the sum of the coefficient on the dummy for whether the school gets a balsakhi and on the dummy for whether the child is assigned to the balsakhi group. In both Mumbai and Vadodara, the coefficient for whether or not the child is assigned to the balsakhi is positive, significant, and large (0.89 standard deviations in Vadodara, and 1.3 standard deviations in Mumbai.) This suggests that the children actually assigned to the balsakhi benefit much more than others in their school. In Mumbai, the treatment effect on those who were not sent to the balsakhi is negligible. In Vadodara, it is 0.21 standard deviations. Unfortunately, the estimate is too noisy to be distinguished from zero. Pooling both sample, we find an effect of 0.23 for children not assigned to the balsakhi in treatment school, and an additional 0.95 standard deviation for children who are assigned.

Clearly, while there is some (not very conclusive) evidence of an indirect effect for children who remained in class, the main reason the program is effective seems to be that it delivers more effective education to the children who were actually with the balsakhi. Part of this is probably due to the fact that the children assigned to work with the balsakhi ended up in much smaller classes than those who weren't, given the initial large class sizes in both cities. The reduction in class size was somewhat larger for children who were assigned to the balsakhi, because they were moved to classes whose average size was 20. However, if one believes the point estimate, the effect of the program for children assigned to the balsakhi group is more than five time the size of the effect for other children, which suggest that the balsakhi is a more effective teacher for these children.

In Vadodara, a mid-test was conducted, as well as a pre- and post- test. We conduct the analysis for both tests separately. It appears that, during the period between the pre and the mid test, both the balsakhi and the non-balskahi children benefitted equally from the program (the balsakhi dummy is not significant). Between the mid-test and the post-test, however, only children allocated to the balsakhi group benefitted.

Finally, we examine whether there is any evidence of a positive (or negative) impact of tracking by interacting the dummy for being in the treatment group with the average score of the bottom 20 children in the school. In Vadodara pre to mid-test, the effect of being in a balsakhi school for children who were not allocated to the balsakhi group declines with the initial test score of the children, which is consistent with a positive impact of tracking.

5 Cost Benefit Analysis

In seeking to improve the academic performance of schoolchildren, governments could potentially hire additional teachers, or hire balsakhis. Data on the cost of teachers, combined with the results presented above, give an idea of the cost effectiveness of each option. Table 12 shows the cost per student per year of the balsakhi program. While the data available from the municipalities about the cost of the program in Vadodara and Mumbai are not directly comparable to each other (for Vadodara we have the total cost of schoolteachers, while for Mumbai we know the starting salary of a new teacher), we use measures of the Pratham program cost that are comparable within cities.[12] The program cost 91 rupees, (about 2 dollars) per student [13] in Vadodara, and 54 Rupees (or 1 dollar) per child in Mumbai, while the teacher salary cost per student of the

[12] We also have cost data from Vadodara in 2001-2002, and from Mumbai in 2002-2003. As costs did not change appreciably from one year to the next, we use these figures for cost-benefit comparisons for both years.

[13] The denominator includes all students in standards three and four in the treatment schools, since we will also use average test scores.

Vadodara and Mumbai Municipal Corporations are 3,168 Rupees and 1318 Rupees, respectively. (The cost per student/year increased in Mumbai in year 2 because the wage of the balsakhi was increased from 500 to 750 Rs./month).

Table 13 combines these numbers with the test score improvements over the pre- to post-test in years 1 and 2. Since the post-test used the same structure as the pre-test, improvement of the control children between the pre- and the post-test provides a measure of the effect of being in school for one year, which can be compared with the effect of having a balsakhi over the same period. Clearly, this is only suggestive, since children's competency over a year may change for reasons other than being in school. In year 1 in Vadodara, the estimate of the treatment effect was about 40% as large as the improvement between the pre-test and the post-test in the comparison group. The ratio of the cost per child is however 35. This calculation suggests that the average balsakhi is about 12-16 times more cost effective (in terms of improvement in test scores) than the average teacher. In the second year, the balsakhi program appears to be 11-12 times more cost effective than the average teacher. The results in Mumbai also suggest the Balsakhi program is dramatically more cost effective. In the first year, the treatment effect was half as large as the gains made by the comparison group, while the cost ratio was 24, suggesting Balsakhis are thirteen times more cost effective than teachers. In the second year, the treatment effect was $\frac{1}{10}$ (respectively $\frac{1}{4}$) as large in the third (fourth) standard as the comparison groups gains, giving a cost advantage of 2 in the third standard, and of six in the fourth standard. It is important to note that this results *do not* suggest that teachers should be replaced by balsakhis, since balsakhis are always complementing the teachers. The results do provide evidence that if the Mumbai and Vadodara Municipalities wanted to spend additional resources, hiring balsakhis may be a more effective way to do it than hiring additional teachers.

6 Conclusion

This paper reports the preliminary results of a remedial education program. The program has already shown that it can be brought to scale, since it is already reaching tens of thousands of children across India. Evaluations conducted in two cities over two years suggest that this is a remarkably effective and cost effective program: Test scores of children whose schools benefited from the program improved by .12 to 0.16 standard deviations in the first year, and .15 to 0.3 standard deviations in the second year. At the margin, the program is up to 12-16 times more effective than resources spent on teachers. Results are even stronger for children in the bottom of the distribution (in the bottom third of the distribution, the program improved tests score by 0.22 standard deviations in the first year, and 0.58 in the second year).

References

Angrist, Joshua D. (1995) "Conditioning on the Probability of Selection to Control Selection Bias," NBER Technical Paper #181

Angrist, Joshua, et. al. "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment." American Economic Review 92 (2002): 1535-1558

Banerjee, Abhijit, Suraj Jacob and Michael Kremer. "Promoting School Participation in Rural Rajasthan: Results from Some Prospective Trials." MIT Department of Economics Working Paper (2002): .

Chin, Amiee (2002) "The Returns to School Quality When School Quality is Very Low: Evidence from Operation Blackboard in India." University of Houston

Duflo, Esther, (2001). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." American Economic Review 91 (2001): 795-813

Glewwe, Paul, Michael Kremer and Sylvie Moulin (1997) "Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya." Manuscript, Harvard University Department of Economics

Glewwe, Paul, Michael Kremer, Sylvie Moulin and Eric Zitzewitz. (2003) "Retrospective vs. Prospective Analyses of School Inputs: The Case of Flip Charts in Kenya." Journal of Development Economics, forthcoming

Kremer, Michael and Ted Miguel (2003) "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." Econometrica, forthcoming

Kremer, Miguel, and Thornton (2002) "Merit Scholarships," manuscript, Harvard Economics.

Spohr, Chirstopher. (2003) "Formal schooling and workforce participation in a rapidly developing economy: evidence from 'compulsory' junior high school in Taiwan." Journal of Development Economics 70: 291-321

Vermeersch, Christel (2002) "School Meals and Educational Achievement: Evidence from a Randomized Evaluation,." Manuscript, Harvard University Department of Economics

Table 1: Sample Design

	Standard	Study Group	Number of Schools	Number of Divisions	Number of Children	Number of Schools Treated	Number of Children With Balsakhi
Vadodara							
Year 1	Three	Balsakhi	48	78	2596	-	-
		No Balsakhi	48	80	2527	-	-
	Four	Balsakhi	48	73	2414	-	-
		No Balsakhi	49	77	2619	-	-
Year 2	Three	Balsakhi	61	101	3146	61	952
		No Balsakhi	61	93	2906	0	0
	Four	Balsakhi	61	96	3167	61	1011
		No Balsakhi	61	94	3170	0	0
Mumbai							
Year 1	Three	Balsakhi	32	70	2592	-	-
		No Balsakhi	35	65	2182	-	-
Year 2	Three	Balsakhi	39	74	2530	28	636
		No Balsakhi	38	79	2943	.	.
	Four	Balsakhi	38	77	2812	27	688
		No Balsakhi	39	71	2460	.	.

Table 2: Summary Statistics: Vadodara Year 1

	PRE TEST			MID TEST			POST TEST		
	Treatment	Control	Difference	Treatment	Control	Difference	Treatment	Control	Difference
STANDARD 3									
A. OBSERVATIONS	2596	2527	69	2204	2082	122	2094	2069	25
B. SCORES (PERCENTAGE)									
Math	0.262	0.264	-0.002	0.407	0.382	0.025	0.354	0.315	0.039
			(0.019)			(0.020)			(0.023)
Verbal	0.233	0.222	0.011	0.400	0.380	0.020	0.385	0.356	0.030
			(0.017)			(0.019)			(0.021)
C. NORMALIZED TEST SCORES									
Math	-0.008	0.000	-0.008	0.619	0.511	0.108	0.391	0.221	0.170
			(0.084)			(0.089)			(0.100)
Verbal	0.059	0.000	0.059	0.916	0.814	0.102	0.840	0.688	0.152
			(0.089)			(0.099)			(0.106)
D. PERCENTAGE OF CHILDREN PASSING EACH STANDARD COMPETENCY									
Math Standard 1	0.177	0.159	0.017	0.338	0.287	0.051	0.262	0.216	0.046
			(0.024)			(0.028)			(0.027)
Math Standard 2	0.012	0.018	-0.005	0.047	0.047	0.001	0.038	0.024	0.014
			(0.006)			(0.009)			(0.008)
Math Standard 3	0.042	0.031	0.012	0.132	0.113	0.020	0.091	0.064	0.027
			(0.010)			(0.020)			(0.017)
Verbal Standard 1	0.238	0.214	0.024	0.529	0.526	0.003	0.519	0.496	0.023
			(0.028)			(0.032)			(0.035)
Verbal Standard 2	0.158	0.150	0.008	0.331	0.317	0.014	0.284	0.252	0.032
			(0.024)			(0.028)			(0.031)
Verbal Standard 3	0.038	0.030	0.008	0.134	0.139	-0.005	0.095	0.076	0.019
			(0.011)			(0.022)			(0.020)
STANDARD 4									
A. OBSERVATIONS	2414	2619	-205	2127	2341	-214	1960	2178	-218
B. SCORES (PERCENTAGE)									
Math	0.441	0.452	-0.011	0.533	0.509	0.024	0.506	0.473	0.033
			(0.019)			(0.022)			(0.022)
Verbal	0.340	0.353	-0.013	0.511	0.497	0.014	0.498	0.485	0.013
			(0.017)			(0.022)			(0.023)
C. NORMALIZED TEST SCORES									
Math	-0.046	0.000	-0.046	0.349	0.247	0.102	0.231	0.088	0.142
			(0.080)			(0.096)			(0.095)
Verbal	-0.059	0.000	-0.059	0.738	0.672	0.067	0.677	0.617	0.060
			(0.078)			(0.101)			(0.108)
D. PERCENTAGE OF CHILDREN PASSING EACH STANDARD COMPETENCY									
Math Standard 1	0.365	0.386	-0.021	0.471	0.469	0.002	0.434	0.404	0.030
			(0.028)			(0.034)			(0.033)
Math Standard 2	0.064	0.057	0.007	0.108	0.097	0.011	0.084	0.078	0.006
			(0.010)			(0.016)			(0.013)
Math Standard 3	0.106	0.106	0.000	0.225	0.193	0.032	0.174	0.145	0.029
			(0.022)			(0.026)			(0.027)
Verbal Standard 1	0.437	0.459	-0.022	0.677	0.630	0.047	0.717	0.670	0.046
			(0.027)			(0.031)			(0.033)
Verbal Standard 2	0.274	0.316	-0.042	0.482	0.452	0.030	0.444	0.433	0.010
			(0.029)			(0.035)			(0.038)
Verbal Standard 3	0.111	0.123	-0.012	0.261	0.251	0.010	0.216	0.212	0.004
			(0.021)			(0.034)			(0.031)

Table 3: Summary Statistics: Vadodara Year 2

	PRE TEST			MID TEST			POST TEST		
	Treatment	Control	Difference	Treatment	Control	Difference	Treatment	Control	Difference
STANDARD 3									
A. OBSERVATIONS	3146	2906	240	2572	2409	163	3019	2790	229
B. SCORES (PERCENTAGE)									
Math	0.167	0.160	0.007	0.423	0.343	0.080	0.478	0.396	0.082
			(0.014)			(0.022)			(0.022)
Verbal	0.221	0.216	0.004	0.467	0.399	0.068	0.428	0.386	0.042
			(0.014)			(0.018)			(0.018)
C. NORMALIZED TEST SCORES									
Math	0.039	0.000	0.039	1.407	0.980	0.427	1.698	1.259	0.439
			(0.074)			(0.119)			(0.116)
Verbal	0.025	0.000	0.025	1.470	1.070	0.400	1.245	0.998	0.247
			(0.082)			(0.106)			(0.103)
D. PERCENTAGE OF CHILDREN PASSING EACH STANDARD COMPETENCY									
Math Standard 1	0.258	0.249	0.009	0.587	0.483	0.104	0.616	0.541	0.075
			(0.025)			(0.030)			(0.032)
Math Standard 2	0.016	0.018	-0.002	0.144	0.086	0.058	0.183	0.119	0.064
			(0.005)			(0.020)			(0.023)
Math Standard 3	0.001	0.000	0.001	0.023	0.009	0.014	0.057	0.034	0.023
			(0.001)			(0.005)			(0.010)
Verbal Standard 1	0.508	0.500	0.008	0.820	0.710	0.110	0.776	0.740	0.036
			(0.035)			(0.027)			(0.028)
Verbal Standard 2	0.073	0.088	-0.015	0.322	0.247	0.075	0.332	0.319	0.013
			(0.016)			(0.029)			(0.029)
Verbal Standard 3	0.016	0.009	0.007	0.137	0.124	0.013	0.146	0.120	0.026
			(0.005)			(0.020)			(0.021)
STANDARD 4									
A. OBSERVATIONS	3167	3170	-3	2893	2935	-42	3003	3007	-4
B. SCORES (PERCENTAGE)									
Math	0.309	0.297	0.012	0.529	0.436	0.094	0.575	0.498	0.076
			(0.018)			(0.016)			(0.020)
Verbal	0.348	0.332	0.017	0.561	0.476	0.085	0.517	0.457	0.060
			(0.017)			(0.017)			(0.018)
C. NORMALIZED TEST SCORES									
Math	0.051	0.000	0.051	1.002	0.598	0.405	1.197	0.869	0.329
			(0.077)			(0.069)			(0.087)
Verbal	0.083	0.000	0.083	1.135	0.716	0.418	0.916	0.621	0.295
			(0.083)			(0.083)			(0.089)
D. PERCENTAGE OF CHILDREN PASSING EACH STANDARD COMPETENCY									
Math Standard 1	0.470	0.449	0.021	0.737	0.642	0.095	0.779	0.695	0.084
			(0.030)			(0.020)			(0.029)
Math Standard 2	0.055	0.051	0.004	0.232	0.144	0.088	0.289	0.221	0.069
			(0.010)			(0.022)			(0.029)
Math Standard 3	0.015	0.012	0.003	0.052	0.040	0.012	0.110	0.082	0.028
			(0.003)			(0.009)			(0.014)
Verbal Standard 1	0.749	0.740	0.009	0.897	0.811	0.086	0.859	0.848	0.011
			(0.026)			(0.017)			(0.016)
Verbal Standard 2	0.186	0.181	0.005	0.441	0.325	0.116	0.481	0.363	0.118
			(0.020)			(0.026)			(0.028)
Verbal Standard 3	0.082	0.079	0.003	0.260	0.196	0.064	0.223	0.180	0.043
			(0.015)			(0.024)			(0.024)

Table 4: Summary statistics, Mumbai year 1

	PRE TEST			POST TEST		
	Treatment	Control	Difference	Treatment	Control	Difference
STANDARD 3						
A. OBSERVATIONS	2592	2182	410	2417	2027	390
B. SCORES (PERCENTAGE)						
Math	0.470	0.470	0.001	0.571	0.530	0.041
			(0.029)			(0.033)
Verbal	0.596	0.569	0.027	0.666	0.626	0.040
			(0.029)			(0.028)
C. NORMALIZED TEST SCORES						
Math	0.002	0.000	0.002	0.383	0.227	0.156
			(0.108)			(0.126)
Verbal	0.100	0.000	0.100	0.359	0.210	0.149
			(0.108)			(0.102)
D. PERCENTAGE OF CHILDREN PASSING EACH STANDARD COMPETENCY						
Math Standard 1	0.326	0.337	-0.012	0.397	0.357	0.040
			(0.038)			(0.039)
Math Standard 2	0.126	0.147	-0.021	0.211	0.195	0.017
			(0.025)			(0.036)
Math Standard 3	0.024	0.023	0.001	0.089	0.091	-0.003
			(0.007)			(0.022)
Verbal Standard 1	0.856	0.837	0.019	0.937	0.913	0.025
			(0.025)			(0.018)
Verbal Standard 2	0.486	0.473	0.013	0.577	0.526	0.050
			(0.045)			(0.042)
Verbal Standard 3	0.517	0.470	0.047	0.631	0.584	0.047
			(0.039)			(0.039)

Table 5: Summary Statistics: Mumbai Year 2

	PRE TEST			POST TEST			
	Treatment	Control	Difference	Treatment	Control	Difference	Implied Difference
STANDARD 3							
A. OBSERVATIONS	2530	2943	-413	2337	2731	-394	
B. SCORES (PERCENTAGE)							
Math	0.221	0.233	-0.012	0.502	0.470	0.031	0.044
			(0.016)			(0.028)	(0.038)
Verbal	0.351	0.344	0.007	0.588	0.569	0.018	0.025
			(0.022)			(0.025)	(0.034)
C. NORMALIZED TEST SCORES							
Math	-0.070	0.000	-0.070	1.509	1.333	0.176	0.245
			(0.087)			(0.155)	(0.213)
Verbal	0.025	0.000	0.025	0.898	0.831	0.067	0.093
			(0.082)			(0.091)	(0.126)
D. PERCENTAGE OF CHILDREN PASSING EACH STANDARD COMPETENCY							
Math Standard 1	0.137	0.167	-0.030	0.421	0.339	0.082	0.114
			(0.025)			(0.043)	(0.058)
Math Standard 2	0.082	0.090	-0.008	0.412	0.412	0.001	0.001
			(0.015)			(0.053)	(0.073)
Math Standard 3	0.003	0.006	-0.003	0.136	0.099	0.037	0.052
			(0.003)			(0.023)	(0.031)
Math Standard 4	0.007	0.013	-0.006	0.123	0.088	0.035	0.048
			(0.004)			(0.024)	(0.034)
Verbal Standard 1	0.653	0.648	0.005	0.820	0.817	0.004	0.005
			(0.036)			(0.022)	(0.031)
Verbal Standard 2	0.165	0.147	0.017	0.388	0.363	0.024	0.034
			(0.022)			(0.033)	(0.046)
Verbal Standard 3	0.137	0.131	0.005	0.317	0.307	0.010	0.013
			(0.021)			(0.034)	(0.047)
STANDARD 4							
A. OBSERVATIONS	2812	2460	352	2635	2290	345	
B. SCORES (PERCENTAGE)							
Math	0.409	0.396	0.013	0.642	0.564	0.079	0.106
			(0.019)			(0.027)	(0.034)
Verbal	0.555	0.530	0.025	0.721	0.683	0.038	0.051
			(0.021)			(0.021)	(0.026)
C. NORMALIZED TEST SCORES							
Math	0.053	0.000	0.053	0.995	0.678	0.317	0.426
			(0.076)			(0.111)	(0.136)
Verbal	0.083	0.000	0.083	0.641	0.513	0.127	0.171
			(0.071)			(0.069)	(0.086)
D. PERCENTAGE OF CHILDREN PASSING EACH STANDARD COMPETENCY							
Math Standard 1	0.300	0.240	0.060	0.474	0.387	0.087	0.117
			(0.031)			(0.036)	(0.046)
Math Standard 2	0.245	0.243	0.003	0.554	0.464	0.090	0.121
			(0.023)			(0.055)	(0.072)
Math Standard 3	0.042	0.041	0.001	0.241	0.171	0.069	0.093
			(0.010)			(0.033)	(0.042)
Math Standard 4	0.074	0.063	0.011	0.335	0.242	0.093	0.125
			(0.013)			(0.035)	(0.044)
Verbal Standard 1	0.825	0.796	0.029	0.923	0.900	0.023	0.030
			(0.022)			(0.014)	(0.017)
Verbal Standard 2	0.338	0.333	0.005	0.576	0.512	0.064	0.086
			(0.027)			(0.034)	(0.043)
Verbal Standard 3	0.355	0.317	0.038	0.532	0.485	0.047	0.064
			(0.031)			(0.033)	(0.043)

Table 6 : Attrition patterns

	Bombay, year 1			Bombay, year 2			Bombay 2 years			Vadodara, year 1			Vadodara, year 2		
	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference
Standard 3, All															
Percent attrition	0.070	0.075	-0.004	0.077	0.073	0.005	0.255	0.250	-0.006	0.193	0.181	0.012	0.040	0.040	0.000
Difference in score at pretest attriters-stayers			(0.015)			(0.010)			(0.009)			(0.020)			(0.008)
	-0.146	-0.274	0.128	-0.330	-0.193	-0.137	-0.445	-0.594	0.149	-0.130	-0.222	0.092	-0.060	-0.075	0.015
			(0.169)			(0.129)			(0.151)			(0.093)			(0.131)
Standard 4, All															
Percent attrition				0.063	0.070	-0.006				0.188	0.168	0.020	0.052	0.051	0.000
Difference in score at pretest attriters-stayers						(0.010)						(0.021)			(0.020)
				-0.180	-0.427	0.247				-0.178	-0.176	-0.002	-0.098	0.015	-0.113
						(0.139)						(0.077)			(0.155)

Table 7: Differences in differences estimate of the impact of the balsakhi program, by city and sample

	Number of Observations	Math	Verbal	Total
Mumbai and Vadodara together Year 1	12730	0.188	0.069	0.138
		(0.047)	(0.056)	(0.047)
Mumbai and Vadodara together Year 2	21805	0.319	0.153	0.250
		(0.067)	(0.050)	(0.059)
Pooling Both Standards				
Vadodara Year 1	8301	0.196	0.109	0.164
		(0.059)	(0.058)	(0.058)
Vadodara Year 2	11819	0.342	0.225	0.309
		(0.077)	(0.064)	(0.073)
Mumbai Year 2	9986	0.279	0.032	0.150
		(0.124)	(0.076)	(0.099)
Mumbai Year 2 Specification Check	9986	0.285	0.063	0.173
		(0.112)	(0.067)	(0.088)
Standard 3				
Vadodara Year 1	4163	0.198	0.101	0.162
		(0.092)	(0.090)	(0.089)
Vadodara Year 2	5809	0.399	0.224	0.342
		(0.111)	(0.094)	(0.103)
Mumbai Year 1	4429	0.163	0.060	0.118
		(0.072)	(0.072)	(0.067)
Mumbai Year 2	5063	0.312	0.043	0.163
		(0.164)	(0.108)	(0.133)
Mumbai Year 2 Specification Check	5063	0.276	0.073	0.168
		(0.149)	(0.097)	(0.121)
Standard 4				
Vadodara Year 1	4138	0.186	0.114	0.160
		(0.074)	(0.075)	(0.072)
Vadodara Year 2	6010	0.269	0.211	0.258
		(0.089)	(0.073)	(0.081)
Mumbai Year 2	4923	0.376	0.084	0.232
		(0.111)	(0.077)	(0.093)
Mumbai Year 2 Specification Check	4923	0.403	0.104	0.257
		(0.099)	(0.066)	(0.081)
Two Year 01-03				
Mumbai Pre-test Year 1 to Pre-test Year 2	3124	-0.129	0.003	-0.066
		(0.093)	(0.100)	(0.094)
Mumbai Pre-test Year 1 to Post-test Year 2	3299	0.521	0.145	0.351
		(0.142)	(0.113)	(0.115)

Table 8: Differences in differences for standard competencies, by city and year

	Math Competencies for				Verbal Competencies for		
	Standard 1	Standard 2	Standard 3	Standard 4	Standard 1	Standard 2	Standard 3
Vadodara							
Year 1							
Both Standards	0.050	0.011	0.021	-	0.039	0.040	0.014
	(0.022)	(0.007)	(0.019)	-	(0.023)	(0.019)	(0.015)
Standard Three	0.047	0.018	0.016	-	0.007	0.024	0.011
	(0.033)	(0.009)	(0.019)	-	(0.036)	(0.027)	(0.017)
Standard Four	0.051	0.004	0.027	-	0.070	0.058	0.019
	(0.031)	(0.011)	(0.033)	-	(0.028)	(0.027)	(0.022)
Year 2							
Both Standards	0.064	0.063	0.022	-	0.019	0.073	0.031
	(0.023)	(0.019)	(0.009)	-	(0.023)	(0.021)	(0.015)
Standard Three	0.065	0.066	0.022	-	0.030	0.029	0.020
	(0.032)	(0.022)	(0.010)	-	(0.034)	(0.028)	(0.021)
Standard Four	0.061	0.064	0.024	-	0.002	0.115	0.042
	(0.033)	(0.029)	(0.013)	-	(0.023)	(0.030)	(0.021)
Mumbai							
Year 1							
Standard Three	0.058	0.037	-0.003	-	0.004	0.045	0.007
	(0.032)	(0.030)	(0.020)	-	(0.021)	(0.030)	(0.032)
Year 2							
Both Standards	0.088	0.058	0.078	0.094	-0.013	0.041	0.009
	(0.038)	(0.047)	(0.025)	(0.032)	(0.030)	(0.025)	(0.037)
Standard Three	0.147	0.005	0.053	0.054	-0.008	0.003	0.003
	(0.056)	(0.066)	(0.029)	(0.032)	(0.049)	(0.035)	(0.043)
Standard Four	0.042	0.125	0.094	0.110	-0.003	0.084	0.016
	(0.048)	(0.064)	(0.039)	(0.045)	(0.030)	(0.036)	(0.061)

Table 9: Differences in differences, by third on the intitial test score distribution

	Year 1				Year 2					Two Year Analysis			
	N	Math	Verbal	Total	Fraction of Children who go to the Balsakhi	N	Math	Verbal	Total	N	Math	Verbal	Total
Vadodara and Mumbai together													
Bottom Third	4147	0.250	0.146	0.211	0.22	7293	0.507	0.133	0.316				
		(0.055)	(0.061)	(0.057)			(0.155)	(0.093)	(0.123)				
Middle Third	4271	0.179	0.036	0.115	0.16	7086	0.319	0.034	0.167				
		(0.057)	(0.068)	(0.054)			(0.136)	(0.089)	(0.111)				
Top Third	4312	0.127	0.016	0.079	0.06	7426	0.039	-0.038	-0.003				
		(0.062)	(0.072)	(0.060)			(0.151)	(0.076)	(0.106)				
Vadodara													
Bottom Third	2661	0.224	0.120	0.185	0.22	4007	0.458	0.308	0.417				
		(0.059)	(0.064)	(0.061)			(0.088)	(0.074)	(0.084)				
Middle Third	2784	0.186	0.140	0.173	0.18	3836	0.407	0.217	0.342				
		(0.075)	(0.070)	(0.069)			(0.094)	(0.072)	(0.083)				
Top Third	2856	0.156	0.060	0.118	0.09	3976	0.205	0.183	0.210				
		(0.084)	(0.083)	(0.081)			(0.083)	(0.086)	(0.082)				
Mumbai													
Bottom Third	1486	0.297	0.212	0.269	0.22	3286	0.507	0.133	0.316	949	0.754	0.329	0.572
		(0.109)	(0.099)	(0.102)			(0.155)	(0.093)	(0.123)		(0.196)	(0.146)	(0.167)
Middle Third	1487	0.147	-0.032	0.060	0.14	3250	0.319	0.034	0.167	1079	0.504	0.057	0.295
		(0.083)	(0.087)	(0.069)			(0.136)	(0.089)	(0.111)		(0.160)	(0.128)	(0.110)
Top Third	1456	0.049	0.005	0.029	0.03	3450	0.039	-0.038	-0.003	1096	0.342	0.069	0.217
		(0.069)	(0.070)	(0.063)			(0.151)	(0.076)	(0.106)		(0.121)	(0.101)	(0.093)

Table 10: Disentangling balsakhi and class size effects

	Balsakhi Assignment				Improvement in Test scores Pre to Post				Improvement in Test scores Vadodara	
	Mumbai	Vadodara	Both		Mumbai	Vadodara	Both		Pre to Mid	Mid to Post
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. First stages and reduced form										
Treatment School	0.259	0.457	0.363	0.471	0.434	0.636	0.590	0.524	1.039	-0.393
	(0.064)	(0.045)	(0.037)	(0.052)	(0.194)	(0.185)	(0.132)	(0.176)	(0.238)	(0.226)
Treatment * Rank <20	0.075	0.177	0.138	0.063	0.124	0.152	0.132	0.158	-0.047	0.201
	(0.029)	(0.022)	(0.017)	(0.020)	(0.086)	(0.080)	(0.063)	(0.065)	(0.082)	(0.096)
Treatment * Rank	-0.002	0.001	-0.001	-0.003	-0.003	0.003	0.000	0.005	-0.021	0.023
	(0.001)	(0.001)	(0.001)	(0.003)	(0.004)	(0.005)	(0.003)	(0.009)	(0.012)	(0.013)
Treatment * Pre-test score	-0.083	-0.088	-0.082	-0.077	0.004	-0.100	-0.055	-0.057	-0.152	0.072
	(0.027)	(0.013)	(0.013)	(0.014)	(0.097)	(0.087)	(0.065)	(0.066)	(0.076)	(0.070)
Treatment * Number of students	0.001	-0.004	-0.001	-0.001	-0.002	-0.007	-0.005	-0.005	-0.008	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.002)
Rank <20					-0.219	-0.031	-0.097	0.002	0.065	-0.047
					(0.063)	(0.059)	(0.045)	(0.046)	(0.058)	(0.066)
Rank					0.003	0.004	0.004	0.015	0.033	-0.008
					(0.002)	(0.004)	(0.002)	(0.005)	(0.008)	(0.008)
Pre-test score					-0.368	-0.340	-0.349	-0.356	-0.297	-0.032
					(0.069)	(0.060)	(0.042)	(0.043)	(0.061)	(0.043)
Number of students					0.000	-0.002	0.000	0.000	0.002	-0.004
					(0.001)	(0.003)	(0.001)	(0.001)	(0.004)	(0.001)
B. Instrumental variable estimates										
Saw a Balsakhi					1.383	0.891	0.954	2.785	0.105	0.833
					(0.753)	(0.490)	(0.408)	(1.644)	(0.385)	(0.504)
Balsakhi School					0.080	0.213	0.237	-0.877	0.753	-0.576
					(0.363)	(0.330)	(0.212)	(0.991)	(0.276)	(0.339)
Treatment * Rank					-0.001	0.003	0.001	0.015	0.003	-0.001
					(0.004)	(0.005)	(0.003)	(0.014)	(0.004)	(0.003)
Treatment * Pre-test score					0.134	-0.019	0.031	0.147	-0.149	0.160
					(0.128)	(0.104)	(0.079)	(0.134)	(0.086)	(0.096)
Treatment * Number of students					-0.003	-0.004	-0.004	-0.003	-0.007	0.004
					(0.002)	(0.005)	(0.002)	(0.002)	(0.004)	(0.003)
C. Instrumental Variable estimate (with interactions with the average pre-test score of balsakhi children)										
Saw a Balsakhi					1.281	0.813	1.029		0.335	0.524
					(0.717)	(0.538)	(0.447)		(0.412)	(0.511)
Balsakhi School					-0.211	0.299	0.128		0.469	-0.211
					(0.498)	(0.389)	(0.277)		(0.297)	(0.376)
Treatment school * Pre-test score					-0.168	0.150	-0.013		-0.365	0.486
					(0.356)	(0.283)	(0.203)		(0.224)	(0.242)

Notes: Column (4) has as additional controls Treatment*Rank Squared, Treatment*Rank Cubed, and Treatment*Rank^4
Column (8) includes fourth order polynomials in rank, and in treatment * rank. Space limitations preclude reporting of the polynomial coefficients

Table 11: Attendance

			Standard 3			Standard 4			Standard 3 and 4 Together	
		Avg Attendance Sept-Oct	Simple Differences	Difference in Difference	Avg Attendance Sept-Oct	Simple Differences	Difference in Difference	Avg Attendance Sept-Oct	Simple Differences	Difference in Difference
Year 1	**RA Attendance**									
	Whole Sample	0.88	-0.002	0.006						
			(0.012)	(0.017)						
	Roster Attendance									
	Whole Sample	0.93	-0.004	-0.014						
			(0.010)	(0.011)						
Year 2	**RA Attendance**									
	Whole Sample	0.88	0.013	0.005	0.88	-0.016	0.000	0.88	-0.001	0.001
			(0.017)	(0.023)		(0.014)	(0.021)		(0.011)	(0.016)
	Roster Attendance									
	Whole Sample	0.91	0.001	0.009	0.91	-0.021	0.011	0.91	-0.010	0.010
			(0.013)	(0.013)		(0.012)	(0.011)		(0.009)	(0.009)

Table 12: Cost Comparison

		Cost Per Year (Dollars)	Cost Per Year (Rupees)	Students	Rps/student per year
Vadodara					
year 1 & 2	Balsakhi	11830	520384	5730	91
year 1 & 2	Primary School Teachers	3924300	172740888	54525	3168
Mumbai					
year 1	Primary School Teachers (L Ward)	89488.6364	3937500	2988	1318
year 1	Balsakhi	4200	184800	3433	54
year 2	Balsakhi	7770	341880	4225	81

Note: Cost of teachers for Vadodara is calculated by dividing the total wage bill for teachers by the number of students, using data from year 1. In Mumbai, it is calculated using the starting salary of a teacher (7500 Rs/month) and the number of divisions in the schools in the study. In Vadodara, balsakhis were paid 500 Rs./month in both years. In the first year in Mumbai, balsakhis were paid 500 Rs./month, while in the second they were paid 750 Rs./month) (Schoolteacher salary figures are preliminary.)

Table 13: Cost Benefit Analysis

		Improvements in Test Scores		Rupees per standard deviation	
		Standard 3	Standard 4	Standard 3	Standard 4
Vadodara	Year 1				
	Balsakhi treatment effect	0.162	0.16	561	568
	Pre to Post difference (comparison)	0.46	0.36	6887	8800
	Year 2				
	Balsakhi treatment effect	0.356	0.249	255	365
	Pre to Post difference (comparison)	1.144	0.731	2769	4334
Mumbai	Year 1				
	Balsakhi treatment effect	0.115	n/a	468	n/a
	Pre to Post difference (comparison)	0.215	n/a	6129	n/a
	Year 2				
	Balsakhi treatment effect	0.151	0.226	536	358
	Pre to Post difference (comparison)	1.094	0.607	1205	2171