Remedying Education: Evidence from Two Randomized Experiments in India

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

November 13, 2005

Abstract

Many efforts to improve school quality by adding school resources have proven to be ineffective. This paper presents the results of two experiments conducted in Mumbai and Vadodara, India, designed to evaluate ways to improve the quality of education in urban slums. A remedial education program hired young women from the community to teach basic literacy and numeracy skills to children lagging behind in government schools. We find the program to be very effective: it increased average test scores of all children in treatment schools by 0.14 standard deviations in the first year, and 0.28 in the second year, relative to comparison schools. A computer-assisted learning program provided each child in the fourth standard with two hours of shared computer time per week, in which students played educational games that reinforced mathematics skills. The program was

*The authors are respectively from MIT (Department of Economics and Poverty Action Lab), Harvard Business School, MIT(Department of Economics and Poverty Action Lab), and Columbia University (Department of Economics, School of International and Public Affairs, and Poverty Action Lab). 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 requests for new data: Pratima Bandekar, Rukmini Banerji, Lekha Bhatt, Madhav Chavan, Shekhar Hardikar, Rajashree Kabare, Aditya Natraj, and many others. We thank Jim Berry and Marc Shotland, Mukesh Prajapati and Nandit Bhat for their excellent work coordinating the fieldwork, and for their remarkable work in developing and improving the CAL program. Kartini Shastry provided superb research assistance. Joshua Angrist, Angus Deaton, Rachel Glennerster, Michael Kremer, Alan Krueger, Victor Lavy, and Caroline Minter-Hoxby provided very helpful comments. For financial support, we thank the ICICI corporation, the World Bank, the Alfred E. Sloan Foundation and the MacArthur Foundation Network on the Costs of Inequality.

also very effective, increasing math scores by 0.36 standard deviations the first year, and 0.54 the second year. These results were not limited to the period in which students received assistance, but persisted for at least one year after leaving the program. Two instrumental variable strategies suggest that while remedial education benefited the children who attended the remedial classes, their classmates, who did not attend the remedial courses but did experience smaller classes, did not post gains, confirming that resources alone may not be sufficient to improve outcomes.

1 Introduction

The recent World Development Report on "Making Services Work for Poor People" (World Bank, 2004) illustrates well the essential tension in the public conversation about primary education in developing countries. On the one hand the report embraces the broad agreement, now enshrined in the Millennium Development Goals, that primary education should be universal. On the other hand, it describes in detail the dismal quality of the educational services that developing countries offer to the poor.

In rural India, for example, 25% of teachers were absent during random visits to schools throughout the country, and only 50% were actually teaching (Michael Kremer, Karthik Muralidharan, Nazmul Chaudhury, Jeffrey Hammer, 2004). In Uttar Pradesh (a large Indian State), a recent survey found that 41% of the children of primary school age cannot read a simple paragraph, 56% cannot write, and 63% cannot do simple additions (Abhijit Banerjee, Rukimini Banerji, Esther Duflo, Rachel Glennerster, Sendhil Mullainathan and Marc Shotland, 2005). Even in urban India, where widespread absenteeism by students and teachers is not an issue, the learning levels are very low: in Vadodara, a major Indian city and a site for the study in this paper, only 5.4% of the students enrolled in grade[1] three can correctly answer questions testing grade one math competencies.

In these conditions, policies that promote school attendance may not be in children's best interests. And indeed, the recent evidence suggests that many interventions which increase school participation do not improve test scores for the average student.[2] Students simply did

[1]In India, the term "standard" is used instead of the U.S. term "grade". We use grade throughout the paper.

[2]These include giving children deworming drugs (Edward Miguel and Kremer (2004) and providing school meals for children (Kremer and Christel Vermeersch, 2004).

not learn anything in the additional days that they spent at school.[3]

It is therefore clear that efforts to get children into school must be accompanied by significant improvements in the quality of schools that serve these children. The problem is that while we now know a reasonable amount about how to get children into school, much less is known about how to improve school quality in a cost-effective way. Worse still, a number of rigorous, randomized, evaluations have confirmed that spending more on resources like textbooks (Paul Glewwe, Kremer, and Sylvie Moulin 2002), flip charts (Glewwe, Kremer, Moulin and Eric Zitzewitz, 2004), or additional teachers (Banerjee, Kremer and Suraj Jacob, 2004) has no impact on children's test scores (see Kremer, 2003 for discussions and more references). These results have led to a general skepticism about the ability of interventions focusing on inputs to make a difference (echoing Hanusehk's (1986 and 1995) earlier assessment of both the US and developing countries), and have led many to advocate more systemic reforms designed to change the incentives faced by teachers, parents and children. The World Development Report, once again, embraces this view, and proceeds to propose various ways to improve incentives (most of which have either not been rigorously tested or when tested, have also proven rather disappointing).[4]

It is not clear, however, that we know enough to entirely give up on inputs. Based on the existing evidence it seems possible that the problem is rather that we are not providing inputs that address a specific unmet needs in the school.

Ironically the difficulty in improving the quality of education may in part be a by-product of the success in getting more children to attend school. Neither the infrastructure, nor the curriculum, has been adapted to take into account the influx of children and their characteristics: many of these children are first generation learners whose parents are not in a position to follow what is happening in school or react if their child is falling behind. For example, in the Uttar Pradesh survey mentioned previously, it was found that while 41% of the children cannot read a simple paragraph, only 21% of the parents think that their children cannot read.

Meanwhile, in many countries, the school system continues to operate as if it were catering to the elite. This may explain why just providing more inputs to the existing system, or more

[3]This is true even if we restrict our attention to children who were enrolled before the intervention, suggesting this result is not due to a change in the composition of the children.

[4]With the important exception of incentives for children (Kremer, Edward Miguel, and Rebecca Thornton, 2005).

school days, is often ineffective. For many children, neither more inputs nor an extra day makes much of a difference, because what is being taught in class is too hard for them. For example, Glewwe and Kremer (1997) found that new textbooks make no difference for the test scores of the average child, but do help those who had already done well on the pretest. The authors suggest that this is because the textbooks were written in English (the language of instruction, in theory), which most children do not speak very well at all. If most parents have little understanding of what is supposed to happen in school, this may also explain why giving more power to the parents has proven disappointing.

Taken together, these results suggest the possibility that inputs specifically targeted to helping weaker students learn may potentially have large effects.

This paper reports the results from randomized evaluations of two programs that provide supplementary inputs to children at the bottom of the class. The first intervention is specifically targeted to the weakest children: it is a remedial education program, where a young woman ("balsakhi") from the community works on basic skills with children who have reached grade three or four without having mastered them. These children are taken out of the regular classroom to work with this young woman for 2 hours per day (the school day is about 4 hours). The second intervention could potentially benefit all children, but is adapted to a child's current level of achievement: it is a computer-assisted learning program, where children in grade four are offered two hours of shared computer time per week, during which they play games that involve solving math problems at varying levels of difficulty. Both programs are provided by Pratham, a very large NGO operating in conjunction with government schools in India.

The programs that we evaluated were run in Mumbai (formerly known as Bombay) and Vadodara (formerly known as Baroda), two of the most important cities in Western India. Both cities have a reputation for having relatively well-run government school systems by Indian standards. Teacher attendance was not seen as a big problem and most children come to school regularly. Nevertheless, the state of education when the evaluation began was rather abysmal: as we noted, only 5.4% of third grade children in Vadodara pass the grade one competencies in math (14% do so in Mumbai). These poor results occur despite the fact that both teachers and children were coming to school. Improvements therefore require changing what children learn while in school.

In contrast to the disappointing results of the earlier literature, we find that both programs had a substantial positive effect on children's academic achievement. This is true in all years and cities, despite the instability of the environment (notably a major riot in one of the cities in 2003). The remedial education program increased average test scores in the treatment schools by 0.14 standard deviations in the first year, and 0.28 in the second year. Moreover, the weaker students, who are the primary target of the program, gained the most. In the second year, children in the bottom quintile of the initial distribution gained over 0.40 standard deviations. The computer-assisted learning increased math scores by 0.36 standard deviations the first year, and 0.51 the second year, and was equally effective for all students. These results persist over time: one year after leaving the program, initially low scoring students who were in balsakhi schools still do better than their untreated counterparts, though the difference is smaller. Students of all levels perform better in math if they were in schools where the math computer assisted learning program was implemented.

Moreover, using the assignment rules for the remedial education program to form instruments, we are able to separately identify the direct effect of attending the remedial education classes, and the indirect effect on children whose weakest peers were removed from the classroom for part of the day. The results suggest that the remedial education program benefited only the children who went to the balsakhi and that the effect on those students is very high (0.6 to 1 standard deviation). Since this was a pull-out program, the children who remained in class could have also benefited from a reduced student teacher ratio. The striking fact is that they did not suggests that reducing class size without changing pedagogy may not help, a conclusion that fits with the view that inputs per se are not sufficient.

Our confidence in these results is bolstered by the way the evaluation was carried out. We had the opportunity to implement an evaluation design that is often recommended but rarely utilized. Both programs were allocated using random assignment across the same set of schools, in a way that allowed us to separately estimate the effects of the two programs as well as compare their impacts. Second, the programs were run on a very large scale (over 15,000 students were included in the study over 3 years). In addition, the remedial education program in particular had already been taken to scale in a number of cities. Third, we simultaneously carried out randomized evaluations of the remedial education program in two different cities, each of which had its own management team, and got similar results. This reinforces our confidence in the external validity of these results. Fourth, we conducted each evaluation over two years, using several tests to assess learning, making it less likely that the results are a consequence of the newness of the program, or the effect of implementing an evaluation. In other words, there is little risk that what we are evaluating cannot be reproduced elsewhere. Finally, we did not only measure short term effect, but measured the persistence of the program as well.

Taken together, these results suggest that it is possible to substantially improve the quality of education in developing countries with cheap interventions, and without a fundamental transformation of the system. Computers provided children some relief from the indifferent teaching they have come to expect, but did not raise the ire of the teacher's union. Indeed, no teachers objected to the program, and many believed it was beneficial . Even the balsakhis, who pose a more obvious threat to teachers, and did meet some initial resistance, became more and more welcome as the teachers realized that they were actually taking some of the more difficult children off the teacher's hands.

The fact that both the interventions had an effect but the reduction in class size did not, suggests that we cannot simply go with a conventional educational production function approach. It is not the number of teachers that seems to matter but how they are deployed, and what they do. Our results are therefore consistent with the earlier literature, which found no effects of additional inputs. The computer-assisted learning program illustrates the same point even better. As we will see below, the program took advantage of computers that were already in the schools, but were not used. The program found a way to make these computers pedagogically useful in the treatment schools, without putting more demands on the teachers' time. It is the utilization in this specific way and not the possession of the computers that had the impact.

In other words, we read the evidence as saying that resources can make a difference. But to improve the quality of education for the children brought into school by "education for all" campaigns, resources must target students' needs, rather than be used to just provide more of what the schools already have.

Our paper is organized as follows. In Section 2, we describe the remedial education and computer assisted learning interventions in detail. Section 3 describes the evaluation design. Sections 4 checks for preexisting differences between the treatment and comparison groups as discusses well as attrition patterns. In Section 5, we present the results of the evaluation and in Section 6, we distinguish the direct effects of being taught by a remedial education instructor from the indirect effects of being left behind in a smaller, more homogenous class. Section 7 evaluates the cost effectiveness of the interventions.

2 The Programs

The interventions evaluated in this study were implemented in conjunction with Pratham, a network of India-based NGO's. 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 300,000 children in 13 States in India, employing thousands. It works closely with the government: most of its programs are conducted in the municipal schools, and Pratham also provides technical assistance to the government.

2.1 Remedial Education: The Balsakhi Program

One of Pratham's core programs at the time of this study was a remedial education program, called the Balsakhi Program (balsakhi means "the child's friend"). This program, in place in many municipal schools, provides government schools with a teacher ("balsakhi," 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 instructor typically meets with a group of approximately 15-20 children in a class for two hours a day during school hours (the school day is about 4 hours long). Instruction focuses on the core competencies the children should have learned in the first and second grades, primarily basic numeracy and literacy skills. The instructors are provided with a standardized curriculum that was developed by Pratham. They received 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 not capable of following the standard curriculum.

Children may feel more comfortable with women from their own communities than the teachers, who are often from different backgrounds. As the balsakhi works with a relatively small class, 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 for intervention. Removing children from the classroom for two hours means the effective student-teacher ratio in the main classroom drops, and may allow the teacher to focus on more advanced material with the other children. Finally, if the balsakhis are indeed effective, when children return to the main classroom, the teacher may not need to re-teach 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 stays 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, playground, or even hallways when necessary), the program has very low overhead and capital costs.

2.2 Computer-Assisted Learning

The Computer-Assisted Learning (CAL) program takes advantage of a policy put in place by the government of Gujarat and the established infrastructure of the Balsakhi program. In 2000, the government delivered four computers to each of 100 municipal primary schools in the city (80% of the public schools). A survey conducted by Pratham in June 2002 suggested that very few of these computers were actually used by children in elementary grade levels. While some schools may have run programs for older students or allowed teachers to use the computers for administrative tasks, most computers remained in their boxes, for want of anyone capable of operating them.

This situation is not unique to Gujarat. Many in India see computer assisted learning as a supplement to regular instruction, that is, as a way to improve the quality of education. Good educational software can be reproduced at nominal cost, and well-designed educational games can sustain interest and curiosity in an otherwise dull school environment. The excitement seems to be particularly strong in India, where the high-tech sector is both successful and visible. Many local governments have started providing computers in schools, but without guidance about how the schools should use them. The idea of using computers is particularly attractive in areas where the number of qualified teachers is limited and the quality of existing teachers is notoriously poor. Computers have the potential to both directly improve learning and indirectly increase attendance by making school more attractive.

Unfortunately, despite this general excitement, there exists very little rigorous evidence of the impact of computers on educational outcomes and no reliable evidence for India or other developing countries. What's more, the little evidence available is not encouraging. For example, Joshua Angrist and Victor Lavy (2002) evaluate a computer-assisted learning program in Israeli schools with disappointing results. Among the fourth and eighth grade students evaluated with math and Hebrew exams, the data show no benefits for computer-assisted instruction and provide some evidence that children who received such instruction are actually at a disadvantage. Alan Krueger and Cecilia Rouse (2004) report a randomized evaluation of the language software "Fast ForWord" commonly used in US classrooms, and find no impact.

It is not clear, however, that these results apply to the use of computers in schools in developing countries: in Israel and in the US, the computer-assisted learning replaces time spent in well equipped classrooms with high quality instructors. It is easy to imagine that computers can make a significant improvement in schools in developing countries even if they are not useful in the developed world.

Pratham had previous experience with computer-assisted learning, having run a small computer-assisted learning program in Mumbai for several years. In particular, they had developed instructional software in the local language, Gujarati. After consultation with the Vadodara Municipal Corporation (VMC), they introduced a computer-assisted learning program in half of the VMC schools, using the computers already present when possible and replacing or adding computers where necessary.

Pratham hired a team of instructors from the local community and provided them with five days of computer training. These instructors provided children with two hours of shared computer time per week (two children shared one computer) - one hour during class time and one hour either immediately before or after school. During that time, the children played a variety of educational computer games selected because they emphasized basic competencies in the VMC mathematics curriculum. In the second year of the program, Pratham teamed up with Media-pro, a company that develop instruction software, to develop a suite of software that more closely followed the curriculum. Children also completed simple worksheets designed to track their progress at the beginning of each session.

Pratham designed the program to allow the children to learn as independently as possible. The instructors encouraged each child to play games that challenged the student's level of comprehension, and when necessary, they helped individual children understand the tasks required of them by the game. All interaction between the students and instructors was driven by the child's use of the various games, and at no time did any of the instructors provide general instruction in mathematics.

Schools at which the CAL program was not implemented were free to use the computer on their own, but we did not observe those schools employ them for instructional purpose.

3 Evaluation Design

3.1 Sample: Vadodara

• Balsakhi

In 2000, Pratham expanded their remedial education program (Balsakhi) with a view to progressively cover the entire city of Vadodara, meanwhile taking advantage of the expansion, which affected 98 schools, to evaluate the efficacy of the program. In November, 2000, they administered an academic test (designed by the Pratham team) to all children in the third grade. They then hired and trained balsakhis, who were sent to half of the schools in Vadodara. Assignment was random, with schools stratified by language ("medium" in the official terminology) 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]

[5]Throughout the paper, we will refer to the year 2001-2002 as "year 1", year 2002-2003 as "year 2" and year 2003-2004 as "year 3."

In July, 2001 (the beginning of "year 1"), the group of schools that had received a balsakhi in the previous year of the program received the balsakhi in the fourth grade, and the remaining schools received a balsakhi in the third grade.

The program continued during the school year 2002-2003, with the addition of the 25 remaining primary schools. Schools where the balsakhi was assigned in grade three in the year 2001-2002 were now assigned a balsakhi in grade four, so that by the end of year 2, grade four children in the treatment group benefited from two years of the Balsakhi program if they stayed in the same school. Schools where the balsakhi was assigned to grade four in year 1 received balsakhi assistance for grade three in year 2. The 25 remaining primary schools were also added by randomly assigning them to the research groups 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.

Given this design, children in grade three in schools that received the program in grade four form the comparison group for children that receive the program in grade three, and vice versa. While the assignment strategy ensures treatment and comparison groups are comparable, 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 (an in particular in Vadodara) makes this extremely unlikely: schools have only one class (a group of students and a teacher) per grade. All students are automatically promoted, so that the principals have no discretion in the number of students per class or the number of teachers per grade. Most schools have just enough class rooms for each class, and the balsakhi class typically met outside, under a tree. There are essentially no other resources to speak of that the head teacher could allocate to the grade that did not receive the balsakhi. Thus, we are confident that there was no reallocation of resources to the grade that did not receive the balsakhi, which makes these students a good comparison group.

• Computer-Assisted Learning

The CAL program was implemented in half of the municipal primary schools in Vadodara in 2002-2003, focusing exclusively on children in grade four. The sample was stratified according to treatment or comparison status for the grade four Balsakhi program, as well as gender, language of instruction of the school, and average math test scores in the posttest in the previous year.

Table 1 summarizes the allocation of schools across different groups in the program. In some 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. Thus, in the final sample for the study, 55 schools received the CAL program and 56 serve as the comparison group. The program was continued in 2003-2004, after switching the treatment and comparison groups.

3.2 Sample: Mumbai

To ensure the results from the Vadodara study would be generalizable, the Balsakhi program was also evaluated in Mumbai, 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, Urdu, and Marathi schools. In total, 62 schools were included in the study. Schools were stratified according to their scores in a pretest, as well as by the medium of instruction. Half the schools were randomly selected to receive a balsakhi in grade two, and half the schools were randomly selected to receive a balsakhi in grade three. Unfortunately, students in grade two were not accustomed to taking standardized exams and no satisfactory test could be developed for them,[6] obliging us to remove them from the evaluation immediately after the initial pretest. In 2001-2002, data were collected only for grade three children, while in 2002-2003, we expanded the study to include students in grade four. As in Vadodara, children kept their treatment assignment status as they moved from grade 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.[7]. Throughout the paper, the schools that

[6]Students did have experience with exams administered by the teachers, but in these cases the teacher often gave substantial assistance (including writing answers on the board).

[7]All the children were tested, however: Schools could not refuse testing, because Pratham had obtained written

were assigned balsakhis but did not get them are included in the "intention to treat" group. The analysis then adjusts for the fraction of the treatment group that was effectively treated.

3.3 Outcomes

The main outcome of interest is whether the interventions resulted in any improvement in learning levels.

In the Vadodara pilot year, children were given a pretest in November, 2000, and posttest in March, 2001. In the first full year, a pretest was given at the beginning of the school year (August 2001), a midtest in October 2001, and a posttest was 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 grade three and four children, so that the scores can be directly compared across grades. Scores on the pre- and posttest can also be directly compared within a given year, as the format of the questions and the competencies tested were the same.[8] The exam contains a math and a language section. In Vadodara, both sections focused on competencies that the Vadodara Municipal Corporation (VMC) prescribe for children in grades 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. Testing was similar in Mumbai. In the first year, tests focused on competencies in grades one through three, while in the second year they included grades one through four. In the second year, the same test was used for third and fourth grade 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. To minimize attrition, the

permission for testing from the city administration

[8]The tests did, however, change between years.

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 a make-up test outside of school.

Another outcome of interest is attendance and dropout rates. These were collected weekly by Pratham employees who made randomly timed visits to classrooms to take attendance with a roll call.

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 in a session).

4 Results: Pre-Intervention Difference and Attrition Patterns

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

Tables 2 through 4 present descriptive statistics of the test scores for all samples used in this analysis (year 1 and 2 in Vadodara and Mumbai). The scores are normalized relative to the distribution of the pretest score in the comparison group in each city, grade, and year.[9] The appendix Tables 1 to 5 show the raw scores as well as the percentage of children who correctly answered the questions in the test relating to the competencies in each grade.

The randomization appears to have been successful: with the exception of the CAL program in year 3 in Vadodara, none of the differences between the treatment and comparison groups prior to the implementation of the program are statistically distinguishable from zero. The point estimates are also very small, with each difference less than a tenth of a standard deviation.

Table 5 implements bootstrap tests of equality of distributions proposed by Alberto Abadie (2002).[10] The first row in Table 5 reports the p-value for the hypothesis that the two distributions are equal, while the second row reports the p-value for the hypothesis that the treatment

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

[10] The test uses the Kolmogorov-Smirnov statistic to measure the discrepancy between the hypothesis of equality of distributions and the data.

distribution stochastically dominates the comparison distribution. The third line presents the p-value for the hypothesis that the comparison distribution stochastically dominates the treatment distribution. For the pretest scores, the distributions between treatment and comparison can never be statistically distinguished from each other, except in the case of the CAL program in year 2. Figure 1 shows an illustrative example of the cumulative distributive function of the test score in the first year (Vadodara, year 2): the distributions are clearly on top of each other.

The raw scores (presented in appendices), and the percentage of children correctly answering the questions relating to the curriculum in each grade give an idea of how little these children actually know. In grade three in Vadodara in the second year, for example, the average student in math scores about 16%, both in the comparison and treatment groups. Since one of the math questions is multiple-choice, on average a student who picks a random answer to that question will score 1.8%. 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 grade performance. Only 5.4% of third grade children in Vadodara pass the grade one competencies in math in grade three in Vadodara (and 14% in Mumbai). Grade one competencies cover number recognition, counting and one digit addition and subtraction.

The results are more encouraging in verbal competencies: 50% of the grade three children pass the grade one competencies in Vadodara (reading a single word, choosing the right spelling among different possible spellings for a word), and 65% do so in Mumbai.

4.2 Attrition and Transfers

Differential attrition between the treatment and comparison groups could potentially bias the results. The testing procedure (the survey team visited children who were not present at the posttest in their home, and administered the test then) was designed to minimize attrition, and was largely successful.

Table 6A and 6B present the levels of attrition in Mumbai and Vadodara for both programs. We present attrition that occurred between the pretest and posttest for both cities in both years, as well as the two-year attrition (in Mumbai, for grade four only), broken down by treatment status.

Attrition was generally very low, except for Vadodara in year 1. The high attrition in that year is likely attributable to civil unrest (severe riots affected the city in 2002). The posttest was conducted after the riots and while the research team attempted to track down all of the children who had not appeared for the exam, many families had left Vadodara for their native villages. Nevertheless, the attrition rates did not vary by treatment status in that year or any other: 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 both the balsakhi and non balsakhi 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. In the CAL program, attrition was 4.1% in the treatment group and 4.8% in the comparison group in year 1, and 7.3% and 6.8% respectively in year 2.

The fact that there was no differential attrition in the treatment and comparison 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 comparison groups (Angrist, 1996). This does not seem to occur in our study: the second row in each panel presents differences in pretest scores of children who were not present at the posttest, by treatment status. The third column of each sample group presents 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. However, the difference is very similar in the treatment and comparison groups. In Mumbai in the second year, there is some evidence that the attritors may have had lower pretest scores than the stayers in the treatment group, compared to the comparison group. In the CAL program in year 2 and year 3, we find the opposite, with the attritors in the treatment group seeming to perform better than the non-attritors (which is different from the other results in the tables). This latter difference may bias the results obtained from simple differences upwards (the effect on the difference-in-difference estimate and the lagged dependent variables specification are unclear), although since the attrition is very low, this is unlikely to have a large effect.

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 here to save space).

In what follows, the treatment status of a child will be assigned based upon the school in which the child took the pretest. Student transfer could theoretically introduce two sources of bias. First, if students were able to transfer prior to the pretest, then treatment schools may have gained students likely to experience a significant improvement in test scores over the following year, generating an upward bias. Second, if motivated students transferred during the academic year, then some of the comparison group would have experienced the treatment causing us to underestimate the treatment effect.

These biases are not a concern in our setting. The program was not announced prior to the start of the school year. In addition, parents rarely inquire about programs offered through the school. And even if they were interested, school transfers are very unlikely in both Vadodara and Mumbai. Administrators provide them only reluctantly, and parents have a limited number of alternative schools. Most areas have only a few schools of the same medium within a large radius. Finally, since we were sensitive to the potential problems that could arise due to transfers, we checked for students that took the pretest in a comparison school and the posttest in a treatment school and found none.

5 Effects of the Balsakhi and the CAL Programs

5.1 Attendance

Part of the goal of the Balsakhi program was to make it easier for parents to play a role in their children's education, by allowing the balsakhis to serve as an intermediary between parents and the school environment. One could therefore have expected the program to affect attendance. In practice, it did not seem to: Table 13 shows the effect of the program of attendance in both cities (attendance was not collected in year 1 in Vadodara). In no city and no grade do we see any impact of the program on attendance.[11] The CAL program could have affected attendance as well, by making school more attractive for students, at least on days where they are scheduled to go to the CAL program. Table 13 also shows the effect of the CAL program on attendance. In year 2 (the first year of the program), CAL appears to have no effect on attendance. In year

[11] The data on attendance was obtained by roll calls during unannounced visits. In Mumbai, we also collected attendance from rosters filled out by the teachers. The rosters generally show a higher attendance rate than Pratham roll calls, but there is no difference in the measure between treatment and comparison schools.

3, we see a small, positive effect, significant at the 10 percent level (the effect 2.5 percentage points, with a standard errors of 1.5 percent). This may indicate a small effect of the CAL program on attendance, but the fact that the treatment schools scored slightly higher on test scores before the program warrants a cautious interpretation.

The low impact of both of these programs on attendance may be due to the relatively high attendance levels of children before the program. They make the result on learning, to which we now turn, easier to interpret: the results are entirely attributable to what happen in the school when children are already attending.

5.2 Test Scores: Balsakhi Program

Tables 2 and 3 present the first estimates of the effect of the Balsakhi program, as simple differences between the posttest scores in the treatment and comparison groups.

The Balsakhi program appears to be successful: in all years and grades, for both tests, and in both cities, and for all subgroups, the difference in posttest scores between treatment and comparison groups is positive.

In the first year in Vadodara (Table 2), the difference in posttest scores between treatment and comparison groups in grade three was 0.18 standard deviations in grade three for math, 0.16 for language. In grade four the corresponding differences for math and language were 0.16 and 0.09. Note that between the mid test and the posttest in year 1, scores actually declined. This is likely due to the riots, which severely disturbed the schools and the children. The results in Mumbai (Table 4) are remarkably similar, with the math and language test scores improving by 0.16 and 0.15 standard deviations, respectively.

In the second year of the program, the effects are larger. In Vadodara (Table 2), the difference in grade three total test scores is .44 for math and 0.25 for language; in grade four the differences are .34 and 0.30 for math and language respectively. To obtain the estimate for Mumbai in year 2, because one third of treatment schools did not get a balsakhi, the simple difference according to initial treatment status (intention to treat) is divided by the probability of treatment: in other words, this is an IV (or Wald) estimate where the initial treatment assignment is used as an instrument for the actual treatment status. In Mumbai in year two (Table 3), the Wald estimates (presented in the last column) of the impact of the program on test score differences for math and language respectively are .26 and .11 in grade three and .49 and .20 in grade four. In year two in Vadodara, all of the differences between treatment and comparison groups are statistically significant, while for Mumbai, the grade four results are significant.[12]

Table 5 compares the entire distribution of the posttest scores. The hypothesis that the two distributions are equal can be rejected at the one percent level in year 2 for both grades in Vadodara, and at the 1.4% level for grade four in Mumbai in year two, and at the 11% level in year 1 grade three in Vadodara. (Equality cannot be rejected for grade four grade three in Mumbai years 1 and 2, or Vadodara grade four in year 1.) The hypothesis that the comparison group distribution stochastically dominates the distribution of the treatment group can be rejected at the 10% level or better for all groups except Vadodara grade four in year 1, while the hypothesis that the distribution in the treatment group dominates that of the comparison group can never be rejected. Figure 2 illustrates this result in one case (Vadodara, year 2). The contrast with Figure 1 is striking: while the two curves were on top of each other in the pretest, the distribution of test scores in year 2 has clearly shifted to the left in the treatment schools, relative to the comparison schools.

Because test scores have a strong persistent component, the precision of the estimated program effect can be increased substantially by controlling for a child's pretest score or employing a difference-in-difference estimator. Since the randomization appeared to be successful, and attrition was low in both the treatment and comparison groups, the point estimates should be similar to the simple differences in these two specifications, but the confidence interval around these point estimates should be much tighter.

Table 7 presents the results, for various years, cities, grades, and sub-groups, from the estimation of a difference-in-differences (columns 1 to 3) and value added (column 4 to 6) specification.

The value added specification simply regresses the posttest score on the treatment status, controlling for the pretest score of child i in grade g and school j.

[12]All standard errors reported in the paper are adjusted for clustering at the school-grade level. Using nested random effects (classroom effects nested within school effects) yields very similar point estimates and generally higher t-statistics.

$$y_{igjPOST} = \lambda + \delta D_{jg} + \theta y_{igjPRE} + \epsilon_{igjPOST}, \quad (1)$$

where D_{jg} is a dummy equal to 1 if the school received a balsakhi in the child's grade g, and 0 otherwise.

This specification asks whether children improved more, relative to what they would have been expected to on the basis of their pretest score, in treatment schools than in comparison schools.

For all years and samples except Mumbai, equation (1) is estimated with OLS. However, for Mumbai in year two (and when both cities are pooled), Equation 1 is estimated by two stage least squares, using the initial treatment status as an instrument (to take into account that not all schools that should have received a balsakhi received one).

Second, we stack the pre and post data and use the following difference-in-difference specification, letting k index the individual tests (pre and post):

$$y_{igjk} = \lambda + \delta D_{jg} + \theta POST_k + \gamma(D_{ig} * POST_k) + \epsilon_{igjk}, \quad (2)$$

where $POST_k$ is a dummy indicating whether the test is the post test.

For Mumbai in year two (and when both cities are pooled), equation 2 is estimated with instrumental variables, with the initial assignment to the treatment group and its interaction with the posttest dummy serving as instruments. We also present a specification check where we include only schools that received the balsakhi in the grade for which she was assigned, and estimate the regression using OLS. The estimates using either specification are very similar.

In accord with the simple difference results, the point estimates from both specifications suggest a substantial treatment effect. Pooling both cities and grades together (in the first two rows of Table 7), the impact of the program was 0.14 standard deviations overall in the first year, and 0.27 standard deviations in the second year (0.28 using the value added specification). All estimates for total score are significant at the 1% level.

The impact is bigger in the second year than the first, for both math (0.34 vs. 0.19) and verbal (0.17 vs..0.06); all but first-year verbal scores are significant at the 1% level. For both years and both subjects pooled, the effects are a little larger in Vadodara than in Mumbai (with a total-score effect of 0.14 standard deviations versus 0.12 in the first year (grade three), and

0.31 versus 0.20 in the second year (both grades). The difference is the weakest for language, where there is a significant impact in both years for Vadodara (0.11 and 0.23 standard deviations respectively), but no significant impact in either year in Mumbai for grade three (0.06 standard deviations in year 1, and 0.051 standard deviations in year 2). For both cities and both subjects, the effects are very similar in grade three and grade four. [13] Results are also very similar when the analysis is conducted separately for girls or boys (results not reported).

In the second to last panel of the table, we present an estimate of the impact of two years of exposure to the program. These are estimates of the difference between the year 1 (2001-02) pretest and year 2 (2002-03) posttest for students that were in the third grade during the 2001-02 academic year.[14] Table 6a also shows that there was more attrition in this group, due to the movement of children across schools during the summer (25% in Mumbai, 33% in Vadorara, though the differences are also insignificant). In Mumbai, the effect of two years of treatment (from year 1 pretest score to year 2 posttest score) is substantially larger than that of either individual year (0.60 standard deviations in math, for example, versus 0.40 for year 2 in grade 4). 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, but larger than the one-year effect for the first year of the program. One possible explanation for this difference is the riots at the end of year 1, which severely disturbed the children.

Compared to the other educational interventions, this program thus appears to be quite effective. The Tennessee STAR experiment, for example, for which class size was reduced by 7 to 8 children (from 22 to about 15), improved test scores by about 0.21 standard deviations (Krueger and Diane Whitmore, 2001). The Balsakhi program improved test scores by 0.27 standard deviations in the second year, by reducing effective class size from 40 to 20 children for part of the day. However, the balsakhis were paid less than one tenth the teacher's salary (a starting teacher earned about Rs. 7,500 at the time, while balsakhi's were paid between Rs. 500 and Rs. 750). Section 7 offers a more detailed cost benefit analysis.

[13] Results disaggregated by grade are shown in Appendix, table 5.

[14] Only children who were in grade 3 in year 1 can be exposed for 2 years. Thus, the two-year effect is estimated using substantially fewer students than the one-year effect.

5.3 Test Scores: Computer-Assisted Learning

Table 4 shows the simple difference in the mid and post test in the CAL program. In the posttest, the math test scores are significantly greater in the treatment schools than in comparison schools in both years. In year 2, the math posttest score is on average 0.33 standard deviations higher in the CAL schools (with a standard deviation of 0.087). In year 3, it is 0.63 standard deviations higher, but this does not account for the fact that pretest scores were already 0.15 higher in the treatment group in year 3. Table 8 addresses this problem, by showing the difference-in-difference and value added specifications of the effect of the CAL program. The CAL program has a strong effect on math score (0.36 standard deviations in the first year, and 0.51 standard deviations in the second year, using the value added specification). It has no discernible impact on language scores (the point estimates are also very close to zero). This is not surprising, since the software targeted exclusively math skills, although one could have expected some spillover effects on language skills (through increased attendance, the practice of reading instructions, or if the teachers had reallocated time away from math to reading). The effect on the sum of language and math test scores is 0.18 standard deviations in year 2, and 0.19 standard deviations in year 3. Like those of the Balsakhi program, these are not temporary gains. Panel C in the table shows the estimated treatment effect of the students in the 2002-03 cohort from the pretest in 2002 through a posttest administered to them at the end of their fifth grade year in 2004. In both specifications, the treatment effect remains very strong (0.35 standard deviation in math) a year after the students leave the program. Again, this indicates that the CAL program provided children with a durable increment in learning. Panel B compares the Balsakhi and the CAL effects, and examines their interactions, in the year where they were implemented at the same time (the randomization was stratified). When not interacted, CAL has a larger effect on math test scores than Balsakhi (although this difference is not significant) and a smaller effect on overall test score (although again the difference is not significant). The programs appear to have no interaction with each other: the coefficients on the interaction on the math and overall test score are small, insignificant, and negative.

5.4 Effect on Specific Competencies and Distributional Effects

The Balsakhi program was primarily intended to help children at the lower end of the ability distribution, by providing targeted instruction to them. However, 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.

The program could also have, perversely, 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). While this would result in an improvement in average test score, this should probably not be construed as a success of the program. It is therefore important to check at what level children are affected. In practice, the program appears to have helped children who were initially lagging behind. As we discussed earlier, Table 5 shows that the distribution of the test scores in the treatment schools stochastically dominates the distribution of the test scores in the comparison schools in all cases for which the simple average was significant. Figures 1 and 2 show one example of how the distributions transformed, for Vadodara year 2. While the distributions are not distinguishable in year 1, they are very different in year 2, with the distribution in the treatment schools clearly dominating the distribution in year 2.

Table 9 provides estimates of changes in the probability of a student mastering individual competencies.[15] Estimates in this table suggest that, for math, the biggest effect of the Balsakhi program was on grade one and two competencies: in Vadodara, for example, the program increased the fraction of children who mastered the competencies of the first grade in math by 4.0 percentage points in the first year, and 7.3 percentage points in the second year. In Mumbai the effect was 4.5 percentage points and 13.1 percentage points, respectively, in year 1 and 2. This last number represents a 40% increase in the number of children who are able to pass the grade one competencies (the fraction of children who can pass the competencies of grade one in the posttest in year 2 is 34%). The effect on the fraction of children demonstrating knowledge of grade three competencies is much smaller.

In language, the most important effect seems to have been to help children master the

[15]To save space, these estimates are presented only for the lagged dependent variable specification. The difference in differences specification delivers very similar results.

competencies of grade two. This is not surprising, since many children had already demonstrated knowledge of grade one competencies. The effect of the program thus appears to be the strongest on the easiest competencies not already mastered by many pupils. These results correspond well with a stated goal of the program, to work with children on basic competencies they had not yet acquired.

The CAL program affected only math competencies, and seems to have had an equal effect on the number of children able to pass grade one and grade two competencies (about 13 percentage points for each in year 2). It also affected grade three competencies, especially in year 3 (it increased the fraction of students that had achieved grade three competencies by 7.9 percentage points, when only 1.3% of fourth-grade students passed these competencies in the pretest. Among the comparison schools in the posttest, only 5.4% passed these competencies). The CAL program, unlike the Balsakhi program, has the potential to help children at all levels.

To illustrate the size of the gains throughout the distribution, Figure 3 plots the posttest scores as a function of the pretest score rank in the overall distribution (using a Fan locally weighted regression), for treatment and comparison schools in year 2 (both cities and grades are pooled). Children do on average better on the posttest in treatment schools than in comparison schools for any level of pretest achievement. However the difference is largest for children who were initially doing poorly, and the two curves seem to join for children with high initial test scores.

Poor initial scorers registered the largest gains, and were also most likely to be sent to the balsakhi. Figure 4 plots, for the distribution of pretest scores, the difference in test-score gain between treatment and comparison students (the solid line) and the probability of a treatment child being sent to the balsakhi in year 2 (the dashed line)[16]. The treatment effect estimates are obtained as the difference between the functions presented in Figure 3. Assignment probability is estimated using a locally weighted regression. Table 10 (panel A for Balsakhi, and C for CAL) summarizes these patterns by showing the results broken down by third of the initial test score distribution in year 2 in Vadodara[17]. For the balsakhi program, the effect is about twice as large for the bottom third than for the top third (0.42 standard deviation, versus 0.22 standard

[16] Unfortunately, we have no data on assignment to the balsakhi in year 1.

[17] The results are very similar for other years and city, and are available from the authors. This table focus on the year 2 Vadodara results, because these are the children who were followed the following year

deviation). The probability to be assigned to the balsakhi is respectively 0.22 and 0.09 in these two groups. For the CAL program, the impact is also a little bit higher for the bottom third, but the difference is much smaller (0.41 versus 0.32 standard deviation in math in the bottom versus top groups).

The magnitude of the effect of the Balsakhi program at various levels of the pretest score distribution follows closely the probability of being assigned to the balsakhi. This result leads to our next question: How much of the program effect is directly attributable to the children who visited the balsakhi, and how much is indirect (i.e., benefitting the children whose classmates visited the balsakhi)? The fact that both the program impact and the probability of being assigned to a balsakhi decline with a child's position in the test score distribution suggests that the impact of the program may be due to those who were actually assigned to the balsakhi (otherwise, one would see a positive treatment effect even for children with very low assignment probability). However, an alternative explanation for this pattern is that the direct (or indirect) effects of the Balsakhi program are lower for children with higher pretest scores, in ways that track the decrease in the probability of assignment. This question is further investigated in the next section.

5.5 Long Run Impacts

An important consideration in the evaluation of educational interventions is whether or not the changes generated by the interventions last beyond the period in which the intervention is administered. In Vadodara, we were able to track students who received the Balsakhi or the CAL program in the fourth grade in during the 2002-03 academic year in to their fifth grade year, a year in which they no longer received assistance from the program (and also after they have left primary school to either drop out or enter middle school). We are also using the result of children who had benefitted from the Balsakhi program in grade 3, and are enrolled in grade 4 (but do not benefit from the balsakhi program any more at this stage, since there was a gap in the program after the end of the experiment). We are able to track a substantial fraction of children, (see table 6A: the attrition rates is only 20%, both for treatment and comparison children). The attrition rate remain comparable in grades 3 and 4 and the characteristics of the students who attrited are not different that that of those who did not attrit.

Panels B and D of table 10 show the treatment effect estimates using the difference between the August 2002 pretest and a posttest administered to students in March of 2004 at the end of their fifth grade year for the balsakhi and CAL program. For the balsakhi program, the average effect becomes insignificant. However, the effect for the bottom third of the children, which had been the biggest to start with, remains significant, and is around 0.10 standard deviation both for math and for language. For the CAL program, the effects on math also diminish, but they are still significant, on average for for all ability groupings.

These results are very important, since it gives us some indication that the effect we obtained at the end of the enrollment in the program are not artificial and temporary. However, the gains erode relatively quickly over time (after a year, they are still non trivial–0.10 standard deviation would be a good impact for the program in its first year) but they are divided by four. This suggests that it is probably important to follow up the programs continuously.

6 Inside the Box: Direct and Indirect Effects

Estimating equations (1) and (2) generates estimates of the average impact of the program on all children whose school-grade 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 and these may affect other children. This indirect effect could 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 (peer effect).

6.1 Statistical Framework

The ideal experiment to separate the direct and indirect effects of remedial education would have identified the children who would work with the balsakhi in all schools, before randomly assigning treatment and comparison groups (and not allowed substitution after the initial allocation). The balsakhi effect could then be estimated by comparing children designated for the balsakhi in the treatment group with their peers in 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 feasible in this setting, since teachers were not prepared to assign the children in the abstract, without knowing whether or not they were going to get a balsakhi.

We do know, however, that the assignment to the balsakhi group was based in part on pretest score, and that a maximum of twenty children per school in Vadodara, and twenty per class in Mumbai were assigned to a balsakhi. We use these facts to implement two different empirical strategies to disentangle direct and indirect effects.

6.1.1 Exploiting Assignment Probabilities

The first strategy is directly inspired by Figure 4, which suggests that the effect of the program closely tracks the probability of assignment to the balsakhi.

We start by predicting assignment probability in the treatment schools as a flexible function of the rank in the pretest score distribution.

$$P_{ijg} = (\pi_1 + \pi_2 Q_{ij} + \pi_2 Q_{ij}^2 + \pi_3 Q_{ij}^3 + \pi_4 Q_{ij}^4) * D_{ijg} + \omega_{ijg} \tag{3}$$

where P_{ijg} is a dummy indicating that the child was assigned to the program (i.e., worked with the balsakhi) and Q_{ij} is the child's rank in the initial test score distribution.[18]

Denote by M_{ij} the vector $[1, Q_{ij}, Q_{ij}^2, Q_{ij}^3, Q_{ij}^4]$.

We then estimate the treatment effect as a function of the same variables, interacted with the treatment status of the schools.

$$y_{ijgPOST} = \theta y_{ijgPRE} + M_{ij}\lambda + (D_{ij} * M_{ij})\mu + \epsilon_{ijg} \tag{4}$$

Equation 3 and 4 form the first stage and the reduced form, respectively, of the following structural form equation:

$$y_{ijgPOST} = \theta y_{ijgPRE} + \gamma D_{ijg} + \delta P_{ijg} + M_{ij}\alpha + \epsilon_{ijg} \tag{5}$$

The four instruments allow us to estimate γ and δ. The identification assumption is that γ and δ are both constant. Under the maintained assumption that the indirect treatment effect γ is

[18] We use class rank in Vadodara, as each treatment classroom received a balsakhi, and school-grade rank in Mumbai, as balasakhis were assigned to treatment school-grades.

constant, an overidentification test allows us to test whether the remedial education treatment effect δ is indeed constant.

This strategy relies on the assumptions that the indirect treatment effect of the program does not vary with the initial test score of the child. The posttest score can be related to the pretest score in any way, but the treatment effect cannot vary with pretest score in an unrestricted way. If, for example, the indirect treatment effect declined in a way that exactly tracked how the assignment probability changes with the test score, we would mistakenly conclude that the program has no indirect effect.

Since this assumption could be violated, to complement the first strategy, we implement another strategy which does not rely on this assumption.

6.1.2 Exploiting the Non-Linearity in Assignment Rules

This strategy exploits the discrete change in assignment probability at rank 20 in a given class. It estimates the direct remedial education effect and indirect class size or tracking effects for children whose test scores could place them either below rank 20 or above rank 20, depending on their class size. Estimating these parameters does not require any assumption about the constancy or the regularity of the direct and indirect effect at rank 20. The effect is estimated for children who are close to rank 20.

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 pretest, the rank in the class, and a variable indicating whether the child is among the bottom 20 children in his class.

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

where S_{ijg} is the number of students in the class or the school, y_{ijgPRE} is the score of the child at the pretest, 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 a treatment school) predicts assignment to the balsakhi in the whole sample.

$$P_{ijg} = (\pi_1 + \pi_2 Z_{ijg} + X_{ijg}\pi_3) * D_{ijg} + \epsilon_{ijg} \tag{7}$$

We can then regress the posttest scores on the same variables (controlling for pretest score), and examine whether being one of the bottom 20 children is associated with a bigger effect for those whose school is in the treatment group:

$$y_{ijgPOST} = \pi_4 + \pi_5 Z_{ijg} * D_{ijg} + \pi_6 D_{ijg} + \pi_7 Z_{ijg} + X_{ijg}\pi_8 + (X_{ijg} * D_{ijg})\pi_9 + \epsilon ijg \tag{8}$$

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

$$y_{ijgPOST} = \alpha + \beta P_{ijg} + \gamma D_{ijg} + \mu Z_{ijg} + X_{ijg}\kappa + (X_{ijg} * D_{ijg})\lambda + \epsilon_{ijg} \tag{9}$$

where P_{ijg} and D_{ijg} are the independent variables of interest and D_{ijg} and $Z_{ijg} * D_{ijg}$ are the excluded instruments. The identification assumption underlying this estimation strategy is that the only reason 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, test score, and the rank of the child. The only identification assumption is therefore that the *treatment effect* does not vary discontinuously at rank 20. Under this assumption, this equation allows us to estimate the direct and indirect effect of the program for children whose test scores place them in the neighborhood of the rank 20 in the class.

6.2 Results

Columns 1 to 3 in Table 11 show the first stage and the reduced form for the second estimation strategy (equations 7 and 8).

The fact that a child has a rank lower than 20 in his class predicts assignment to the balsakhi remedial education group, even after controlling continuously for his rank, score at the pretest and the number of students in his class. Not surprisingly, because some schools in Mumbai were not assigned a balsakhi, all coefficients are smaller in Mumbai. In columns 4 to 6, we present the reduced form estimates for test score gain. The coefficient on the interaction between the dummy for being among 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 bigger if the child is more likely to be assigned to the balsakhi.

In Table 12, we present instrumental variables estimates of the direct and indirect impact of being in a balsakhi group, using the two strategies described above. The first 3 columns use the treatment dummy (equal to 1 for every child assigned to the treatment group) and this dummy interacted with the pretest score, its square, cube and quartic, as instruments for the balsakhi school and balsakhi assignment. The last lines in the table show the F statistic for the excluded interactions, which are highly significant, and the p-value for the overidentification test.[19]

Based on these results, we cannot reject the hypothesis that being in a balsakhi school has no effect for children who were not themselves sent to the balsakhi: the effect of the program is concentrated on children who were indeed assigned. The effect on these children is large: they gain 0.6 standard deviations in overall test scores (which is over half of the test score gain a comparison child realizes from one year of schooling). The overidentification test indicates that we cannot reject the hypothesis that the treatment effect is constant: 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.

Columns 4 to 6 present the estimate of the program effect using the discontinuity in the assignment rule at rank 20. Once again, based on these estimates, we cannot reject the hypothesis that the program had no effect on children who were not sent to the balsakhi. The point estimates of the direct effect (the effect of visiting a balsakhi) are even larger than above (1 standard deviation), but they are also less precise and cannot be statistically distinguished from the estimate in columns 1 to 3.

Both strategies lead to the same conclusion: the direct effect of the Balsakhi program is very large, and the reduction in class size induced by the program had no indirect effect on children who stayed in the regular class. The second strategy shows that this is true for middle-ranked

[19]We omitt to present the coefficient of the first stage equation, since it is graphically presented in Figure 4.

children, who benefit from a Balsakhi program but not from a reduction in class size with the regular teacher. This helps rule out the possibility that our results are explained by the fact that a reduction in class size is beneficial to a low-scoring child, but not a high-scoring child.

Since the average class size is 45 (though the average student has 63 students in his class), the Balsakhi program effectively halved class size. We can therefore compare the effect of doing more of the same in the schools—represented by the group of twenty students who worked with the regular teacher–with that of providing focused remedial education. The results suggest that class size reductions would be much more effective if done through a balsakhi type program, involving separating children of different levels, than by simply dividing a large class into two smaller classes. What these estimates do not tell us is whether the same impact on the scores of the low scoring children could be obtained by hiring a regular teacher to work with the low scoring children.

Table 13 investigates whether the direct and indirect impacts of the Balsakhi program vary with school characteristics.

The first characteristics we consider is class size. While the size of the balsakhi group is always 20, the size of the class that remains with the teacher depends on the original class size. The reduction in the class size for balsakhi–attenders is thus larger in big schools, while the proportional reduction in the size of the class for the children who do not attend the Balsakhi program is smaller in big schools. We divide the schools into two groups, those with more than 40 students and those with fewer. In small schools, the balsakhi group is actually larger than the size of the group that remains in the regular classroom. We may therefore expect a larger effect of the Balsakhi program in large schools. The prediction on the indirect effect is ambiguous, as it depends on the functional form of the class size effect.

As expected, the effect of the Balsakhi program appears to be about twice as large in large schools than in small schools, though the estimates are noisy. The effect of the program on unassigned children is smaller in large schools. The coefficient on (balsakhi school * big school) is negative, and significant in panel A. This estimate also suggests that non-attending children in small schools may have benefited from the class size reduction (the effect is 0.2 standard deviations, and is significant in the combined sample). This suggests that class size reduction may be effective if it results in very small class size (less than 20) for the regular teacher. In such small classes, teachers may be able to change the way in which they teach.

The two other characteristics we consider, variance in initial test score and average test score of the bottom 20 children, are meant to capture possible benefits of tracking. However, neither the variance in initial test score nor the average test score of the bottom 20 children appears to influence either the direct or indirect effects.

7 Cost Benefit Analysis

In seeking to improve the academic performance of schoolchildren, governments could potentially hire additional teachers, hire balsakhis or put computers in classrooms.

Since we do not detect any effect of reducing class size on test scores, hiring new teachers (who are paid at least 10 times more than balsakhis) does not appear to be a cost-effective strategy. Even using the most optimistic estimate of reduced class size of a 0.2 standard deviations gain, hiring balsakhis would be several times more cost effective than hiring new teachers.[20]

A more interesting exercise is to compare the cost of one year of the Balsakhi program with one year of the CAL program. The cost per student per year of the Balsakhi program is 107 rupees, or approximately 2.25 dollars. The recurring expenditures of the CAL program are 367 rupees, but once the start-up costs of the computers and software are included (assuming five-year depreciation), the program costs 722 rupees. Thus, using the estimates from Table 8, we can calculate the relative cost effectiveness of each program. CAL increases the math score by 0.41 standard deviations and the overall test score by 0.25 standard deviations whereas the Balsakhi program increases the math score by 0.31 and the total score by 0.28. Since CAL costs 6.7 times as much as the Balsakhi program per student, the Balsakhi program is 5 times more cost effective for math and 7.7 times more cost effective for the total score.

The cost per student of the remedial education program is about $2.25 per child per year. The computer-assisted learning program (including the cost of the computers) costs about $15 per child per year. In terms of cost per standard deviation, the remedial education program appears extremely cost effective. Kremer, Miguel and Rebecca Thornton (2004) provide cost per standard deviation for a range of programs. The cost of the most cost effective program they

[20] Of course, we need to remember that this is an argument only about the *marginal* teacher. The Pratham model is easy to replicate precisely because it takes advantage of the existing government machinery.

consider (a children incentive programs) was between \$1.77 and \$3.53 per 0.1 standard deviations (depending on the region). Using the same assumptions (and the figure of the average effect for a child in the treatment school), the remedial education program cost between \$0.89 and \$1.79 per 0.1 standard deviations (depending on the year). The Balsakhi program thus stands up as the most cost effective program for learning improvement.

Moreover, the cost of the Balsakhi programs and those of the girls' scholarship program in Kremer, Miguel and Thornton (2004) are mainly due to payment to teachers or to students, and these transfers should therefore probably not be considered as real costs. This is not the case for the CAL program, whose main cost is computers. Moreover, the cost would be higher in places which do not have reliable electricity supply. When doing this comparison, we have to keep in mind that, in the Balsakhi program, the effect is stronger at the bottom of the distribution, while the other programs (including Computer-Assisted Learning) affect equally the entire distribution of children, or affect most children at the top. The overall assessment of the program therefore depends on how improvements at different places in the distribution are valued. Despite its popularity, and the effectiveness that we demonstrate in this study, computer-assisted learning may not be the most cost-effective intervention to improve the quality of education in India at this stage. However, turning computers already in the schools to productive use, as Pratham did in this program, is clearly a very cost effective proposition and, according to our results, would lead to improvements in learning.

8 Conclusion

This paper reports the results of a remedial education and a computer-assisted learning programs. The remedial education program has already shown that it can be brought to scale, since it has already reached 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 0.14 standard deviations in the first year, and 0.28 in the second year, at a cost of about two dollars per child per year. We also estimate that children who were directly affected by the program improved their test scores by at least 0.6 standard deviations in the second year, while children remaining in the regular classroom did not benefit.

A computer-assisted learning program provided each child in the fourth grade with two hours of shared computer time per week, in which students played educational games that reinforced mathematics skills. Contrary to what has been found in developed-country settings, the program was also very effective, increasing math scores by 0.36 standard deviations the first year, and 0.54 the second year.

These results show that it is possible to dramatically increase the quality of education in urban India, a very important result since a large fraction of Indian children cannot read when they leave school. However, we also find that education is not likely to improve if schooling resources are simply increasing without changing the way teaching is conducted.

References

Abadie, Alberto (2002) "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variables Models." Journal of the American Statistical Association 97: 284-292

Angrist, Joshua D. (1996) "Conditioning on the Probability of Selection to Control Selection Bias," NBER Technical Paper #181

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

Angrist, Joshua and Victor Lavy (2002). "New Evidence on Classroom Computers and Pupil Learning" The Economic Journal. 112(October): 735-765.

Abhijit Banerjee, Rukimini Banerji, Esther Duflo, Rachel Glennerster, Stuti Khemani, Sendhil Mullainathan and Marc Shotland. (2005) "The Impact of Information, Awareness, and Participation on Learning Outcomes." Manuscript, MIT Department of Economics

Banerjee, Abhijit, Suraj Jacob and Michael Kremer. (2004) "Promoting School Participation in Rural Rajasthan: Results from Some Prospective Trials." MIT Department of Economics Working Paper.

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: 795-813

Fan, Jianqing (1992). "Design-Adaptive Non-Parametric Regression," Journal of the American Statistical Association. 87(420): 998-1004.

Glewwe, Paul, Michael Kremer and Sylvie Moulin. (2003) "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. (2004) "Retrospective vs. Prospective Analyses of School Inputs: The Case of Flip Charts in Kenya." Journal of Development Economics. 74(1): 251-268.

Hanushek, Eric A. (1986) "The Economics of Schooling: Production and Efficiency in Public Schools" Journal of Economic Literature. 24(3): 1141-1177.

Hanushek, Eric A. (1995) "Interpreting Recent Research on Schooling in Developing Countries" World Bank Research Observer. 10(2): 227-246.

Kremer, Michael (2003) "Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons" American Economic Review. 93(2): 102-106.

Kremer, Michael and Ted Miguel (2003) "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." Econometrica. 72(1): 159-217.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. (2004) "Incentives to Learn" Manuscript. Havard University Department of Economics.

Kremer, Michael, Karthik Muralidharan, Nazmul Chaudhury, Jeffrey Hammer, and F. Halsey Rogers (2004) "Teacher Absence in India: A Snapshot" Forthcoming in Journal of the European Economic Association

Krueger, Alan and Cecilia Rouse (2004) "Putting Computerized Instruction to the Test: A Randomized Evaluation of a "Scientifically-based" Reading Program". NBER Working Paper No. 10315.

Krueger, Alan and Diane M. Whitmore (2001) "The Effect of Attending Small Class in Early Grades on College Test-Taking and Middle School Test Results: Evidence from Project STAR" The Economic Journal 111(January): 1-28.

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

World Bank (2004) World Development Report 2004: Making Services Work for the Poor. Oxford University Press: New York, NY.

Christel Vermeersch and Michael Kremer (2004) "School Meals, Educational Achievement, and School Competition: Evidence from a Randomized Evaluation" World Bank Policy Research Working Paper No. 3523, November.

Table 1: Sample Design

	Grade	Study Group	Number of Schools	Number of Divisions	Number of Children	Number of Schools Assigned a Balsakhi	Number of Children With Balsakhi
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Mumbai							
Year 1	Three	Balsakhi	32	70	2592	-	-
		No Balsakhi	35	65	2182	-	-
Year 2	Three[1]	Balsakhi	39	74	2530	28	636
		No Balsakhi	38	79	2943	-	-
	Four[2]	Balsakhi	38	77	2812	27	688
		No Balsakhi	39	71	2460	-	-
B. Vadodara							
Year 1	Three	Balsakhi	48	78	2595	-	-
		No Balsakhi	48	80	2539	-	-
	Four	Balsakhi	48	72	2395	-	-
		No Balsakhi	49	77	2669	-	-
Year 2	Three	Balsakhi	61	101	3146	61	951
		No Balsakhi	61	93	2906	-	-
	Four	Balsakhi and CAL	28	44	1415	28	454
		Balsakhi; no CAL	26	42	1457	26	445
		CAL; no Balsakhi	27	44	1435	-	-
		No CAL, no Balsakhi	30	47	1638	-	-
		Balsakhi, not in CAL study	7	9	293	7	111
		No Balsakhi, not in CAL study	4	4	125	-	-
Year 3	Four	CAL	56	82	3131	-	-
		No CAL	55	81	2814	-	-

Notes: This table gives the number of treatment and comparison schools, classrooms (or "division"), and children in the study.

1. The number of schools in column 3 is the number of schools that were intended to be treated. 28 schools were actually treated in year 2 in standard 3.
2. The number of school in column 3 is the number of schools that were intended to be treated. 27 schools were actually treated in year 2 in standard 4.

Table 2: Summary Statistics, Vadodara, Balsakhi Program

	PRE TEST			MID TEST			POST TEST		
	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Vadodara, Year 1									
Grade 3									
Math	0.028	0.000	0.028	0.666	0.541	0.125	0.434	0.254	0.181
			(0.085)			(0.092)			(0.102)
Verbal	0.088	0.000	0.088	0.962	0.851	0.110	0.874	0.715	0.159
			(0.091)			(0.102)			(0.106)
Observations	2595	2539	56	2285	2174	111	2122	2108	14
Grade 4									
Math	-0.044	0.000	-0.044	0.363	0.259	0.104	0.254	0.092	0.162
			(0.081)			(0.096)			(0.096)
Verbal	-0.044	0.000	-0.044	0.763	0.674	0.089	0.707	0.621	0.086
			(0.080)			(0.101)			(0.108)
Observations	2395	2669	-274	2175	2402	-227	1962	2234	-272
Vadodara, Year 2									
Grade 3									
Math	0.040	0.000	0.040	1.389	0.935	0.454	1.698	1.259	0.438
			(0.074)			(0.120)			(0.116)
Verbal	0.026	0.000	0.026	1.451	1.028	0.423	1.245	0.999	0.246
			(0.082)			(0.107)			(0.103)
Observations	3146	2906	240	2843	2609	234	3027	2792	235
Grade 4									
Math	0.053	0.000	0.053	1.005	0.594	0.411	1.201	0.856	0.346
			(0.077)			(0.069)			(0.085)
Verbal	0.084	0.000	0.084	1.132	0.710	0.422	0.919	0.614	0.305
			(0.082)			(0.082)			(0.087)
Observations	3165	3198	-33	2953	2969	-16	3053	3078	-25

Notes: This table gives the mean normalized test score for pre-, and posttest scores for treatment and comparison students in Vadodara. Standard errors of the difference, corrected for clustering, are given in parentheses. The normalized test score is obtained by subtracting the mean pretest score of the comparison group, and dividing by the standard deviation of scores of the the pretest comparison group.

Table 3: Summary Statistics, Mumbai

	PRE TEST			POST TEST			
	Treatment	Comparison	Difference	Treatment	Comparison	Difference	IV
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Mumbai, Year 1							
Grade 3							
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)	
Observations	2592	2182	410	2417	2027	390	
Mumbai, Year 2							
Grade 3							
Math	-0.070	0.000	-0.070	1.509	1.333	0.176	0.276
			(0.087)			(0.155)	(0.240)
Verbal	0.025	0.000	0.025	0.898	0.831	0.067	0.105
			(0.082)			(0.091)	(0.142)
Observations	2530	2943	-413	2337	2731	-394	
Grade 4							
Math	0.053	0.000	0.053	0.995	0.678	0.317	0.494
			(0.076)			(0.111)	(0.154)
Verbal	0.083	0.000	0.083	0.641	0.513	0.127	0.198
			(0.071)			(0.069)	(0.097)
Observations	2812	2460	352	2635	2290	345	

Notes: This table gives the mean normalized test score for pre-, and posttest scores for treatment and comparison students in Mumbai. Standard errors of the difference, corrected for clustering, are given in parentheses. The normalized test score is obtained by subtracting the mean pretest score of the comparison group, and dividing by the standard deviation of scores of the the pretest comparison group.

Table 4: Summary Statistics: Vadodara

	PRE TEST			MID TEST			POST TEST		
	CAL	No CAL	Difference	CAL	No CAL	Difference	CAL	No CAL	Difference
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Vadodara, Year 2									
Math	-0.054	0.000	-0.054	0.765	0.707	0.058	1.128	0.810	0.318
			(0.076)			(0.081)			(0.087)
Verbal	-0.009	0.000	-0.009	0.867	0.872	-0.006	0.719	0.709	0.010
			(0.083)			(0.095)			(0.093)
Observations	2850	3095	-245	2671	2886	-215	2741	2991	-250
Vadodara, Year 3									
Math	0.127	0.000	0.127	0.301	0.066	0.235	0.811	0.236	0.575
			(0.073)			(0.076)			(0.088)
Verbal	0.118	0.000	0.118	0.048	-0.011	0.059	0.121	0.019	0.102
			(0.079)			(0.082)			(0.080)
Observations	3129	2807	322	2963	2642	321	2906	2613	293

for the CAL program. Standard errors of the difference, corrected for clustering, are given in parentheses. The normalized test score is obtained by subtracting the mean pretest score of the comparison group, and dividing by the standard deviation of scores of the the pretest comparison group.

Table 5: Tests for First Order Stochastic Dominance Among Test-score Distributions

	Year 1				Year 2				Year 3	
	Grade 3		Grade 4		Grade 3		Grade 4		Grade 4	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
Vadodara: Balsakhi v Non-Balsakhi										
Test of Equality: p-value	0.9035	0.1090	0.6040	0.2315	0.6790	0.0035	0.2345	0.0000		
Test for FOSD: p-value	0.7145	0.9235	0.3035	0.8055	0.7130	0.9535	0.7980	0.9715		
Check: Test Whether Comparison FO	0.4735	0.0585	0.7250	0.1130	0.3160	0.0010	0.1155	0.0000		
Vadodara: CAL v Non-CAL										
Test of Equality: p-value							0.8310	0.0735	0.1850	0.0000
Test for FOSD: p-value							0.4395	0.8265	0.9335	0.9690
Check: Test Whether Comparison FOSD Treatment: p-value							0.5305	0.0385	0.0865	0.0000
Mumbai										
Test of Equality: p-value	0.6785	0.1645			0.9915	0.3150	0.5095	0.0135		
Test for FOSD: p-value	0.5295	0.6895			0.6250	0.7315	0.9665	0.8745		
Check: Test Whether Comparison FO	0.3275	0.0800			0.6500	0.1490	0.2540	0.0080		

Note: The test for FOSD tests the hypothesis that the treatment distribution first order stochastically dominates the comparison distribution.

Table 6A : Attrition Patterns-Balsakhi

	Year 1			Year 2			2 Years (pre year 1-post year 2) Student matched			One year out (pre year 2-post year 3)		
	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
A. Vadodara												
Grade 3, All												
Percent Attrition	0.182	0.170	0.013	0.038	0.039	-0.001	0.321	0.345	-0.025	0.206	0.196	0.010
			(0.020)			(0.008)			(0.024)			(0.015)
Difference in Score at Pretest	-0.131	-0.260	0.129	-0.045	-0.058	0.013	-0.056	-0.079	0.022	0.048	0.010	0.037
Attriters-Stayers			(0.096)			(0.134)			(0.072)			(0.069)
Grade 4, All												
Percent Attrition	0.181	0.163	0.018	0.035	0.038	-0.002				0.193	0.207	-0.014
			(0.021)			(0.008)						(0.028)
Difference in Score at Pretest	-0.190	-0.168	-0.022	-0.281	0.046	-0.327				-0.120	0.034	-0.154
Attriters-Stayers			(0.080)			(0.118)						(0.081)
B. Mumbai												
Grade 3, All												
Percent Attrition	0.070	0.075	-0.004	0.077	0.073	0.005	0.255	0.250	-0.006			
			(0.015)			(0.010)			(0.009)			
Difference in Score at Pretest	-0.146	-0.274	0.128	-0.330	-0.193	-0.137	-0.445	-0.594	0.149			
Attriters-Stayers			(0.169)			(0.129)			(0.151)			
Grade 4, All												
Percent Attrition				0.063	0.070	-0.006						
						(0.010)						
Difference in Score at Pretest				-0.180	-0.427	0.247						
Attriters-Stayers						(0.139)						

Note: This table describes the attrition patterns and test score results of attriters and stayers for the Balsakhi program. The standard error (corrected for clustering) of differences is given in parentheses. Columns (7)-(9) give attrition rates for the two-year span of the program.

Table 6B: Attrition, CAL Program

	Vadodara, Year 2			Vadodara, Year 3			Vadorara Year 2 pre, year 3 post		
	CAL	No CAL	Difference	CAL	No CAL	Difference	CAL	No CAL	Difference
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Grade 4, All									
Percent Attrition	0.038	0.034	0.005	0.0713	0.0691	0.00216	0.209	0.211	-0.002
			(0.008)			(0.00970)			(0.019)
Difference in Score at	-0.198	-0.025	-0.173	-0.194	0.026	-0.220	-0.142	0.035	-0.177
Pretest Attriters-Stayers			(0.127)			(0.112)			(0.072)

Note: This table describes the attrition patterns and test score results for the CAL program. The standard error (corrected for clustering) of differences is given in parentheses.

Table 7: Estimates of the Impact of the Balsakhi Program, by City and Sample

	Number of Observations	Difference in Difference			Dependent Variable: Test Score Improvement		
		Math	Verbal	Total	Math	Verbal	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mumbai and Vadodara Together Year 1	12855	0.187	0.063	0.135	0.182	0.076	0.138
		(0.047)	(0.057)	(0.047)	(0.046)	(0.056)	(0.047)
Mumbai and Vadodara Together Year 2	21936	0.341	0.163	0.267	0.353	0.187	0.284
		(0.071)	(0.053)	(0.062)	(0.069)	(0.050)	(0.060)
Pooling Both Grades							
Vadodara Year 1	8426	0.195	0.104	0.161	0.189	0.109	0.161
		(0.060)	(0.059)	(0.058)	(0.057)	(0.057)	(0.057)
Vadodara Year 2	11950	0.347	0.226	0.312	0.371	0.246	0.331
		(0.077)	(0.065)	(0.073)	(0.073)	(0.061)	(0.070)
Vadodara Year 2 Oral Test	1286	0.261	0.169	0.240	0.267	0.177	0.247
		(0.073)	(0.077)	(0.071)	(0.062)	(0.061)	(0.060)
Mumbai Year 2	9986	0.327	0.038	0.175	0.324	0.069	0.188
		(0.145)	(0.089)	(0.115)	(0.145)	(0.081)	(0.112)
Mumbai Year 2 Specification Check	9986	0.285	0.063	0.173	0.287	0.086	0.184
		(0.112)	(0.067)	(0.088)	(0.113)	(0.062)	(0.087)
Two Year 01-03							
Mumbai Pretest Year 1 to Posttest Year 2	3188	0.629	0.136	0.394	0.612	0.185	0.407
		(0.162)	(0.134)	(0.133)	(0.141)	(0.094)	(0.106)
Vadodara Pretest Year 1 to Posttest Year 2	3425	0.271	0.150	0.229	0.282	0.181	0.250
		(0.117)	(0.093)	(0.104)	(0.094)	(0.079)	(0.088)

Notes: This table gives the difference in difference and value added specification for the Balsakhi program, for different groups and years. Standard errors (corrected for clustering) are given in parentheses. The dependent variable is the normalized test score.

Table 8: Differences in Differences Estimate of the Impact of the CAL Program, by Year

	Number of Observations	Difference in Differences Math	Difference in Differences Language	Difference in Differences Overall	Value Added Specification Math	Value Added Specification Language	Value Added Specification Overall
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
A. Effect of the CAL Program							
Vadodara Both Years	11251	0.373	-0.043	0.175	0.389	-0.027	0.187
		(0.080)	(0.084)	(0.086)	(0.074)	(0.082)	(0.083)
Vadodara Year 2	5732	0.365	0.014	0.215	0.346	0.013	0.207
		(0.083)	(0.073)	(0.078)	(0.076)	(0.069)	(0.074)
Vadodara Year 3	5519	0.434	-0.031	0.201	0.466	-0.010	0.218
		(0.069)	(0.044)	(0.051)	(0.067)	(0.042)	(0.050)
B. Balsakhi and Computer Assisted Learning Program: Main Effects and Interactions (Vadodara, Year 2)							
CAL	5732	0.412	0.023	0.246	0.408	0.017	0.242
		(0.096)	(0.083)	(0.090)	(0.087)	(0.084)	(0.087)
Balsakhi		0.319	0.199	0.281	0.371	0.229	0.315
		(0.121)	(0.111)	(0.118)	(0.112)	(0.104)	(0.112)
CAL*Balsakhi		-0.111	-0.030	-0.078	-0.145	-0.021	-0.087
		(0.160)	(0.144)	(0.153)	(0.141)	(0.134)	(0.141)

Notes: Panel A gives the difference in difference and value added specification for the CAL program, for different years. Panel B gives the main effects and interactions of the Balsakhi and CAL programs. Standard errors (corrected for clustering) are given in parentheses. The dependent variable is the normalized test score.

Table 9: Lagged Dependent Variable Specification for Competencies by Grade, by City and Year

	Math Competencies for				Verbal Competencies for		
	Grade 1	Grade 2	Grade 3	Grade 4	Grade 1	Grade 2	Grade 3
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
A. Balsakhi Program							
Vadodara							
Year 1							
Both Grades	0.038	0.013	0.026		0.034	0.028	0.012
	(0.023)	(0.007)	(0.017)		(0.023)	(0.022)	(0.017)
Grade Three	0.048	0.016	0.029		0.015	0.031	0.018
	(0.029)	(0.008)	(0.017)		(0.032)	(0.027)	(0.017)
Grade Four	0.034	0.011	0.028		0.060	0.032	0.013
	(0.029)	(0.011)	(0.027)		(0.029)	(0.030)	(0.025)
Year 2							
Both Grades	0.074	0.065	0.023		0.022	0.068	0.032
	(0.021)	(0.019)	(0.009)		(0.015)	(0.020)	(0.015)
Grade Three	0.072	0.064	0.022		0.035	0.017	0.023
	(0.030)	(0.023)	(0.010)		(0.026)	(0.028)	(0.021)
Grade Four	0.080	0.070	0.027		0.012	0.118	0.043
	(0.026)	(0.028)	(0.013)		(0.015)	(0.027)	(0.021)
Mumbai							
Year 1							
Grade Three	0.045	0.023	-0.003		0.021	0.049	0.030
	(0.032)	(0.031)	(0.021)		(0.015)	(0.028)	(0.030)
Year 2							
Both Grades	0.131	0.077	0.093	0.116	0.023	0.067	0.045
	(0.040)	(0.057)	(0.031)	(0.039)	(0.021)	(0.031)	(0.040)
Grade Three	0.136	0.003	0.060	0.058	0.003	0.019	0.011
	(0.062)	(0.081)	(0.034)	(0.038)	(0.034)	(0.043)	(0.049)
Grade Four	0.119	0.141	0.108	0.139	0.028	0.099	0.055
	(0.050)	(0.076)	(0.046)	(0.049)	(0.019)	(0.039)	(0.053)
B. CAL Program, Vadodara							
Year 2							
Grade Four	0.125	0.128	0.038		-0.006	0.010	0.014
	(0.025)	(0.027)	(0.013)		(0.015)	(0.029)	(0.022)
Year 3							
Grade Four	0.148	0.111	0.077		0.000	-0.089	0.009
	(0.021)	(0.019)	(0.012)		(0.023)	(0.026)	(0.019)

Notes: This table presents the lagged dependent variable specification for the math and verbal portions of th exam. The dependent variable is the fraction of students who have mastered the competencies associated with the given grade. Standard errors, corrected for clustering, are given in parentheses.

Table 10: Results by initial level--persistence of results

	Number of Observations	Probability of assignment to balsakhi	Difference in Difference			Dependent Variable: Test Score Improvement		
			Math	Verbal	Total	Math	Verbal	Total
	(1)		(2)	(3)	(4)	(5)	(6)	(7)
PANEL A Balsakhi, Year 2								
All Children	11950	0.159	0.348	0.227	0.313	0.371	0.246	0.331
			(0.077)	(0.065)	(0.073)	(0.073)	(0.061)	(0.070)
Bottom Third	4053	0.217	0.468	0.317	0.427	0.469	0.317	0.425
			(0.087)	(0.074)	(0.084)	(0.088)	(0.074)	(0.084)
Middle Third	3874	0.176	0.407	0.213	0.340	0.374	0.240	0.339
			(0.093)	(0.073)	(0.083)	(0.082)	(0.069)	(0.080)
Top Third	4023	0.086	0.214	0.187	0.217	0.229	0.174	0.216
			(0.082)	(0.085)	(0.081)	(0.076)	(0.076)	(0.077)
PANEL B: Balsakhi, After one year								
All Children	9925	0.165	0.030	0.014	0.023	0.053	0.033	0.040
			(0.051)	(0.047)	(0.045)	(0.047)	(0.041)	(0.041)
Bottom Third	3356	0.227	0.088	0.104	0.102	0.096	0.097	0.103
			(0.045)	(0.041)	(0.040)	(0.045)	(0.038)	(0.040)
Middle Third	3226	0.182	0.050	-0.049	0.002	0.021	-0.024	0.001
			(0.063)	(0.059)	(0.052)	(0.056)	(0.054)	(0.052)
Top Third	3343	0.086	-0.005	0.024	0.007	0.015	0.006	0.009
			(0.072)	(0.071)	(0.064)	(0.069)	(0.062)	(0.061)
PANEL C: CAL, Year 2								
All Children	5732		0.365	0.014	0.215	0.346	0.013	0.207
			(0.083)	(0.073)	(0.078)	(0.076)	(0.069)	(0.074)
Bottom Third	1962		0.414	0.080	0.276	0.423	0.085	0.277
			(0.107)	(0.090)	(0.102)	(0.106)	(0.089)	(0.102)
Middle Third	1844		0.341	-0.023	0.183	0.316	0.005	0.183
			(0.088)	(0.082)	(0.082)	(0.081)	(0.081)	(0.082)
Top Third	1926		0.319	-0.026	0.169	0.266	-0.033	0.146
			(0.086)	(0.089)	(0.085)	(0.073)	(0.081)	(0.078)
PANEL C CAL: After one year								
All Children	4694		0.098	-0.078	0.009	0.093	-0.072	0.009
			(0.053)	(0.054)	(0.050)	(0.045)	(0.048)	(0.045)
Bottom Third	1588		0.088	-0.005	0.041	0.109	0.007	0.050
			(0.050)	(0.054)	(0.048)	(0.047)	(0.047)	(0.046)
Middle Third	1511		0.102	-0.132	-0.018	0.081	-0.109	-0.018
			(0.061)	(0.073)	(0.058)	(0.055)	(0.070)	(0.058)
Top Third	1595		0.114	-0.098	0.007	0.075	-0.104	-0.011
			(0.079)	(0.074)	(0.073)	(0.072)	(0.064)	(0.068)

Table 11: Disentangling Balsakhi and Class Size Effects

	Balsakhi Assignment			Improvement in Test Scores Pre to Post		
	Mumbai	Vadodara	Both	Mumbai	Vadodara	Both
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
A. First Stages and Reduced Form						
Treatment School	0.185	0.476	0.463	0.404	0.615	0.232
	(0.075)	(0.046)	(0.037)	(0.242)	(0.177)	(0.129)
Treatment * Rank <20	0.078	0.181	0.146	0.128	0.157	0.179
	(0.024)	(0.023)	(0.020)	(0.073)	(0.078)	(0.060)
Treatment * Rank	-0.007	0.000	-0.001	-0.004	0.003	0.001
	(0.003)	(0.001)	(0.001)	(0.008)	(0.005)	(0.004)
Treatment * Pretest Score	-0.062	-0.087	-0.090	-0.016	-0.093	-0.056
	(0.024)	(0.014)	(0.012)	(0.094)	(0.086)	(0.062)
Treatment * Number of Students	0.004	-0.004	-0.003	-0.007	-0.007	-0.002
	(0.003)	(0.001)	(0.001)	(0.008)	(0.004)	(0.003)
Rank <20				-0.079	-0.029	-0.120
				(0.050)	(0.057)	(0.042)
Rank				0.008	0.004	0.003
				(0.006)	(0.004)	(0.003)
Pretest Score				-0.338	-0.342	-0.332
				(0.074)	(0.060)	(0.047)
Number of Students				0.007	-0.002	-0.003
				(0.005)	(0.003)	(0.003)

Notes: Column (1)-(3) present the first stage for the estimation strategy to measure balsakhi and class size effects. The dependent variable is a dummy for whether a child visited a Balsakhi. Columns (4)-(6) present the reduced form for this strategy. The dependent variable is normalized test scores. In both specifications, standard errors (corrected for clustering) are given in parentheses.

Table 12: Estimation of the Direct and Indirect Effect of the Balsakhi Program

	Improvement in Test Score: 2SLS Regressions					
	Using f(Pre-Test Score)*Balsakhi as Instrument			Using Rank<20*Treatment as Instruments		
	Mumbai	Vadodara	Both	Mumbai	Vadodara	Both
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Balsakhi School	-0.029	0.133	0.056	0.220	0.193	-0.127
	(0.085)	(0.106)	(0.068)	(0.623)	(0.307)	(0.236)
Saw a Balsakhi	0.574	0.614	0.606	1.477	0.880	1.102
	(0.240)	(0.292)	(0.189)	(0.803)	(0.456)	(0.333)
Treatment * Rank				0.007	0.002	0.001
				(0.016)	(0.005)	(0.004)
Treatment * Pre-test score				0.164	-0.016	0.060
				(0.190)	(0.099)	(0.093)
Treatment * Number of students				-0.019	-0.003	0.000
				(0.014)	(0.004)	(0.004)
F-stat (first stage)	29.491	78.037	87.586			
p-value	0.000	0.000	0.000			
Over Id Test 2: p-value	0.598	0.477	0.476			

Notes: This table presents instrumental variables estimates of the direct and indirect effect of the Balsakhi program. The dependent variable is the difference between normalized posttest and normalized pretest scores. Standard errors, corrected for clustering, are given in parentheses.

Table 12: Estimation of the Direct and Indirect Effect of the Balsakhi Program

	Improvement in Test Score: 2SLS Regressions					
	Using f(Pre-Test Score)*Balsakhi as Instrument			Using Rank<20*Treatment as Instruments		
	Mumbai	Vadodara	Both	Mumbai	Vadodara	Both
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Balsakhi School	-0.029	0.133	0.056	0.220	0.193	-0.127
	(0.085)	(0.106)	(0.068)	(0.623)	(0.307)	(0.236)
Saw a Balsakhi	0.574	0.614	0.606	1.477	0.880	1.102
	(0.240)	(0.292)	(0.189)	(0.803)	(0.456)	(0.333)
Treatment * Rank				0.007	0.002	0.001
				(0.016)	(0.005)	(0.004)
Treatment * Pre-test score				0.164	-0.016	0.060
				(0.190)	(0.099)	(0.093)
Treatment * Number of students				-0.019	-0.003	0.000
				(0.014)	(0.004)	(0.004)
F-stat (first stage)	29.491	78.037	87.586			
p-value	0.000	0.000	0.000			
Over Id Test 2: p-value	0.598	0.477	0.476			

Notes: This table presents instrumental variables estimates of the direct and indirect effect of the Balsakhi program. The dependent variable is the difference between normalized posttest and normalized pretest scores. Standard errors, corrected for clustering, are given in parentheses.

Table 13: Disentangling Balsakhi and Class Size Effects: Instrumental Variable Estimates with Interactions

Dependent Variable:	Improvement in Test scores: Pre to Post								
	Mumbai			Vadodara			Both		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Panel A: Using Rank<20*Balsakhi as Instrument									
Saw a Balsakhi	1.616	1.523	2.248	0.624	1.713	0.611	0.578	1.368	1.086
	(0.866)	(2.774)	(0.880)	(0.631)	(1.554)	(0.591)	(0.392)	(1.323)	(0.464)
Saw Balsakhi*Big School	0.787			-0.138			1.236		
	(0.559)			(1.287)			(0.702)		
Saw Balsakhi*Variance in Pretest Score		-0.053			-0.884			-0.275	
		(2.757)			(1.630)			(1.425)	
Saw Balsakhi*Average Pretest Score of Bottom 20			1.814			-0.247			0.131
			(0.942)			(0.940)			(0.713)
Balsakhi School	-0.022	0.177	0.022	0.428	0.042	0.372	0.050	-0.121	-0.073
	(0.760)	(1.382)	(0.631)	(0.342)	(0.738)	(0.364)	(0.251)	(0.602)	(0.265)
Balsakhi School*Big School	-0.394			-0.252			-0.644		
	(0.309)			(0.585)			(0.309)		
Balsakhi School*Variance in Pretest Score		0.048			0.148			-0.017	
		(1.266)			(0.706)			(0.615)	
Balsakhi School*Average Pretest Score of Bottom 20			-0.653			0.290			0.039
			(0.423)			(0.459)			(0.332)
Panel B: Using Balsakhi*f(Test Score) as Instruments									
Saw a Balsakhi	0.422	-1.405	0.918	0.694	3.217	0.681	0.403	0.831	0.818
	(0.327)	(1.327)	(0.430)	(0.327)	(1.894)	(0.423)	(0.260)	(1.288)	(0.290)
Saw Balsakhi*Big School	0.335			0.034			0.380		
	(0.435)			(0.489)			(0.358)		
Saw Balsakhi*Variance in Pretest Score		2.177			-2.595			-0.209	
		(1.435)			(1.851)			(1.326)	
Saw Balsakhi*Average Pretest Score of Bottom 20			0.296			-0.082			0.155
			(0.589)			(0.628)			(0.413)
Balsakhi School	0.102	0.366	-0.009	0.302	-0.790	0.241	0.191	-0.096	0.074
	(0.102)	(0.517)	(0.108)	(0.213)	(0.766)	(0.219)	(0.103)	(0.474)	(0.102)
Balsakhi School*Big School	-0.287			-0.245			-0.216		
	(0.149)			(0.241)			(0.130)		
Balsakhi School*Variance in Pretest Score		-0.427			0.929			0.161	
		(0.552)			(0.767)			(0.493)	
Balsakhi School*Average Pretest Score of Bottom 20			0.081			0.186			0.061
			(0.149)			(0.268)			(0.132)
F-stat (First Stage)	22.098	3.208	11.098	91.894	3.028	110.202	39.076	5.164	46.311
p-value	0.000	0.015	0.000	0.000	0.018	0.000	0.000	0.000	0.000
Over Id Test 1: p-value	0.990	0.638	0.749	0.032	0.364	0.239	0.183	0.211	0.198
Over Id Test 2: p-value	0.390	0.182	0.685	0.009	0.114	0.319	0.002	0.338	0.295

Notes: This table tests whether the direct and indirect effects of the Balakhi program vary by school characteristics. The dependent variable is difference between pretest and posttest score. Each column within a panel represents a separate regression. Standard errors, corrected for clustering, are given in parentheses.

Table 14: Attendance

	Year 1			Year 2			Year 3		
	Treatment	Comparison	Difference	Treatment	Comparison	Difference	Treatment	Comparison	Difference
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
A. Mumbai									
Grade 3									
RA Attendance	0.861	0.870	-0.009	0.854	0.850	0.007			
			(0.014)			(0.022)			
RA Attendance	0.915	0.925	-0.010	0.893	0.895	-0.004			
			(0.011)			(0.018)			
Observations	2463	1786	-677	2499	2836	337			
Grade 4									
RA Attendance				0.859	0.867	-0.012			
						(0.017)			
RA Attendance				0.886	0.900	-0.023			
						(0.015)			
Observations				2742	2388	-354			
B. Vadodara, Balsakhi Program									
Grade 3									
RA Attendance	0.745	0.764	-0.019	0.735	0.739	-0.005			
			(0.012)			(0.013)			
Observations	2593	2535	-58	3131	2892	-239			
Grade 4									
RA Attendance	0.769	0.759	0.010	0.752	0.743	0.009			
			(0.013)			(0.011)			
Observations	2389	2595	206	3155	3172	17			
C. Vadodara, CAL program									
RA Attendance				0.749	0.743	0.006	0.708	0.684	0.025
						(0.011)			(0.015)
Observations				2826	3082	256	3124	2794	-330

Note: This table reports the effect of the Balsakhi and CAL program on classroom attendance. The dependent variable is the fraction (from 0 to 1) of days the child is recorded as present by a Pratham research assistant.

Appendix Table 1: Summary Statistics: Vadodara Year 1, Balsakhi

	PRE TEST			MID TEST			POST TEST		
	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
GRADE 3									
A. OBSERVATIONS	2595	2539	56	2285	2174	111	2122	2108	14
B. SCORES (PERCENTAGE)									
Math	0.263	0.257	0.006	0.407	0.379	0.028	0.355	0.314	0.041
			(0.019)			(0.021)			(0.023)
Verbal	0.234	0.217	0.017	0.402	0.381	0.021	0.385	0.354	0.031
			(0.017)			(0.020)			(0.020)
C. NORMALIZED TEST SCORES									
Math	0.028	0.000	0.028	0.666	0.541	0.125	0.434	0.254	0.181
			(0.085)			(0.092)			(0.102)
Verbal	0.088	0.000	0.088	0.962	0.851	0.110	0.874	0.715	0.159
			(0.091)			(0.102)			(0.106)
D. PERCENTAGE OF CHILDREN PASSING ALL COMPETENCIES FOR EACH GRADE									
Math Grade 1	0.214	0.195	0.019	0.406	0.350	0.056	0.327	0.273	0.053
			(0.027)			(0.031)			(0.032)
Math Grade 2	0.012	0.015	-0.003	0.045	0.043	0.002	0.036	0.020	0.016
			(0.005)			(0.009)			(0.008)
Math Grade 3	0.043	0.032	0.011	0.133	0.109	0.024	0.092	0.062	0.030
			(0.010)			(0.019)			(0.017)
Verbal Grade 1	0.237	0.209	0.028	0.531	0.524	0.007	0.520	0.497	0.024
			(0.028)			(0.033)			(0.035)
Verbal Grade 2	0.158	0.142	0.017	0.332	0.317	0.015	0.284	0.246	0.037
			(0.023)			(0.028)			(0.031)
Verbal Grade 3	0.038	0.028	0.010	0.131	0.133	-0.003	0.095	0.073	0.022
			(0.011)			(0.022)			(0.020)
GRADE 4									
A. OBSERVATIONS	2395	2669	-274	2175	2402	-227	1962	2234	-272
B. SCORES (PERCENTAGE)									
Math	0.441	0.451	-0.010	0.535	0.511	0.024	0.510	0.473	0.037
			(0.019)			(0.022)			(0.022)
Verbal	0.343	0.352	-0.009	0.516	0.497	0.019	0.504	0.486	0.019
			(0.017)			(0.022)			(0.023)
C. NORMALIZED TEST SCORES									
Math	-0.044	0.000	-0.044	0.363	0.259	0.104	0.254	0.092	0.162
			(0.081)			(0.096)			(0.096)
Verbal	-0.044	0.000	-0.044	0.763	0.674	0.089	0.707	0.621	0.086
			(0.080)			(0.101)			(0.108)
D. PERCENTAGE OF CHILDREN PASSING ALL COMPETENCIES FOR EACH GRADE									
Math Grade 1	0.405	0.447	-0.041	0.561	0.537	0.024	0.506	0.486	0.019
			(0.030)			(0.034)			(0.034)
Math Grade 2	0.061	0.053	0.008	0.106	0.095	0.011	0.085	0.072	0.013
			(0.010)			(0.016)			(0.013)
Math Grade 3	0.109	0.106	0.003	0.224	0.195	0.029	0.173	0.144	0.029
			(0.022)			(0.027)			(0.027)
Verbal Grade 1	0.441	0.458	-0.017	0.684	0.631	0.053	0.725	0.670	0.056
			(0.028)			(0.031)			(0.033)
Verbal Grade 2	0.278	0.315	-0.037	0.488	0.449	0.039	0.450	0.437	0.013
			(0.029)			(0.035)			(0.038)
Verbal Grade 3	0.114	0.122	-0.008	0.264	0.251	0.013	0.217	0.209	0.007
			(0.022)			(0.034)			(0.031)

Notes: This table presents summary statistics for the various testing rounds. Standard errors of differences (corrected for clustering) are given in parentheses. The normalized test score is obtained by subtracting from the raw score the mean of the normalized comparison group score, and then dividing this difference by the standard deviation of the pretest scores of the comparison group. "SCORES (PERCENTAGE)" indicates the share of possible points a child scored on the exam.

Appendix Table 2: Summary Statistics: Vadodara Year 2, Balsakhi

	PRE TEST			MID TEST			POST TEST		
	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
GRADE 3									
A. OBSERVATIONS	3146	2906	240	2843	2609	234	3027	2792	235
B. SCORES (PERCENTAGE)									
Math	0.167	0.160	0.007	0.420	0.335	0.085	0.478	0.396	0.082
			(0.014)			(0.022)			(0.022)
Verbal	0.221	0.216	0.004	0.463	0.391	0.072	0.428	0.386	0.042
			(0.014)			(0.018)			(0.018)
C. NORMALIZED TEST SCORES									
Math	0.040	0.000	0.040	1.389	0.935	0.454	1.698	1.259	0.438
			(0.074)			(0.120)			(0.116)
Verbal	0.026	0.000	0.026	1.451	1.028	0.423	1.245	0.999	0.246
			(0.082)			(0.107)			(0.103)
D. PERCENTAGE OF CHILDREN PASSING ALL COMPETENCIES FOR EACH GRADE									
Math Grade 1	0.258	0.249	0.009	0.591	0.475	0.117	0.616	0.542	0.075
			(0.025)			(0.031)			(0.032)
Math Grade 2	0.016	0.018	-0.002	0.140	0.081	0.059	0.183	0.119	0.063
			(0.005)			(0.019)			(0.023)
Math Grade 3	0.001	0.000	0.001	0.021	0.009	0.013	0.056	0.034	0.023
			(0.001)			(0.004)			(0.010)
Verbal Grade 1	0.508	0.500	0.008	0.813	0.698	0.115	0.776	0.740	0.036
			(0.035)			(0.027)			(0.028)
Verbal Grade 2	0.073	0.088	-0.014	0.323	0.236	0.087	0.331	0.319	0.013
			(0.016)			(0.028)			(0.029)
Verbal Grade 3	0.016	0.009	0.007	0.135	0.120	0.014	0.145	0.120	0.025
			(0.005)			(0.019)			(0.021)
GRADE 4									
A. OBSERVATIONS	3165	3198	-33	2953	2969	-16	3053	3078	-25
B. SCORES (PERCENTAGE)									
Math	0.309	0.297	0.012	0.530	0.435	0.095	0.576	0.496	0.080
			(0.018)			(0.016)			(0.020)
Verbal	0.348	0.331	0.017	0.560	0.475	0.085	0.517	0.456	0.062
			(0.017)			(0.017)			(0.018)
C. NORMALIZED TEST SCORES									
Math	0.053	0.000	0.053	1.005	0.594	0.411	1.201	0.856	0.346
			(0.077)			(0.069)			(0.085)
Verbal	0.084	0.000	0.084	1.132	0.710	0.422	0.919	0.614	0.305
			(0.082)			(0.082)			(0.087)
D. PERCENTAGE OF CHILDREN PASSING ALL COMPETENCIES FOR EACH GRADE									
Math Grade 1	0.470	0.448	0.022	0.739	0.641	0.098	0.779	0.692	0.087
			(0.030)			(0.020)			(0.028)
Math Grade 2	0.055	0.051	0.004	0.231	0.144	0.087	0.292	0.220	0.072
			(0.010)			(0.021)			(0.028)
Math Grade 3	0.015	0.012	0.003	0.052	0.040	0.012	0.110	0.082	0.028
			(0.003)			(0.009)			(0.013)
Verbal Grade 1	0.749	0.739	0.010	0.898	0.811	0.087	0.859	0.846	0.014
			(0.025)			(0.017)			(0.016)
Verbal Grade 2	0.186	0.181	0.005	0.440	0.323	0.117	0.481	0.361	0.120
			(0.020)			(0.026)			(0.028)
Verbal Grade 3	0.082	0.079	0.003	0.256	0.194	0.063	0.224	0.179	0.045
			(0.015)			(0.024)			(0.024)

Notes: This table presents summary statistics for the various testing rounds. Standard errors of differences (corrected for clustering) are given in parentheses. The normalized test score is obtained by subtracting from the raw score the mean of the normalized comparison group score, and then dividing this difference by the standard deviation of the pretest scores of the comparison group. "SCORES (PERCENTAGE)" indicates the share of possible points a child scored on the exam.

Appendix table 3: Summary Statistics: Mumbai Year 1, Balsakhi

	PRE TEST			POST TEST		
	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
GRADE 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 ALL COMPETENCIES FOR EACH GRADE						
Math Grade 1	0.326	0.337	-0.012	0.397	0.357	0.040
			(0.038)			(0.039)
Math Grade 2	0.126	0.147	-0.021	0.211	0.195	0.017
			(0.025)			(0.036)
Math Grade 3	0.024	0.023	0.001	0.089	0.091	-0.003
			(0.007)			(0.022)
Verbal Grade 1	0.856	0.837	0.019	0.937	0.913	0.025
			(0.025)			(0.018)
Verbal Grade 2	0.486	0.473	0.013	0.577	0.526	0.050
			(0.045)			(0.042)
Verbal Grade 3	0.517	0.470	0.047	0.631	0.584	0.047
			(0.039)			(0.039)

Notes: This table presents summary statistics for the various testing rounds. Standard errors of differences (corrected for clustering) are given in parentheses. The normalized test score is obtained by subtracting from the raw score the mean of the normalized comparison group score, and then dividing this difference by the standard deviation of the pretest scores of the comparison group. "SCORES (PERCENTAGE)" indicates the share of possible points a child scored on the exam.

Appendix table 4: Summary Statistics: Mumbai Year 2, Balsakhi

	PRE TEST			POST TEST			
	Balsakhi	No Balsakhi	Difference	Balsakhi	No Balsakhi	Difference	Implied Difference
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
GRADE 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	0.049
			(0.016)			(0.028)	(0.043)
Verbal	0.351	0.344	0.007	0.588	0.569	0.018	0.028
			(0.022)			(0.025)	(0.039)
C. NORMALIZED TEST SCORES							
Math	-0.070	0.000	-0.070	1.509	1.333	0.176	0.276
			(0.087)			(0.155)	(0.240)
Verbal	0.025	0.000	0.025	0.898	0.831	0.067	0.105
			(0.082)			(0.091)	(0.142)
D. PERCENTAGE OF CHILDREN PASSING ALL COMPETENCIES FOR EACH GRADE							
Math Grade 1	0.137	0.167	-0.030	0.421	0.339	0.082	0.128
			(0.025)			(0.043)	(0.064)
Math Grade 2	0.082	0.090	-0.008	0.412	0.412	0.001	0.001
			(0.015)			(0.053)	(0.083)
Math Grade 3	0.003	0.006	-0.003	0.136	0.099	0.037	0.058
			(0.003)			(0.023)	(0.035)
Math Grade 4	0.007	0.013	-0.006	0.123	0.088	0.035	0.054
			(0.004)			(0.024)	(0.038)
Verbal Grade 1	0.653	0.648	0.005	0.820	0.817	0.004	0.006
			(0.036)			(0.022)	(0.034)
Verbal Grade 2	0.165	0.147	0.017	0.388	0.363	0.024	0.038
			(0.022)			(0.033)	(0.052)
Verbal Grade 3	0.137	0.131	0.005	0.317	0.307	0.010	0.015
			(0.021)			(0.034)	(0.053)
GRADE 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	0.122
			(0.019)			(0.027)	(0.038)
Verbal	0.555	0.530	0.025	0.721	0.683	0.038	0.059
			(0.021)			(0.021)	(0.029)
C. NORMALIZED TEST SCORES							
Math	0.053	0.000	0.053	0.995	0.678	0.317	0.494
			(0.076)			(0.111)	(0.154)
Verbal	0.083	0.000	0.083	0.641	0.513	0.127	0.198
			(0.071)			(0.069)	(0.097)
D. PERCENTAGE OF CHILDREN PASSING ALL COMPETENCIES FOR EACH GRADE							
Math Grade 1	0.300	0.240	0.060	0.474	0.387	0.087	0.136
			(0.031)			(0.036)	(0.053)
Math Grade 2	0.245	0.243	0.003	0.554	0.464	0.090	0.140
			(0.023)			(0.055)	(0.081)
Math Grade 3	0.042	0.041	0.001	0.241	0.171	0.069	0.108
			(0.010)			(0.033)	(0.047)
Math Grade 4	0.074	0.063	0.011	0.335	0.242	0.093	0.144
			(0.013)			(0.035)	(0.050)
Verbal Grade 1	0.825	0.796	0.029	0.923	0.900	0.023	0.035
			(0.022)			(0.014)	(0.020)
Verbal Grade 2	0.338	0.333	0.005	0.576	0.512	0.064	0.099
			(0.027)			(0.034)	(0.049)
Verbal Grade 3	0.355	0.317	0.038	0.532	0.485	0.047	0.074
			(0.031)			(0.033)	(0.050)

Notes: This table presents summary statistics for the various testing rounds. Standard errors of differences (corrected for clustering) are given in parentheses. The normalized test score is obtained by subtracting from the raw score the mean of the normalized comparison group score, and then dividing this difference by the standard deviation of the pretest scores of the comparison group. "SCORES (PERCENTAGE)" indicates the share of possible points a child scored on the exam.

Appendix Table 5: Estimates of the Impact of the Balsakhi Program, by City and Sample

	Number of Observations	Difference in Difference			Dependent Variable: Test Score Improvement		
		Math	Verbal	Total	Math	Verbal	Total
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Grade 3							
Vadodara Year 1	4230	0.179	0.082	0.142	0.179	0.102	0.152
		(0.092)	(0.090)	(0.089)	(0.086)	(0.085)	(0.085)
Vadodara Year 2	5819	0.397	0.222	0.341	0.418	0.233	0.354
		(0.111)	(0.094)	(0.103)	(0.107)	(0.089)	(0.100)
Mumbai Year 1	4429	0.163	0.060	0.118	0.161	0.086	0.127
		(0.072)	(0.072)	(0.067)	(0.075)	(0.066)	(0.067)
Mumbai Year 2	5063	0.369	0.051	0.193	0.348	0.071	0.193
		(0.195)	(0.128)	(0.158)	(0.197)	(0.118)	(0.155)
Mumbai Year 2 Specification Check	5063	0.276	0.073	0.168	0.259	0.076	0.162
		(0.149)	(0.097)	(0.121)	(0.152)	(0.092)	(0.121)
Grade 4							
Vadodara Year 1	4196	0.201	0.121	0.173	0.190	0.114	0.166
		(0.075)	(0.074)	(0.073)	(0.072)	(0.076)	(0.073)
Vadodara Year 2	6131	0.280	0.213	0.265	0.307	0.240	0.289
		(0.087)	(0.073)	(0.080)	(0.078)	(0.068)	(0.074)
Mumbai Year 2	4923	0.435	0.098	0.269	0.456	0.140	0.299
		(0.125)	(0.087)	(0.104)	(0.124)	(0.074)	(0.097)
Mumbai Year 2 Specification Check	4923	0.403	0.104	0.257	0.429	0.149	0.291
		(0.099)	(0.066)	(0.081)	(0.101)	(0.058)	(0.078)
One Year Out							
Vadodara Pretest Year 2 to Posttest Year 3 Grade 3	4834	0.027	0.007	0.017	0.042	0.010	0.024
		(0.082)	(0.076)	(0.073)	(0.079)	(0.065)	(0.068)
Vadodara Pretest Year 2 to Posttest Year 3 Grade 4	5091	0.027	0.024	0.028	0.058	0.055	0.053
		(0.055)	(0.056)	(0.053)	(0.046)	(0.049)	(0.046)

Notes: This table gives the difference in difference and value added specification for the Balsakhi program by standard, for different groups and years. Standard errors (corrected for clustering) are given in parentheses. The dependent variable is the normalized test score.