The Impact of Mother Literacy and Participation Programs: Evidence from a Randomized Evaluation in India

Rukmini Banerji, James Berry, and Marc Shotland*

September 2015

Abstract

We report the results of a randomized evaluation of three programs designed to improve mother literacy and the home learning environment among households in India. Households were assigned to receive either (1) adult literacy (language and math) classes for mothers, (2) training for mothers on how to enhance their children's learning at home, or (3) a combination of the two interventions. We find that the three interventions resulted in statistically-significant improvements in mother math and language test scores, while the adult literacy and combined interventions also improved women's empowerment. All three interventions had statistically-significant impacts on children's math scores.

JEL Classifications: C93, D13, I21, O15

Keywords: adult education, education inputs, field experiments

*Banerji: Pratham and ASER Centre, rukmini.banerji@pratham.org; Berry: Cornell University, jimberry@cornell.edu; Shotland: Abdul Latif Jameel Poverty Action Lab, shotland@mit.edu. We thank Annie Duflo for collaboration and insights on the study design and implementation. We also thank Eric Edmonds, Karthik Muralidharan, Hugo Reis, and seminar participants at the NBER Education and Children's Meeting, UC Santa Cruz, Cornell University, and the Northeast Universities Development Consortium Conference for useful comments and suggestions. We are also grateful to Jessica Chan, Nandini Gupta, Ravi Gupta, Rachna Nag Chowdhuri, and Nikhil Wilmink for superb work coordinating the field activities in Rajasthan and Bihar. Kevin Kho, Laurel Wheeler, and Raymond Lee provided excellent research assistance. This research was funded by the International Initiative for Impact Evaluation. All errors are our own.

1 Introduction

One of the most important outcomes from the initial stages of education is basic literacy—typically defined to include reading, writing, and basic mathematics.[1] Literacy is considered essential for social participation and forms a building block for higher-level skills (UNESCO, 2015). For this reason, adult literacy is often used as an indicator of the educational progress of a nation (see, e.g., United Nations Development Programme, 2014; United Nations, 2014).[2] To improve basic levels of literacy, policy discussions typically focus on investments that can be made to improve primary schooling systems. However, even if successful, such policies would never reach the 781 million adults around the world who are currently illiterate and past their formal schooling years (UNESCO, 2015).

In a number of developing countries, policy makers have turned to adult literacy campaigns to improve the literacy skills of the current generation of adults. These campaigns often focus on women in order to address gender inequality in literacy and to promote women's empowerment. While government spending on adult literacy education is typically dwarfed by spending on formal schooling, such programs are nonetheless ambitious. In 2009, the government of India launched the Saakshar Bharat adult literacy campaign, with a goal of bringing literacy skills to 70 million adults, including 60 million women. In South Africa, the Kha Re Gude Mass Literacy Programme was launched in 2008, aiming to reach 4.7 million adults by 2012. Most of the program's beneficiaries are women (UNESCO, 2015).

In addition to potentially improving adult learning and women's empowerment, adult literacy education programs may have important impacts on parental involvement in child education (Abadzi, 2003; UNESCO, 2005b). Such programs could, for example, provide parents with the skills to facilitate child learning at home and foster an interest in children's education, thereby increasing child learning. The potential importance of parental education and the home learning environment in children's education outcomes has been long recognized, yet there is little rigorous evidence of effective policies that could influence child learning at home (Avvisati et al., 2010).[3]

While adult literacy education may provide parents with the skills and interest to participate more in their children's education, when parents lack experience with the formal schooling system, as is the case with many parents in developing countries, these programs may not be sufficient to improve the quality or quantity of their participation. If a key constraint to

[1] See UNESCO (2005a) for a discussion of the definition of literacy.

[2] Although adult literacy is typically defined to include math (UNESCO, 2005a; UNESCO, 2015), measurement of literacy for national statistics often includes only reading and writing (United Nations Development Programme, 2014).

[3] In the United States, the "Coleman Report" of 1966 (Coleman et al., 1966) found that family background was more strongly correlated with child test scores than school characteristics and sparked a debate about the relative importance of families and schools. More recent studies have continued to highlight the strong relationship between home inputs and education outcomes (see, e.g., Todd and Wolpin, 2007).

facilitating child learning at home is a lack of understanding of how parents can participate, these parents may instead need parental participation training, providing direct instruction on how to be involved. For example, parental participation programs could train parents to monitor their children's completion of homework, or to understand the symbols teachers use to grade homework. Adult literacy and participation programs may also be complementary if parents need skills and the experience of learning, along with instruction on participation, in order to support child learning.

Although adult literacy and participation interventions could serve as important tools to improve adult and child education outcomes in developing countries, there are real challenges that may limit their effectiveness. Adults in developing countries have many competing demands on their time, and without immediate monetary benefits, these programs may not be seen as valuable relative to work, home production, or leisure activities (Wagner, 2000; Abadzi, 2003). Even when adults are motivated to participate regularly, part-time programs may not be extensive enough for formerly-illiterate parents to develop the skills necessary to support child learning. Whether these programs are effective in improving adult and child education outcomes is thus an empirical question.

To address this question, we conducted a randomized evaluation of three programs in rural India designed to improve mother literacy—again defined to include reading, writing, and basic mathematics—and the home learning environment. Villages were randomly assigned to one of four groups. In the first group, mothers in the village were offered the Mother Literacy (ML) intervention, consisting of daily language and math classes. In the second, mothers were given the Child Home Activities and Materials Packet (CHAMP) intervention: materials, activities, and training were provided each week to promote enhanced involvement in their children's education at home. In the third, mothers were offered both the literacy and home-learning interventions (ML-CHAMP). The fourth group served as a control with no intervention. The evaluation was carried out in 480 villages in the states of Bihar and Rajasthan.

We provide evidence that the programs had positive effects on the learning outcomes and on empowerment of mothers. The ML, CHAMP, and ML-CHAMP interventions had statistically-significant impacts on combined mother language and math test scores of 0.10, 0.05, and 0.13 standard deviations, respectively. The ML and ML-CHAMP interventions also had statistically-significant impacts on an aggregate index of women's empowerment. Across our empowerment measures, the programs had the largest effects on measures of capability to carry out life tasks related to literacy and on beliefs and attitudes related to women's education. The ML-CHAMP intervention also had a significant impact on mother involvement in household decisions.

Turning to chidren's outcomes, we find that the ML, CHAMP, and ML-CHAMP interventions resulted in statistically-significant improvements in child math scores of 0.04, 0.04, and 0.06 standard deviations, respectively. The only significant impacts on language scores came from

the combined ML-CHAMP intervention. We also find that the programs significantly increased mother participation in child learning. The evidence supports our theory that improving the home learning environment, particularly the mothers' literacy skills and participation, results in increased child learning. However, we cannot rule out the possibility that the interventions affected children directly through either child attendance in mother literacy classes or through direct child participation in CHAMP activities without the involvement of the mother.

Our study's key policy implication is that mother literacy and participation programs can be used as tools for policy makers who are interested in influencing mother learning and empowerment, as well as child learning and the home learning environment. In the case of the mother literacy program, we observe relatively low participation and argue that evaluation of innovations to increase participation and sustain learner interest would be a useful area for future research. Across our main outcome measures, the impacts of ML-CHAMP are approximately the sum of the individual ML and CHAMP impacts, suggesting that more intensive interventions could have stronger impacts.

Our study contributes to several strands of literature. First, it is the only randomized evaluation of a developing-country adult literacy program that we know of. In this context, there are several evaluations that attempt to establish the impact of such programs on adults and children, although much of the research suffers from methodological limitations. Some studies find significant impacts of adult literacy programs on adult learning using ex-post comparisons with non-participants (Carron, 1990 ; Ortega and Rodriguez, 2008). Aker et al. (2011) conduct a randomized evaluation of a program that provided cell phones to participants in existing adult education classes in Niger and find significant impacts of the cell phone program on writing and math scores. However, there is no evaluation of the adult literacy program, *per se*. Research on the effects of adult literacy programs on children's outcomes is even more sparse, and existing studies also rely on retrospective selection of a comparison group (Aoki, 2005; Abadzi, 2003).[4]

Second, to our knowledge, our study provides the first large-scale randomized evaluation of a parental participation program in a developing country, where parents may have little or no experience with formal education. Bekman (1998) evaluates a Turkish program that trained mothers to help educate their children at home. Using a matching procedure to construct a comparison group, the study finds large effects of the program on child learning. In developed country contexts, rigorous evidence on the effects of participation programs is also limited, and the existing evidence is mixed (Avvisati et al., 2010). A recent large-scale randomized evaluation of a program in a poor school district in France that aimed to enhance parental involvement in the education of adolescent children found significant positive effects on parental involvement and on student behavior in school, though there were no detectable effects on child test scores (Avvisati et al., 2014).

[4]Although there are numerous evaluations of adult literacy programs in the U.S., much of the research also suffers from methodological limitations (Beder, 1999) .

The remainder of our paper is structured as follows. Section 2 discusses the context and programs. Section 3 covers the evaluation design and data collection. Section 4 describes the results. We provide discussion of the results in Section 5 and conclude in Section 6.

2 Program Description

2.1 Context

The three interventions were designed and implemented by Pratham, a non-governmental organization (NGO) specializing in child literacy and numeracy.[5] Pratham conducted the interventions in two blocks (district subdivisions) of the Purnia district in Bihar and two blocks of the Ajmer district in Rajasthan. Bihar and Rajasthan were selected by Pratham based on the low literacy levels of the two states. According to the latest census, these states have the lowest female literacy rates in India at roughly 53 percent each (Government of India, 2011a). Children's education outcomes are similar between the two states. For example, 48 percent of rural children in grades 3 to 5 can read at a grade-1 level or above in both states, just below the national average of 54 percent (ASER Centre, 2013). The intervention districts within each state were selected because of existing Pratham programs and infrastructure in those areas. Within the intervention districts, the study blocks were selected because they did not have any pre-existing Pratham programs.

While education outcomes are similar between the two states, Rajasthan ranks substantially higher than Bihar along several other key dimensions of economic development. Bihar has the lowest GDP per capita of any state in the country, and while Rajasthan is below the national average, its per capita GDP is double that of Bihar (Central Statistics Office, 2013). Similarly, 67 percent of households in Rajasthan have electricity, about the national average, while Bihar ranks last among Indian states, with only 16 percent of households having electricity (Government of India, 2011b).

As of the baseline survey, households in our sample displayed similar patterns of wealth and education across the two states. Appendix Table 1 displays means of wealth, demographic, and education variables in the Rajasthan and Bihar samples. Households in the Rajasthan sample had substantially more assets and were more likely to have electricity. The average education level for a mother in our sample was under one year in both states, and mothers in both states had similar scores on our baseline test, with Rajasthan mothers scoring slightly higher. Child learning levels were also slightly higher in the Rajasthan sample. Women in the Rajasthan sample spent more time working inside and outside of the home (74 hours per week compared to 51).

[5]For more information, see http://www.pratham.org.

Running the interventions in multiple states in different areas of the country aids external validity of the evaluation. Although the interventions were identical in both states, they were implemented by different local teams and supervised by separate state-level Pratham leadership. And while learning levels in both states were similar, the differences in wealth and preexisting activities of the mothers presented distinct implementation challenges in each area.[6]

2.2 Intervention Descriptions

Each of the three interventions was implemented for approximately 10 months, between September 2011 and June 2012. Recruitment of mothers for each program was focused on women with children aged 5 to 8. It was hypothesized that the programs would have the greatest effects on children that were just beginning to develop the most basic reading and math skills. These children and their mothers (referred to as "target children" and "target mothers") were identified from a census of communities selected to receive the programs (see Section 3).[7]

The Mother Literacy (ML) intervention consisted of daily literacy classes held in the villages and was modeled after Pratham's community-based classes for children (see Banerjee, et al., 2010). Classes were taught by volunteers who were recruited from within the village and were given several days of training from Pratham staff before setting up classes. Classes were to be held for two hours per day, six days per week, at the time and place identified by mothers to be the most convenient. While ML classes were open to any who wished to attend, Pratham staff and volunteers were given a list of target mothers to recruit into the classes.Volunteers received several refresher trainings and ongoing supervision by Pratham staff throughout the intervention.

Within ML classes, volunteers utilized a version of Pratham's "Read India" methodology to teach basic language and math skills. This approach, previously shown effective in teaching children to read (Banerjee et al., 2010), was modified to suit the interests of adults. Because most mothers started with very low language and math skills, the programs focused on the most basic competencies. Pratham's prior experience with adult literacy interventions indicated that mothers were more interested in learning math, and responding to this demand, the program placed more emphasis on basic math skills than on on language skills. Math activities included counting, number recognition, addition and subtraction, and activities with currency. The language component provided instruction on letter recognition, reading basic words, and writing names. Each week, volunteers also presented short topical paragraphs to facilitate discussion among mothers. The paragraphs contained messages about decisions mothers could make for themselves or their families, particularly in relation to family health, women's employment,

[6] In Appendix A we explore heterogeneity in the program effects on child and mother learning by state.

[7] When the household contained more than one child aged 5-8, one was randomly selected to be a target child.

government programs, and to a lesser extent, the value of educating girls.[8]

The primary goal of the ML classes was to increase language and math skills of mothers. The weekly discussion topics, as well as the opportunity to meet in groups of women, were designed to promote empowerment, particularly mothers' ability and confidence to make decisions for themselves or for their households. Skills taught in the classes could also increase mothers' capabilities to make such decisions.

We further hypothesized that the ML intervention would improve child learning. This could arise if the skills learned, exposure to the learning process, and opportunity for discussion about education increased mother interest and participation in her child's education. The intervention could thereby change the type of participation at home, time spent with children, education assets at home, aspirations for children, or schooling behavior outside of the home. We note that while two of the weekly discussion topics related to the benefits of educating girls, there were no specific program components providing instruction on mother participation in child learning.

The Child and Mother Activities Packet (CHAMP) intervention was designed to directly encourage mothers to be engaged with their children's learning. Unlike the volunteers who implemented the ML intervention, CHAMP implementers were paid Pratham staff who often had prior experience as Pratham volunteers or supervisors. These implementers were given about one week of training by senior Pratham staff and met together several times per month to review upcoming activities. Each CHAMP implementer was responsible for conducting CHAMP in 5 villages.[9]

Once per week during the intervention period, the CHAMP implementer visited each target mother and gave her a worksheet to help her child complete and instruction on how to be involved in child learning. The worksheets focused on first-grade language and math skills. In about half of the sessions, the mother received direct instruction such as encouraging the child to do schoolwork at home, reviewing the child's school notebooks, and discussing progress with the child's school teacher. Some visits also included instruction on activities that the mother could do with her child such as sharing stories, telling time, or supervising simple writing exercises. Each weekly visit lasted approximately 10-20 minutes. When the mother was not available, the CHAMP session was conducted with another adult member of the household when possible.

The primary goal of CHAMP was to improve child learning by involving the mother in the process. Mothers were equipped with more ways in which they could participate in child

[8]Of 34 topics, two related to girls' education.

[9]The staff who implemented CHAMP were also responsible for supervising volunteers in ML. This was primarily for logistical reasons, as the evaluation required multiple interventions to be conducted in the same areas. However, because the ML and CHAMP interventions were distinct, we do not view the multiple roles of these staff as important to our results.

learning and were encouraged to spend more time on child educational activities. Beyond increasing participation of the mother, the program's encouragement of home learning could also result in an increase in education assets available at home. Finally, the experience of participating in child learning could foster a broader interest in child education, resulting in increased aspirations for children and increased child school attendance.

In addition to its effects on children, CHAMP could also improve the learning and empowerment of mothers. Although the intervention did not directly provide math or language instruction to mothers, increased mother exposure to child learning and interactions with CHAMP implementers could result in increased mother learning. Mother empowerment could also increase as a result of the program, through increases in mother learning, or through the experience of participating in child learning.

The combined intervention (ML-CHAMP) included both the ML and CHAMP interventions. The combined intervention was not integrated—ML and CHAMP were simply conducted in the same villages with the same target group of mothers. As in the separate interventions, mother literacy classes were conducted by volunteers, and CHAMP visits were conducted by separate Pratham staff. The ML-CHAMP intervention was designed to test whether both mother exposure to learning from ML and instruction on how to facilitate child learning from CHAMP would be necessary for mothers to engage with children and improve child learning. It also allows us to test whether one program "crowded out" another, in which case the impacts would not increase when both programs were implemented together.

3 Research Design

3.1 Treatment Assignment

In each state, 240 hamlets (village subdivisions) were selected for the randomization. Hamlets were selected based on a target number of households (the approximate size that could support one mother literacy class) and geographic distance from other target locations to limit spillovers. In Rajasthan, where villages are geographically distinct, one hamlet per village was randomly selected among those determined to be large enough to support a mother literacy class. In Bihar, where hamlets may be close to one another (whether in the same village or in different villages), hamlets of the target size were included if they were sufficiently far from other included hamlets.[10] For ease of exposition, we refer to the randomization unit as a "village" throughout.

In each state, the 240 study villages were randomly assigned in equal proportions to the control

[10] In Appendix B we provide more detail on how hamlets were selected in each state.

group or to one of the three treatment groups. Randomization was stratified geographically to allow Pratham to organize its training and monitoring structure based on a known number of program villages in a fixed area. Appendix Figure 1 provides a graphic representation of the stratification procedure. The 240 villages in each state were first divided into geographically proximate "sets" of 20 villages. Each set was then split into two "phases" consisting of 10 villages each. These phases determined the order of the rollout of the programs. Within each set of 20 villages, the Pratham team first initiated the interventions in Phase 1 villages and began in Phase 2 villages approximately three weeks later.[11] These set-by-phase groups of 10 villages formed the strata for the randomization. [12]

Within each village in the sample, a census was first conducted to determine a list of target mothers. Twenty-two mothers of children aged 5 to 8 years old were then randomly selected to be targeted. If there were fewer than 22 such mothers in the village, all mothers were targeted. There were a total of 8,888 target mothers in the study, with an average of 18.5 mothers per village.

3.2 Data collection

Baseline data were collected from households of target mothers at the onset of the interventions in summer 2011, and endline data were collected after approximately one year, in summer 2012. The primary evaluation data consisted of standardized tests of mothers and children, as well as household surveys administered to mothers.

The standardized tests, designed to evaluate a basic set of Hindi language and math skills, were developed by the ASER Centre, Pratham's research arm. These tests were expanded versions of the ASER Centre's standard assessment tool used each year in its national assessments of child learning (ASER Centre, 2013). At baseline, the tests were administered to target mothers, children aged 5 to 8 in their households, other children in grades 1 to 4, and children aged 4 and below who were going to be enrolled in school in the next year. The endline testing included all mothers and children tested at baseline, in addition to the remaining children who were aged 3 or 4 at the baseline. The mother test contained the same questions as the child test, but included several additional word problems and other questions testing practical literacy and numeracy skills. Minor changes were made to the testing instruments between baseline and endline. At baseline a shorter test, designed to quickly assess basic reading and math knowledge, was also administered to other household members.

[11] Apart from the small difference in implementation timing, Phase 1 villages were identical to Phase 2 villages in program administration and content.

[12] Because there were four interventions to be assigned within a stratum of 10 villages, each stratum contained two "remainder" villages. These remainders were balanced within the larger set of 20 villages, such that each intervention was implemented in exactly five villages per set.

The household surveys were administered to target mothers. These surveys were designed to measure women's empowerment outcomes as well as a set of indicators of mother interest and participation in child education. The empowerment module contained questions on mobility outside of the home, capability to carry out life tasks related to literacy, attitudes towards women's education, and involvement in household decisions. Mother interest and participation were measured through modules on participation in learning at home, time use of the mother and target child, presence of educational materials in the household, mother aspirations for her child, and schooling status of children in the household. While most questions appeared in both survey rounds, some additions were made to the endline survey based on observations made during implementation.[13]

Column 1 of Table 1 contains descriptive statistics from the baseline surveys and tests in the control group. Mothers in our sample had low levels of education and similarly low levels of literacy and numeracy. The average years of education in the control group was 0.76, with a median of zero. Baseline test scores for target mothers were relatively low: although the tests focused primarily on first-grade competencies, the average test score in the control group was 30 percent in the language section and 22 percent in the math section. At the time of the survey, 12 percent of mothers in the control group reported experience with adult literacy classes in the past.

Columns 2, 3, and 4 of Table 1 present differences in means of the baseline variables between each treatment group and the control group. One of the 21 variables is jointly significant at the 10 percent level between the three treatment groups and the control group, and one more is significant at the 5 percent level. Out of 63 individual comparisons between treatment and control groups, three are significant at the 10 percent level, and six more are significant at the 5 percent level. On the whole, this suggests that the randomization was successful in creating similar groups.

Approximately 3.5 percent of households reached for surveys and testing at baseline were not reached at endline. Endline child tests are available for 94 percent of children tested at the baseline. We do not observe evidence of differential attrition across treatment groups at the household level, but there is some imbalance of attrition levels among child test-takers between the CHAMP and ML-CHAMP groups and the control group. In Appendix C we present a formal analysis of attrition, including the construction of bounds on the child learning estimates using Lee's (2009) trimming method. Because attrition was low overall, these bounds are relatively tight and do not alter the conclusions from our analysis.

[13]We also administered short questionnaires to fathers at baseline and endline. These questionnaires contained questions on participation in child learning, empowerment of mothers, and support for adult literacy programs. At endline only 57 percent of fathers were reached, and we therefore do not use these surveys in the main analysis.

4 Results

Before the endline data were examined, a pre-analysis plan (PAP) was uploaded to the American Economic Association's Randomized-Controlled-Trial Registry website.[14] The PAP includes the weights used in aggregating test scores, the estimating equation for test scores, and the list of intermediate outcomes to be examined. The analysis presented in this paper follows the PAP as closely as possible. In the few cases where we deviate from the PAP, we discuss the choices and implications below.

4.1 Estimating Equation

Throughout the analysis we utilize the following estimating equation:

$$Y_{1iv} = \beta_0 + \beta_1 ML_v + \beta_2 CHAMP_v + \beta_3 MLCHAMP_v + \beta_4 Y_{0iv} + \delta G_v + \pi H_{iv} + \varepsilon_{iv} \qquad (1)$$

In this equation, Y_{1iv} is the outcome for individual i in village v. ML_v, $CHAMP_v$, and $MLCHAMP_v$ are dummy variables indicating the treatment status of the village. Y_{0iv} is the baseline value of the outcome of interest (when measured). G_v is a dummy for stratum, as described in Section 3.1, while H_{iv} represents a vector of household-level characteristics. ε_{iv} is the individual error term, clustered by village, the level of randomization. Equation (1) is estimated using ordinary least squares.

Following the PAP, we estimate treatment impacts on the mother and child learning outcomes using Equation (1) with a set of pre-specified household characteristics as controls. For other outcomes, we estimate Equation (1) without additional household-level controls.

We present our results in four parts. First, we analyze program take-up. Second, we present effects on mother learning and empowerment. Third, we examine child learning. Finally, we estimate program impacts on mother involvement and interest in child learning.

4.2 Program Take-up

Take-up of the mother literacy classes is analyzed in Table 2. We estimate Equation (1) using dummy variables indicating self-reported knowledge of classes, mother attendance, and child attendance as dependent variables. Compared with the control group, mothers in the ML and ML-CHAMP treatments were approximately 32 and 37 percentage points more likely to report

[14] https://www.socialscienceregistry.org/trials/65

having ever attended the classes, respectively.[15] Children attended the classes as well. Target children were 22 percentage points more likely to attend with their mothers in the ML treatment and 28 percentage points more likely to attend with their mothers in the ML-CHAMP treatment, compared with the control group.

In addition to relatively low take-up overall, attendance in the mother literacy classes was sporadic throughout the year. According to Pratham's administrative records, the average target mother who attended classes between September 2011 and February 2012 attended between 3 and 5 days per month.[16] In March, the harvest season began in both Rajasthan and Bihar, and very few classes were held through the end of the program in June.

As shown in Table 2, the likelihood of attendance was similar in the ML and ML-CHAMP groups. We do observe higher attendance of 4.4 percentage points in the ML-CHAMP group, a difference significant at the 5 percent level (p-value = 0.047). However, we note that the magnitude of the difference is relatively small (about 14 percent of the attendance effect of ML). In addition, Pratham's administrative data do not indicate a significant difference in the number of classes attended in each group (results not shown).

In Table 3 we analyze the determinants of mother take-up of mother literacy classes in the ML and ML-CHAMP treatments. We regress self-reported mother attendance in the classes on a set of variables including household characteristics, child schooling behavior, mother education, time spent working, experience with literacy classes, self-help group membership, baseline mother empowerment and participation in child learning, and state. Column 1 includes both the ML and ML-CHAMP treatment groups, while Columns 2 and 3 run the regression separately using the ML and ML-CHAMP samples.

Household composition could affect demands on a mother's time (in the case of very young children), mother's interest in improving child learning (primary-aged children), and other household resources that could free mother's time (older household members). However, we observe no evidence of a relationship between take-up and our three measures of household composition. We also find no evidence of a relationship between mother work hours and take-up. In the full sample of mothers in ML and ML-CHAMP, mothers with children in school were

[15]We observe that 7 percent of mothers in the CHAMP and control groups attended classes. Virtually all of these mothers were in the Bihar sample. We find no evidence of spillovers across program hamlets: there is no relationship between attendance in the CHAMP and control group hamlets and presence of ML or ML-CHAMP hamlets in the larger village (not shown). The likely reasons for non-compliance are a combination of misunderstanding of the survey question and the government's Saakshar Bharat adult literacy mobilization that occurred during the spring of 2012. Initial promotion of the program occurred in many villages across the study area, and our respondents who participated in this mobilization may have considered this as class attendance. After the initial mobilization, only 11 Saakshar Bharat classes were ever set up, most of these classes were held for less than one week, and classes were spread across all four treatment arms. Thus, although promotion of the program may have impacted responses to the attendance question on the survey, the classes themselves were unlikely to have influenced our other results.

[16]Because these data were collected by Pratham volunteers and were not verified, they may overstate actual attendance.

more likely to attend the classes, but child test scores are not significantly related to take-up.

The relationship between mother education and attendance is non-monotonic: mothers were more likely to attend when they had some exposure to formal schooling, but more years of education made them less likely to attend. In addition, mothers scoring higher on the baseline test were significantly more likely to attend. Because the test covered only the most basic competencies, the test and schooling results are consistent in that mothers with a small amount of education were more likely to attend.

Beyond basic demographic characteristics, experience meeting in groups, agency within households, and pre-existing participation in child learning are strong predictors of attendance in the ML classes. Mothers who attended mother literacy classes in the past and members of a self-help group were 7 and 9 percentage points more likely to attend during the intervention, respectively. Mothers with greater empowerment and participation in child learning, as measured by the baseline empowerment and participation indices described below, were also significantly more likely to attend.

Finally, mothers in Bihar had 13 percent higher take-up than those in Rajasthan. This mirrors Pratham and research staff observations that mothers were on average more motivated and had more time to attend in Bihar.

As shown in Columns 2 and 3 of Table 3, we find few differences in the relationships between these variables and attendance in the ML and ML-CHAMP villages. In fact, we cannot reject the hypothesis that the differences between the coefficients in Columns 2 and 3 are zero (p-value = 0.28). This result suggests that the presence of CHAMP in the village did not appreciably change the composition of mothers who attended ML classes.

CHAMP was a door-to-door intervention where Pratham staff visited each selected mother at home. The CHAMP intervention consisted of up to 30 weekly visits that took place between September 2011 and June 2012. If the mother was not at home, then the CHAMP session would be conducted with another adult in the household. Pratham administrative data are not available for all months and villages, but based on the administrative data we have (covering about half of the intervention), 96 percent of target households were visited at least once during the intervention period. Ninety percent of households were reached for at least half of the scheduled visits. We estimate that households were visited 20 times on average. We have more limited data for whether the CHAMP was conducted with the mother or another member of the household (about seven weeks in ML-CHAMP villages in Bihar), but within subsample of data we have, mothers were present during 85 percent of successful visits.

4.3 Mother Test Scores and Empowerment

Table 4 presents the effect of the interventions on mothers' normalized math, language, and total (math plus language) test scores. For each of the three test score categories, scores are normalized based on the control group mean and standard deviation in each round of testing. Impact estimates are computed using Equation (1), controlling for baseline language and math test scores and the household-level control variables listed in the PAP.[17] Missing values of control variables are coded as zero, and additional dummy variables are included to indicate missing values.

All three programs had statistically significant impacts on mother language, math, and total test scores. The ML program improved mother test scores by 0.067 standard deviations in language, 0.12 standard deviations in math, and 0.097 standard deviations overall. All three estimates are significant at the 1 percent level. The last column in the table presents instrumental variables estimates of the effects of take-up of the program on learning, instrumenting take-up with assignment to the ML treatment. To account for spillovers within households, take-up is defined as either the mother or the child attending a class at least once. Using this specification, the estimated effect of class attendance on test scores equals 0.19 standard deviations for language, 0.35 standard deviations for math, and 0.28 standard deviations overall.

The CHAMP program improved mother test scores as well. Test scores improved by 0.024 standard deviations in language (significant at the 5 percent level), 0.062 standard deviations in math (significant at the 1 percent level), and 0.045 standard deviations overall (significant at the 1 percent level).

The ML-CHAMP intervention had impacts of 0.089, 0.15, and 0.13 standard deviations on language, math, and total test scores, respectively. Each result is significant at the 1 percent level. For each category of test score, the effect of ML-CHAMP is approximately equal to the sum of the effects of the individual interventions, suggesting no substitution or complementarity between ML and CHAMP. Indeed, we cannot reject that the effects are additive for any of the three test score categories.

Panel A of Appendix Table 2 examines robustness of these results to inclusion of household-level control variables. As shown in the table, running the regression without these controls leaves the results virtually unchanged.

To understand how the interventions impacted specific skills, Appendix Table 3 presents effects on a subset of the competencies covered on the test. With few exceptions, these competencies correspond to those tested in the ASER national assessments of reading and language (ASER Centre, 2013).[18] In language, ML and ML-CHAMP had the largest impact on letter recog-

[17] The list of control variables is included in the notes to Table 4. Although the PAP lists average non-target child and adult test scores in four separate age groups, we include only one group due to missing data.

[18] The primary difference between the competencies on this list and the ASER test is that we include two digit

nition, most basic skill tested. CHAMP, by contrast, had similar impacts across the range of competencies tested. In math, all three interventions had the largest impacts on one-digit number recognition. The difference in the pattern of impacts between ML and CHAMP in language could be due to the fact that children were typically exposed to higher level competencies at school than mothers were exposed to in the ML classes, and therefore the CHAMP intervention, which encouraged mother-child interaction, gave mothers more exposure to those higher-level competencies.

We next examine the impacts of the interventions on women's empowerment. As described in Section 2, either the ML or CHAMP intervention could promote mother empowerment by providing mothers with the ability and confidence to make decisions for themselves or for their households.

To measure empowerment, we include 26 variables measured in the endline survey.[19] With the exception of two variables that are excluded because of a wording mistake in the questionnaire, these variables are specified in the PAP.[20] The full list of included variables (as well as impacts on each) is shown in Appendix Table 4. To organize the variables, we group them into domains based roughly on Pitt, Khandker, and Cartwright (2006). The "mobility and networks" group includes five variables measuring whether the mother has recently traveled outside of the village on her own, as well as self-help group membership. The "capability" group captures seven measures of self-reported literacy, having signed her name on official documents, counting change during purchases, knowledge of the National Rural Employment Guarantee Scheme wage, and involvement in self-help group savings. The "beliefs and attitudes" group includes four variables relating to beliefs about own education, spousal education, and daughter's education. The "decisions" group includes nine variables indicating mother involvement in various types of household purchases, as well as involvement in child schooling decisions. Finally, we include one measure of self-reported happiness that we analyze separately.

We aggregate all 26 variables into an index using the method specified in Kling, Liebman, and Katz (2007). To construct the index, we compute the normalized value of each variable based on the control group mean and standard deviation at endline. We then take the average of the non-missing normalized variables and re-normalize the resulting index based on the control group mean and standard deviation. The final index has a mean of zero and standard deviation

addition, while the ASER test includes division.

[19]We note that empowerment and the intermediate outcomes in the analysis are based on self-reports from the mother survey. While we cannot directly verify the responses to these questions, the patterns of impacts we find are inconsistent with surveyor demand effects that would influence responses on all questions related to education, or only on questions relating to program content. We discuss this issue in detail in Appendix D.

[20]The PAP includes one variable indicating whether the mother worked under the National Rural Employment Guarantee Scheme and one variable indicating whether she took advantage of other entitlements. The survey measured whether *anyone* in the household took advantage of these entitlements, and thus does not accurately measure empowerment of the mother. We have therefore not included them in our main empowerment index. Inclusion of these variables does not substantively change the magnitudes or significance levels of the impact estimates (results not shown).

of one in the control group. We control for a similarly-constructed baseline index, where this index includes the 12 variables that were measured at baseline.

The first row of Table 5 displays the impact estimates. We find that the ML intervention increased empowerment by 0.069 standard deviations (significant at the 10 percent level), and that the ML-CHAMP intervention increased empowerment by 0.13 standard deviations (significant at the 1 percent level). The estimated impact of CHAMP is 0.050 standard deviations but is not significant at conventional levels.[21] As with the learning results, the impact of ML-CHAMP is approximately equal to the sum of the individual ML and CHAMP interventions, and indeed we cannot reject additivity.

The remaining rows of Table 5 display impacts on the different categories of empowerment. Based on the categorization outlined above, we create separate indices by category and estimate impacts on each. We find that the ML and ML-CHAMP interventions had the largest effects on capabilities and on beliefs and attitudes, and the ML-CHAMP intervention also had a significant impact on involvement in household decisions. The CHAMP intervention had a significant impact only on beliefs and attitudes. We do not find significant impacts of any of the interventions on mobility and networks or on happiness.

Our empowerment measures, as listed in the PAP, contain several indicators that relate to content of the ML classes. In Appendix Table 5 we examine robustness of our empowerment results to inclusion of these indicators. First, we construct an index that excludes the four measures of capabilities that most closely relate to the ML program content: considering oneself literate, signing one's name on official documents, counting change, and having caught a mistake counting change. Using the 22 remaining empowerment measures, the effect size of ML falls to 0.039 standard deviations and is no longer statistically significant, while the effect of ML-CHAMP is 0.85 and is significant at the 5 percent level. Second, we construct an index that only includes the 5 measures of mother involvement in household purchase decisions and the value of goods the mother can buy alone. This domain receives particular attention in the economics literature (Malhotra et al., 2002), and there was no program content on mother involvement in household purchases. When we construct the empowerment index using only these indicators, the impact of ML is 0.050 standard deviations and falls just short of statistical significance (p-value = 0.16), while the impact of ML-CHAMP equals 0.068 standard deviations and is significant at the 10 percent level. Thus, while the point estimates decline slightly from those found using our main index, we do observe some evidence for impacts on empowerment beyond measures closely related to the content the ML classes.

[21] Results from the endline father questionnaire suggest that these empowerment impacts arose without increases in intrahousehold conflict. Eighty-nine percent of fathers indicated that they supported adult literacy classes for their wives, with significantly higher proportions of support in each of the the intervention groups compared with the control group (not shown). We note, however, that the response rate of this questionnaire was only 57 percent.

4.4 Child Test Scores

Table 6 presents the effects of the treatment groups on child language, math, and total test scores. We again estimate Equation (1), controlling for baseline math and language test scores and a similar set of household-level variables that were used in the mother learning analysis. As specified in the PAP, all children tested at the endline are included in the estimation, including the younger children not tested at baseline. We code missing baseline test scores as zero and include dummy variables to indicate missing values.

Column 1 of Table 6 displays the impacts of the ML intervention on language, math, and combined test scores. The intervention had an impact of 0.036 standard deviations on math scores (significant at the 5 percent level). The estimated effects on language and combined scores are smaller and not statistically significant. Column 6 presents instrumental variables estimates of the effects of literacy class attendance on child learning, instrumenting attendance with assignment to the ML treatment. As with the mother estimation in Table 4, attendance is defined as either the mother or the child attending a class at least once. Using this specification, the estimated effect of class attendance is 0.10 standard deviations for math. Because the intention-to-treat estimates are not significant for language or total test scores, it is not surprising that the corresponding instrumental variables estimates are not significant at conventional levels.

Column 2 of Table 6 displays the impacts of the CHAMP intervention. Similar to ML, CHAMP had an impact of 0.039 standard deviations on math test scores (significant at the 5 percent level). The effect on language scores is smaller and not statistically significant, but the effect on total test scores equals 0.030 and is statistically significant at the 5 percent level.

Column 3 of Table 6 displays the impacts of the combined ML-CHAMP intervention. The estimated effects are 0.045, 0.060, and 0.055 standard deviations for language, math, and total test scores (respectively), and all three estimates are significant at the 1 percent level. As with the mother learning and empowerment results, the impacts of the combined intervention on child learning do not suggest substitution or complementarity between the individual interventions: in all three cases, we cannot reject that the ML-CHAMP effect equals the sum of the ML and CHAMP effects.

In Panel B of Appendix Table 2, we evaluate the robustness of the child learning results to two alternative specifications. The first set of regressions limits the control variables to baseline test scores and stratum dummies, and the second drops observations where no baseline test scores were measured. The results are virtually unchanged under these alternative specifications.

In Appendix Table 6 we examine the effects on a subset of competencies on the test, similar to the exercise for mothers in Appendix Table 3. We focus on math, where the main results show significant impacts of all three interventions. Among the math skills in the table, the ML intervention similar impacts across the range of competencies, with the largest impacts on

two-digit addition. The CHAMP and ML-CHAMP interventions had the largest impacts on one-digit number recognition.

4.5 Mother Participation, Attitudes, and Home Inputs

As described in Section 2, we hypothesized that the programs would impact mother involvement and interest in child education. In this section we present results on mother participation in child education, child and mother time use, education assets in the home, mother aspirations for her child and perceptions of her child's ability, and child school participation.

We start by analyzing mother participation in child education. Our participation measures consist of survey questions on mother involvement in the education of her target child, focusing on monitoring and supervision of child learning. We include nine measures of mother participation, including indicators of school visits, whether the mother helps with homework, whether she has looked at her child's notebook recently, and the frequency with which she talks to her children and others about child education.[22] As with the empowerment measures, we construct an index of mother participation using the nine measures, following the aggregation procedure described above.

As shown in Table 7, we find positive and statistically significant impacts of each program on the mother participation index. The magnitudes are 0.071 standard deviations for ML (significant at the 5 percent level), 0.13 for CHAMP (significant at the 1 percent level), and 0.11 for ML-CHAMP (significant at the 1 percent level). While both ML and CHAMP had statistically significant impacts on the mother looking at the child's notebook, talking to the child about school, and talking to others about the child's studies, CHAMP and ML-CHAMP also had significant impacts on mothers helping with homework.[23]

We next turn to the effects of the programs on mother's and children's time use. To the extent that the interventions encouraged participation in child learning, we hypothesized that they would increase mother time spent directly helping children learn and child time spent learning at home. However, we find limited evidence of these effects. Table 8 presents impact estimates of the programs on mothers' weekly time spent with children, working, and doing other activities. Across all measures, we see little evidence that the programs impacted time use. Mother time spent helping children with homework increased by about 0.1 hours for each of the three interventions, but these estimates are not statistically significant. Table 9 presents the estimated

[22]These measures correspond to those listed in the PAP, with two exceptions. First, the question "What have your child and family members taught you?" was included in the list of participation measures in the PAP, but it refers to household participation in mother learning and is therefore not appropriate for the index. Second, the PAP included educational assets in the mother participation measures, and we have chosen to analyze impacts on these assets separately.

[23]We note that the recall period of the survey questions overlapped with the end of the CHAMP intervention, and thus some of the CHAMP results could be driven by activities undertaken by the mother during the intervention.

impacts of the programs on the target child's weekly time use. As with the mother results, we find few significant impacts across the range of measures. There is suggestive evidence that the CHAMP program resulted in modest increases in time spent on homework: the impact of ML-CHAMP was 0.26 hours per week (significant at the 10 percent level), and the impact of CHAMP was a statistically-insignificant 0.19 hours per week.[24]

Table 10 presents estimated treatment effects on the presence of education assets in the home. These assets include pencils, school books, other books, newspapers and magazines, and slates. The first row of Table 10 displays the impacts on an index of the five types of assets, where the index is constructed in the manner described above. Both CHAMP and ML-CHAMP had significant impacts on the index, with magnitudes of 0.10 and 0.12 standard deviations, respectively. Among the individual asset types, the CHAMP intervention had statistically significant effects on the presence of pencils, school books, other books, and newspapers/magazines, while ML-CHAMP significantly increased the presence of school books, other books, and slates. For the ML intervention, the only statistically significant effect is on the presence of school books.[25]

In Table 11 we turn to a set of indicators that reflect mothers' perceptions of parental involvement in education, aspirations for their children, and perceptions of child reading and math ability. As shown in Panel A, we find several impacts of the programs on perceptions of parental involvement. Mothers in the ML-CHAMP intervention identified more educational responsibilities of parents and more ways that mothers can help children with education (both significant at the 5 percent level), although impacts of the ML and CHAMP interventions on the two measures are not statistically significant. All three interventions also significantly increased the mother's belief that she should be involved in her child's education. By contrast, as shown in Panel B, we find we find little evidence that the programs influenced expectations and aspirations for child educational attainment. Across the four measures examined, the only significant impact is an increase of 0.24 years in the mother's aspiration for her child's grade attainment resulting from the ML-CHAMP intervention (significant at the 10 percent level).

Panel C of Table 11 examines mother perceptions of child language and math ability. Perceptions were based on integer scales of 0 to 4, corresponding to five levels of the ASER language

[24]The time-use results provide suggestive evidence that ML and ML-CHAMP increased work hours by both the mother and child. ML-CHAMP increased mother's livestock work by 0.52 hours per week (significant at the 5 percent level), and effects of both ML and ML-CHAMP on paid/agricultural work are 0.9 to 1.0 hours per week, falling just short of statistical significance at conventional levels. ML-CHAMP also increased the child's time spent working in an household business by 0.29 hours (significant at the 5 percent level). We speculate that the skills taught and positive messages about employment in the mother literacy classes could have given mothers more confidence and skills to work independently and to manage the child's time working. Child labor could also have increased as a result of the the increase in child skills. However, given the lack of statistical significance of many of these estimates, and the large number of coefficients estimated in the time use tables, these results may also be due to chance.

[25]We note, however, that ML provided mothers with small workbooks. Although surveyors were instructed to emphasize that this question referred to school books for children, it is possible that some mothers confused their ML workbooks for school books.

and math tests. We find that the CHAMP and ML-CHAMP interventions significantly increased mother perceptions of her child's language and math scores. When compared to the child's actual ability, however, the CHAMP and ML-CHAMP interventions caused mothers to be overly optimistic: both interventions increased the differences (in absolute value) between mother perceptions and measured child ability in language and math. Our initial hypothesis was that through additional mother involvement, these interventions would *increase* the accuracy of the perception, and thus these results are surprising.[26] One possible explanation for this result is that treated mothers overestimated the impacts of their involvement on child learning.

Finally, Table 12 presents the impacts of the programs on school participation. Across the full sample of children, there is little evidence for these effects. When we examine schooling behavior of all children aged 3 to 14 at baseline, we find no evidence that the individual programs affected current enrollment, expected enrollment in the upcoming school year, or regular attendance, though we do find a small positive impact on school attendance of the combined ML-CHAMP program (2.0 percentage points, significant at the 10 percent level). Using our sample of target children, we find little evidence for impacts on time spent in school, recent absences, or monthly tuition expenditures. On balance, this table shows limited, if any, impacts of the interventions on school participation. We note, however, that enrollment is relatively high among our study population. In qualitative interviews, many parents claimed that the most they could do within the formal schooling system was send their children to school and cited poor quality of government schools and poor access to private schools as barriers they could not overcome.

In the last two rows of Table 12, we explore this issue further by analyzing whether the programs changed schooling behavior for children around the age of initial enrollment. When their children are between 4 and 6 years old, parents do have a choice of whether and when to send their children to aanganwadi (community kindergarten) and when to enroll their children in first grade in the formal schooling system. Within this group, we find no evidence for an an increase in regular school or aanganwadi attendance in the ML or CHAMP treatments and a small increase of 2.2 percentage points in the ML-CHAMP treatment (significant at the 10 percent level). However, when we examine only formal school attendance among this group, we find that the ML program increased attendance by 3.9 percentage points and that the ML-CHAMP program increased attendance by 3.7 percentage points (both significant at the 5 percent level). Because the impacts on formal school attendance are larger than the impacts on school or aanganwadi attendance, ML and ML-CHAMP may have encouraged mothers to shift their children from aanganwadi to first grade at earlier ages. We note, however, that this cut of the data was not part of our PAP and should thus be taken as suggestive.

[26] In contrast, Dizon-Ross (2014) finds that providing information on child ability to parents in Malawi increases the accuracy of their beliefs and increases investment in level-appropriate educational materials.

5 Discussion

5.1 Intensity of Mother Literacy Classes

Our take-up results suggest that one key aspect inhibiting the effectiveness of the ML classes was mother interest and attendance. Even though our recruitment strategy aimed to encourage all target mothers to attend the classes, in practice only 42 percent reported having ever attended. According to administrative attendance data, for mothers who did attend, attendance averaged 29 out of 105 classes throughout the entire year. Given the relatively low attendance, and the fact that mothers were starting from very low levels of literacy and numeracy, it is not surprising that the effects we find (for both mothers and children) are relatively modest.[27]

As shown in the analysis of take-up in Section 4.2, mothers' previous experience meeting in groups, agency within their households, and their interest in participating in child learning were important factors in the take-up decision. This suggests that even when recruitment efforts target all illiterate mothers, low interest and motivation to attend may lead some mothers to not participate. According to qualitative interviews, the primary excuses for absence from ML classes were lack of free time, or a perception that there was little value in attending. Indeed, it was often difficult to schedule classes around mothers' busy schedules, and mothers were often not willing to give up their leisure time to attend. In Rajasthan, where a higher proportion of households had televisions, evening classes sometimes had to be scheduled around mothers' favorite programs. Our experience with mother motivation to attend echoes prior work highlighting learner interest as a key impediment to the success of adult literacy interventions in developing countries (Wagner, 2000; Abadzi, 2003).

The pattern of impacts across the three interventions suggest scope for larger impacts with more intensive interventions. In our key learning and empowerment results, the point estimates for ML-CHAMP are approximately equal to the sum of the ML and CHAMP impacts, and we cannot reject that the effect of ML-CHAMP is equal to the sum of the effects of the individual interventions for any of these outcomes. While the two programs may have not been complementary, they also did not appear to be substitutes. Based on these results, we speculate that innovations to the ML model that better sustain learner interest could have larger impacts than what we have found here.

We also note that the interventions we study were implemented by a large and organized NGO, and government implementation of these programs may face other challenges. During our intervention period, the state government of Bihar implemented the Saakshar Bharat adult education program in our study area. The government adopted Pratham ML materials for its intervention, and while its teachers were paid, they were often community members with similar profile to

[27] Appendix E analyzes the cost-effectiveness of the interventions in improving child and mother test scores.

the Pratham volunteers. However, based on observations by our field staff, the implementation was a failure. Very few classes were set up, and of those, the vast majority ran for only a few days. Thus, while we are able to provide evidence on scalable ML model that in concept is similar to that implemented by the government, we caution against interpreting our results as directly applicable to the actual implementation of that government program.

5.2 Mechanisms for Child Learning Impacts

In this section we discuss the mechanisms through which the programs could have impacted child learning. Although the experiment was not designed to isolate a particular mechanism, the results do provide suggestive evidence of the potential pathways for impact.

Roughly one third of mothers targeted by the ML program participated in educational activities, which ultimately resulted in significant impacts on mothers' average learning levels. We hypothesized that their experience in the classes, coupled with an increase in skills, would expand mothers' interest in participating in their children's education and their ability to do so. This could be reflected in an increased ability to monitor children's education (e.g., homework), and/or increased time spent with children on educational activities. We note that child outcomes could be affected independently by the quality of time spent, without seeing any movement in the quantity of time. Indeed, we find significant effects of ML on the participation index, which primarily measured different ways in which mothers could monitor child learning, but we do not find evidence for impacts on time spent helping with homework. Beyond participation at home, an increased interest in child education could have led to the suggestive impacts we find on earlier enrollment of children into formal schooling.

CHAMP targeted mother participation in child learning directly and, as reflected by our participation index, significantly increased the number of activities mothers utilized to monitor their children's education. As with ML, the analysis of time use suggests small but insignificant changes in direct help with homework, implying that if monitoring was the primary mechanism, increases in child learning were again driven by the quality of monitoring rather than time spent doing so.[28] CHAMP also increased the presence of educational materials at home, suggesting that the program's emphasis on home learning led mothers (or other household members) to increase the availability of such materials to facilitate learning at home. Finally, CHAMP had small effects on mother learning, which, as with ML, could have led to increased capability to support child learning at home.

We also cannot rule out that part of the child learning impacts were driven by direct child participation in the ML and CHAMP interventions. Substantial numbers of children attended mother

[28]We also note that our time use question measured "help with studies," which may not have captured time spent on all of the activities in which mothers supported learning at home.

literacy classes at least once, although we do not have precise data on children's frequency of attendance. In the CHAMP treatment, we cannot rule out that child presence during the visits or that the worksheets themselves impacted children directly, apart from the participation of mothers.

6 Conclusion

In developing countries, adult literacy programs have been introduced to to improve literacy and empowerment for a generation of adults who lack experience with formal schooling, as well as to improve child outcomes through the home learning environment. Programs that encourage parental involvement have also been cited as a means of improving child learning through education inputs at home. In spite of the potential of these programs to improve key policy outcomes, there has been little rigorous evidence on whether they are actually effective in achieving their objectives.

Our paper evaluates three variants of these programs, each targeting mothers in rural India: an adult literacy program, a participation program, and a combination of the two. We show that all three interventions had significant impacts on mother learning. The mother literacy and combined mother literacy and participation interventions also increased measures of women's empowerment.

In addition to the effects on mothers, each program improved the math skills of young children, as well as several intermediate outcomes related to the home learning environment. While the evidence is consistent with an improved home learning environment leading to increased child learning, we cannot isolate the precise channels that led to the child learning impacts. Understanding the importance of each potential channel is a key area for future research. Nonetheless, our evaluation provides encouraging evidence for policy makers looking to improve to child learning and the home learning environment in addition to adult literacy and empowerment of women.

Although the experiment was not optimally designed to evaluate mother motivation in these programs, we argue that a key to the effectiveness of the mother literacy intervention was mother interest. Evaluation of strategies that might increase interest and participation in mother literacy and parental involvement programs is an important area for future research. We also note that this paper does not fully analyze the opportunity costs of these programs, or whether they are optimal from the perspective of every household. This paper can serve as a starting point for future work examining the welfare implications of these programs.

References

Helen Abadzi. Adult literacy: A review of implementation experience. *The World Bank Operations Evaluation Department*, 2003.

Jenny C. Aker, Christopher Ksoll, and Travis J. Lybbert. Can mobile phones improve learning? Evidence from a field experiment in Niger. *American Economic Journal: Applied Economics*, 4(4):94–120, 2012.

Aya Aoki. Assessing learning achievements and development impact: Ghana's national functional literacy program. *Australian Journal of Adult Learning*, 45, November 2005.

ASER Centre. *Annual Status of Education Report (Rural) 2012*. 2013.

Francesco Avvisati, Bruno Besbas, and Nina Guyon. Parental involvement in school: A literature review. *Economeie De L'Education*, 120(5):759–778, 2010.

Francesco Avvisati, Marc Gurgand, Nina Guyon, and Eric Maurin. Getting parents involved: A field experiment in deprived schools. *The Review of Economic Studies*, 81(1):57–83, 2014.

Abhijit V. Banerjee, Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani. Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India. *American Economic Journal: Economic Policy*, 2(1):1–30, 2010.

Hal Beder. The outcomes and impacts of adult literacy education in the United States. *National Center for the Study of Adult Learning and Literacy Report 6*, (6), 1999.

Sevda Bekman. A fair chance: An evaluation of the mother-child program. *Mother-Child Education Foundation Publication 13*, 1998.

G. Carron. The functioning and effects of the Kenya literacy program. *African Studies Review*, 33(3):97–120, 1990.

Central Statistics Office of India. Statement: Per capita net state domestic product at constant (2004-2005) prices. 2013.

James S. Coleman, Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. *Equality of Eductational Opportunity*. US Government Printing Office, Washington, DC, 1966.

Rebecca Dizon-Ross. Parents' perceptions and children's education: Experimental evidence from Malawi. *Mimeo, MIT*, 2014.

Government of India. Census of India 2011: Provisional population totals. 2011a.

Government of India. Census of India 2011: Source of lighting: 2001-2011. 2011b.

Jeffrey R. Kling, Jeffrey B. Liebman, and Lawrence F. Katz. Experimental analysis of neighborhood effects. *Econometrica*, 75 (1):83–119, 2007.

Michael Kremer, Conner Brannen, and Rachel Glennerster. The challenge of education and learning in the developing world. *Science*, 340(6130):297–300, 2013.

David S. Lee. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76:1071–1102, 2009.

Anju Malhotra, Sidney Ruth Schuler, and Carol Boender. Measuring womenÿs empowerment as a variable in international development. In Deepa Narayan, editor, *World Bank Workshop on Poverty and Gender: New Perspectives*, pages 71–89. The World Bank, 2002.

Daniel Ortega and Francisco Rodriguez. Freed from illiteracy? A closer look at Venezuela's Mision Robinson literacy campaign. *Economic Development and Cultural Change*, 57(1): 1–30, 2008.

Mark M. Pitt, Shahidur R. Khandker, and Jennifer Cartwright. Empowering women with micro finance: Evidence from Bangladesh. *Economic Development and Cultural Change*, 54(4): 791–831, 2006.

Petra E. Todd and Kenneth I. Wolpin. The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital*, 1(1):91–136, 2007.

UNESCO. *Aspects of Literacy Assessment: Topics and Issues from the UNESCO Expert Meeting*. 2005a.

UNESCO. *Education for All Global Monitoring Report 2006: Literacy for Life*. 2005b.

UNESCO. *Education for All Global Monitoring Report 2015: Education for All 2000-2015: Achievements and Challenges*. 2015.

United Nations. *The Millennium Development Goals Report 2014*. 2014.

United Nations Development Programme. *Human Development Report 2014: Sustaining Human Progress: Reducing Vulnerabilities and Building Resilience*. 2014.

Daniel A. Wagner. EFA 2000 thematic study on literacy and adult education. *International Literacy Institute*, 2000.

Table 1: Randomization Check

	Mean	Relative to Control				
	Control	ML	CHAMP	ML-CHAMP	P-value: All coeffs 0	N
	(1)	(2)	(3)	(4)	(5)	(6)
A. Household-level Variables						
First principal component of durables ownership	-0.0341 [2.250]	0.00735 (0.0859)	0.0939 (0.0944)	0.0306 (0.0899)	0.747	8888
Main source of income: farming	0.500 [0.500]	0.0128 (0.0223)	0.00718 (0.0243)	0.0494** (0.0219)	0.0952	8581
Number of children 0-4	0.929 [0.900]	-0.0112 (0.0298)	0.0179 (0.0332)	0.0675** (0.0297)	0.0370	8888
Number of children 5-8	1.438 [0.602]	-0.0300* (0.0174)	-0.0265 (0.0191)	-0.0139 (0.0171)	0.332	8888
Number of children 9-14	0.976 [0.936]	0.0560* (0.0289)	0.0282 (0.0299)	0.0549** (0.0271)	0.138	8888
Number of children 15-17	0.265 [0.500]	0.0130 (0.0152)	0.0125 (0.0156)	0.00232 (0.0154)	0.769	8888
Number of adults 18 and over	2.914 [1.500]	0.0263 (0.0569)	0.127** (0.0610)	0.117** (0.0562)	0.0709	8888
Father education (years)	3.876 [4.438]	-0.150 (0.203)	0.133 (0.226)	0.234 (0.213)	0.290	8181
Mother education (years)	0.764 [2.282]	0.0475 (0.102)	0.152 (0.118)	0.0694 (0.103)	0.625	8864
Mother age	32.29 [7.102]	-0.253 (0.254)	-0.193 (0.273)	-0.232 (0.254)	0.764	8888
Mother has past experience with literacy classes	0.117 [0.321]	-0.00839 (0.0123)	-0.00825 (0.0130)	-0.0209* (0.0124)	0.378	8635
Fraction of household (10 and over) that can read	0.433 [0.412]	0.000829 (0.0174)	0.0242 (0.0191)	0.0133 (0.0182)	0.544	7876
Fraction of household (10 and over) that can do math	0.285 [0.372]	0.00143 (0.0159)	0.0143 (0.0183)	0.0170 (0.0163)	0.652	7879
Mother language score (fraction)	0.299 [0.247]	0.00442 (0.0110)	0.0134 (0.0131)	0.00548 (0.0115)	0.785	8857
Mother math score (fraction)	0.215 [0.241]	0.00639 (0.0110)	0.0106 (0.0129)	0.00598 (0.0115)	0.860	8857
Mother total score (fraction)	0.250 [0.234]	0.00557 (0.0108)	0.0117 (0.0128)	0.00577 (0.0113)	0.830	8857

Table 1: Randomization Check (continued)

	Mean	Relative to Control				
	Control	ML	CHAMP	ML-CHAMP	P-value: All coeffs 0	N
	(1)	(2)	(3)	(4)	(5)	(6)
B. Child-Level Variables (tested children)						
Child is male	0.520	-0.0132	-0.00781	-0.0213**	0.231	15502
	[0.500]	(0.0108)	(0.0110)	(0.0106)		
Child attends school / aanganwadi	0.824	-0.0006	0.0220	0.0140	0.397	11793
	[0.381]	(0.0167)	(0.0164)	(0.0168)		
Child language score (fraction)	0.350	0.00753	0.00912	0.0103	0.786	15502
	[0.296]	(0.0110)	(0.0113)	(0.0109)		
Child math score (fraction)	0.277	0.0114	0.0131	0.00857	0.669	15502
	[0.303]	(0.0111)	(0.0117)	(0.0109)		
Child total score (fraction)	0.310	0.00971	0.0113	0.00934	0.726	15502
	[0.291]	(0.0108)	(0.0113)	(0.0106)		

Notes:
Standard deviations in square brackets, standard errors in parentheses.
Columns 2, 3, and 4 display the differences in means between each treatment group and the control group.
Column 5 displays the p-value of the F-test that the differences in means between the treatment groups and control group are jointly zero.
Differences in means are computed by OLS regression, controlling for stratum dummies.
Father education variable has fewer observations due to absence of fathers from households.
Average household reading and math variables have fewer observations due to missing test scores for older non-mother household members.
Child school attendance was measured for children aged 5-8, leading to missing values for some tested children.
"Aanganwadi" is community-based kindergarten.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Table 2: Takeup of Mother Literacy Classes

	Mean	Impact of Treatment in Endline			
	Control	ML	CHAMP	ML-CHAMP	*N*
	(1)	(2)	(3)	(4)	(5)
Mother knew about ML classes	0.220	0.405***	-0.00489	0.454***	8581
	[0.414]	(0.0227)	(0.0200)	(0.0216)	
Mother attended ML class	0.0710	0.323***	-0.00143	0.367***	8581
	[0.257]	(0.0184)	(0.0128)	(0.0179)	
Child attended ML class	0.0506	0.217***	-0.00996	0.275***	8522
	[0.219]	(0.0153)	(0.00969)	(0.0166)	

Notes:
Standard deviations in square brackets, standard errors in parentheses.
Columns 2, 3, and 4 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for stratum dummies.
"Child" refers to one randomly selected child per household aged 5-8 at baseline.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Table 3: Determinants of Mother Takeup

	Dependent Variable: Mother Attended		
	Sample		
	ML & ML-CHAMP	ML	ML-CHAMP
	(1)	(2)	(3)
First principal component of durables ownership	-0.00449	0.00131	-0.00869
	(0.00551)	(0.00730)	(0.00819)
Number of children 0-4	-0.00647	-0.0188	0.00396
	(0.0102)	(0.0153)	(0.0137)
Number of children 5-14	0.0106	0.0132	0.00936
	(0.00777)	(0.0114)	(0.0107)
Number of adults 15 and over	-0.000654	0.00505	-0.00793
	(0.00539)	(0.00773)	(0.00719)
Fraction of children 5-8 male	-0.00395	0.00730	-0.0171
	(0.0182)	(0.0256)	(0.0248)
Fraction of children 5-8 in school or aanganwadi	0.0488**	0.0440	0.0410
	(0.0239)	(0.0321)	(0.0352)
Average child total test score (normalized)	-0.00413	-0.0250*	0.0164
	(0.0111)	(0.0134)	(0.0175)
Father education (years)	-0.00377*	-0.00204	-0.00635**
	(0.00218)	(0.00331)	(0.00290)
Mother education > 0	0.147***	0.164***	0.125**
	(0.0432)	(0.0611)	(0.0613)
Mother education (years)	-0.0471***	-0.0380***	-0.0543***
	(0.00741)	(0.0107)	(0.0101)
Mother age	-0.00157	-0.00206	-0.00145
	(0.00137)	(0.00182)	(0.00203)
Mother total test score (normalized)	0.0493***	0.0342*	0.0637***
	(0.0151)	(0.0203)	(0.0221)
Total hours worked (in and out of home) per week	0.000289	0.000654	-0.0000808
	(0.000385)	(0.000510)	(0.000552)
Mother has past experience with literacy classes	0.0655**	0.0739*	0.0535
	(0.0293)	(0.0379)	(0.0453)
Mother is member of self-help group	0.0943***	0.0952***	0.0969***
	(0.0253)	(0.0349)	(0.0358)
Baseline empowerment index	0.0239***	0.0168	0.0324**
	(0.00877)	(0.0121)	(0.0126)
Baseline mother-child participation index	0.0233**	0.0304**	0.0145
	(0.00932)	(0.0120)	(0.0144)
State = Bihar	0.126***	0.172***	0.0844*
	(0.0301)	(0.0397)	(0.0449)
Mean of Dep. Var.	0.425	0.401	0.450
R-Squared	0.0594	0.0706	0.0594
N	3790	1907	1883

Notes:
Each column displays estimated coefficients of an OLS regression of self-reported mother attendance on the variables listed.
Empowerment and mother-child participation indices are defined as in Tables 5 and 7, respectively.
F-test that all differences between Columns 2 and 3 are 0: p-value = 0.28.
Standard errors, in parentheses, are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Table 4: Mother Learning

	Impact of Treatment in Endline						IV
	ML	CHAMP	ML-CHAMP	P-value: ML=CHAMP	P-value: Additive	*N*	Impact of ML class
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Language	0.0666***	0.0241**	0.0893***	0.001	0.940	8552	0.190***
	(0.0125)	(0.0120)	(0.0129)				(0.0349)
Math	0.119***	0.0616***	0.151***	0.000	0.205	8552	0.349***
	(0.0163)	(0.0149)	(0.0164)				(0.0467)
Total	0.0973***	0.0450***	0.126***	0.000	0.378	8552	0.282***
	(0.0126)	(0.0114)	(0.0129)				(0.0354)

Notes:

Standard errors in parentheses.

Columns 1, 2, and 3 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for baseline math and language test scores, stratum dummies, durables ownership, household composition, mother and father education, mother's past experience with adult literacy classes, and average child and other household member reading and math scores. Missing values of baseline variables are coded as 0, with additional dummies indicating missing values.

Column 4 displays the p-value of F-test that the impacts of ML and CHAMP are equal.

Column 5 displays the p-value of F-test that the impact of ML-CHAMP equals the sum of the ML and CHAMP impacts.

Column 7 displays the impact of literacy class attendance (either mother or child) on the dependent variables, using assignment to the ML treatment group as an instrument for attendance.

Test scores are normalized based on the control group means and standard deviations for each category of score and in each round of testing.

Standard errors are clustered at the village level.

* denotes significance at 0.10; ** at 0.05; *** at 0.01

Table 5: Empowerment

	Impact of Treatment in Endline					
	ML	CHAMP	ML-CHAMP	P-value: ML=CHAMP	P-value: Additive	*N*
	(1)	(2)	(3)	(4)	(5)	(6)
Empowerment Index (26 components)	0.0690* (0.0386)	0.0498 (0.0386)	0.133*** (0.0402)	0.608	0.801	8581
Mobility and Networks Index (5 components)	0.00124 (0.0409)	-0.0407 (0.0426)	-0.0174 (0.0439)	0.297	0.721	8581
Capability Index (7 components)	0.0699* (0.0366)	0.0377 (0.0373)	0.141*** (0.0383)	0.387	0.531	8581
Beliefs and Attitudes Index (6 components)	0.0641* (0.0375)	0.139*** (0.0425)	0.143*** (0.0416)	0.0654	0.316	8581
Involvement in Decisions Index (7 components)	0.0426 (0.0361)	0.0266 (0.0350)	0.0698* (0.0373)	0.651	0.991	8581
Happiness Index (1 component)	0.0333 (0.0352)	0.0523 (0.0363)	0.0379 (0.0334)	0.572	0.328	8581

Notes:
Standard errors in parentheses.
Columns 1, 2, and 3 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for stratum dummies and baseline values (where available).
Column 4 displays the p-value of F-test that the impacts of ML and CHAMP are equal.
Column 5 displays the p-value of F-test that the impact of ML-CHAMP equals the sum of the ML and CHAMP impacts.
Each index is an normalized average of z-scores of the component variables in the index, using the control group means and standard deviations. See Appendix Table 4 for components of each index.
Baseline indices only include indicators for which data were collected.
The "Happiness Index" is the answer to the question "Last week, how happy were you?" coded on a scale from 1 (very sad) to 5 (very happy), and then normalized relative to the control group mean and standard deviation.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Table 6: Child Learning

	Impact of Treatment in Endline						IV
	ML	CHAMP	ML-CHAMP	P-value: ML=CHAMP	P-value: Additive	*N*	Impact of ML class
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Language	-0.00487	0.0155	0.0454***	0.188	0.109	18283	-0.0114
	(0.0158)	(0.0146)	(0.0146)				(0.0445)
Math	0.0359**	0.0392**	0.0602***	0.832	0.506	18283	0.102**
	(0.0162)	(0.0157)	(0.0157)				(0.0451)
Total	0.0180	0.0295**	0.0554***	0.430	0.699	18283	0.0522
	(0.0150)	(0.0144)	(0.0141)				(0.0418)

Notes:
Standard errors in parentheses.
Columns 1, 2, and 3 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for baseline math and language test scores, stratum dummies, durables ownership, household composition, mother and father education, mother's past experience with adult literacy classes, and mother and average household member reading and math scores.
Missing values of baseline variables are coded as 0, with additional dummies indicating missing values.
Column 4 displays the p-value of F-test that the impacts of ML and CHAMP are equal.
Column 5 displays the p-value of F-test that the impact of ML-CHAMP equals the sum of the ML and CHAMP impacts.
Column 7 displays the impact of literacy class attendance (either mother or child) on the dependent variables, using assignment to the ML treatment group as an instrument for attendance.
Test scores are normalized based on the control group means and standard deviations for each category of score and in each round of testing.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Table 7: Mother Participation in Child Learning

	Baseline Mean	Endline Mean	Impact of Treatment in Endline			
	All Obs	Control	ML	CHAMP	ML-CHAMP	*N*
	(1)	(2)	(3)	(4)	(5)	(6)
Mother Participation Index	0.0470	0.000	0.0711**	0.129***	0.108***	8573
	[1.011]	[1.000]	(0.0334)	(0.0359)	(0.0360)	
Takes child to school (times/week)	0.243	0.322	-0.0447	0.0360	-0.00832	8451
	[1.082]	[1.127]	(0.0376)	(0.0418)	(0.0356)	
Visited school in past 3 months	0.129	0.157	0.00864	0.0148	0.0136	8451
	[0.336]	[0.364]	(0.0121)	(0.0130)	(0.0124)	
Reason for most recent visit was child learning	0.0670	0.0598	0.00610	0.0158*	0.00916	8480
	[0.250]	[0.237]	(0.00835)	(0.00837)	(0.00787)	
Knows how often child receives homework	0.826	0.783	0.00998	0.0121	0.00852	8253
	[0.379]	[0.412]	(0.0119)	(0.0118)	(0.0125)	
Helps child with homework	0.436	0.789	0.00618	0.0335***	0.0243*	8548
	[0.496]	[0.408]	(0.0130)	(0.0127)	(0.0130)	
Opened child's notebook in past 2 months	0.125	0.218	0.0310**	0.0682***	0.0489***	8572
	[0.331]	[0.413]	(0.0151)	(0.0157)	(0.0151)	
Talks to children about school (times/week)		2.994	0.211**	0.244**	0.277***	8525
		[3.049]	(0.102)	(0.106)	(0.105)	
Talks to friends/family about child's education (times/week)	0.931	1.504	0.264***	0.192**	0.263***	8520
	[1.589]	[2.248]	(0.0778)	(0.0774)	(0.0808)	

Notes:

Standard deviations in square brackets, standard errors in parentheses.

Columns 3, 4, and 5 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for stratum dummies and baseline values (where available).

The "Mother Participation Index" is a normalized average of z-scores of the other variables in the table, using the control group means and standard deviations. The baseline participation index only includes indicators for which data were collected.

"Child" refers to one randomly selected child per household aged 5-8 at baseline.

Standard errors are clustered at the village level.

* denotes significance at 0.10; ** at 0.05; *** at 0.01

Table 8: Mother Weekly Time Use

	Baseline Mean	Endline Mean	Impact of Treatment in Endline			
	All Obs	Control	ML	CHAMP	ML-CHAMP	*N*
	(1)	(2)	(3)	(4)	(5)	(6)
Help with homework	1.700	2.309	0.110	0.141	0.0809	8510
	[2.930]	[2.706]	(0.0924)	(0.0976)	(0.0931)	
Read with children	0.200	0.321	-0.00885	-0.0268	0.00538	8398
	[1.059]	[1.374]	(0.0402)	(0.0368)	(0.0428)	
Play with children	0.255	1.339	0.0793	0.0544	-0.0233	8472
	[1.351]	[3.295]	(0.116)	(0.121)	(0.113)	
Share stories with children	0.378	0.499	0.0304	-0.0150	0.0309	8514
	[1.229]	[1.424]	(0.0472)	(0.0439)	(0.0513)	
Paid / agricultural work	30.83	31.13	0.997	0.444	0.938	8547
	[22.53]	[21.07]	(0.613)	(0.604)	(0.611)	
Livestock work	9.524	9.662	0.151	-0.340	0.520**	8573
	[7.448]	[7.042]	(0.259)	(0.241)	(0.257)	
Collect animal feed		6.963	0.206	-0.280	0.193	8577
		[6.773]	(0.268)	(0.274)	(0.287)	
Collect wood		3.278	0.0975	-0.0981	0.0598	8570
		[5.102]	(0.182)	(0.200)	(0.195)	
Housework	22.27	19.02	0.319	0.196	0.341	8581
	[9.025]	[8.008]	(0.307)	(0.306)	(0.288)	
Buy household supplies	4.888	1.201	-0.0835	-0.0757	-0.0166	8567
	[6.424]	[2.844]	(0.0945)	(0.0960)	(0.0926)	
Look after children	5.743	4.646	-0.198*	-0.147	0.112	8581
	[4.038]	[3.679]	(0.106)	(0.109)	(0.105)	

Notes:
Standard deviations in square brackets, standard errors in parentheses.
Time spent in each activity measured in hours per week.
Columns 3, 4, and 5 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for stratum dummies and baseline values (where available).
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Table 9: Child Weekly Time Use

	Baseline Mean	Endline Mean	Impact of Treatment in Endline			
	All Obs	Control	ML	CHAMP	ML-CHAMP	*N*
	(1)	(2)	(3)	(4)	(5)	(6)
Homework	3.178	4.067	-0.0304	0.185	0.255*	7821
	[4.193]	[4.341]	(0.157)	(0.147)	(0.147)	
Read	0.258	0.430	-0.0344	-0.00746	0.0634	7942
	[1.264]	[1.512]	(0.0489)	(0.0528)	(0.0540)	
Draw / paint	0.465	0.685	0.00707	0.0679	0.0845	7896
	[1.391]	[1.569]	(0.0534)	(0.0530)	(0.0533)	
Play with adult	0.456	0.548	0.0169	-0.0266	-0.105*	8335
	[2.108]	[1.993]	(0.0661)	(0.0662)	(0.0628)	
Tuition classes	1.864	2.287	0.209	0.155	0.0208	8414
	[4.491]	[4.926]	(0.182)	(0.195)	(0.174)	
Television	3.950	3.711	-0.116	0.141	0.0312	8338
	[5.961]	[5.080]	(0.174)	(0.180)	(0.171)	
Housework	3.213	3.605	0.0504	0.107	0.158	8408
	[4.343]	[4.112]	(0.134)	(0.134)	(0.136)	
Household business	1.209	1.883	0.200	0.00339	0.294**	8407
	[3.799]	[4.157]	(0.140)	(0.136)	(0.142)	

Notes:
Standard deviations in square brackets, standard errors in parentheses.
Child time use measured for one randomly selected child per household aged 5-8 at baseline.
Time spent in each activity measured in hours per week.
Columns 3, 4, and 5 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for stratum dummies and baseline values.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Table 10: Home Education Assets

	Baseline Mean	Endline Mean	Impact of Treatment in Endline			
	All Obs	Control	ML	CHAMP	ML-CHAMP	*N*
	(1)	(2)	(3)	(4)	(5)	(6)
Education Asset Index	0.0319	0.000	0.0557	0.101***	0.115***	8581
	[1.001]	[1.000]	(0.0371)	(0.0361)	(0.0371)	
Has pen / pencil	0.930	0.945	0.0000	0.0153**	0.0118	8581
	[0.254]	[0.227]	(0.00850)	(0.00728)	(0.00795)	
Has school books		0.906	0.0180*	0.0168*	0.0260***	8581
		[0.292]	(0.00977)	(0.00932)	(0.00992)	
Has other books / comics	0.231	0.251	0.0145	0.0320*	0.0349**	8581
	[0.421]	[0.434]	(0.0156)	(0.0170)	(0.0158)	
Has newspaper / magazine	0.123	0.0561	0.00840	0.0275***	0.00632	8581
	[0.328]	[0.230]	(0.00806)	(0.00900)	(0.00806)	
Has slate		0.889	0.0130	-0.00348	0.0269***	8581
		[0.314]	(0.0104)	(0.0104)	(0.00980)	

Notes:
Standard deviations in square brackets, standard errors in parentheses.
Columns 3, 4, and 5 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for stratum dummies and baseline values (where available).
The "Education Asset Index" is a normalized average of z-scores of the other variables in the table, using the control group means and standard deviations. The baseline index only includes indicators for which data were collected.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Table 11: Mother Aspirations and Perceptions

	Baseline Mean	Endline Mean	Impact of Treatment in Endline			
	All Obs	Control	ML	CHAMP	ML-CHAMP	*N*
	(1)	(2)	(3)	(4)	(5)	(6)
A. Perceptions of parental involvement						
Number of educational activities for which parents are responsible	0.734 [0.773]	1.375 [1.001]	0.0620 (0.0439)	0.0508 (0.0441)	0.0812** (0.0392)	8581
Number of things mother can do to help child improve studies	1.247 [0.949]	1.911 [1.116]	0.0671 (0.0447)	0.0358 (0.0486)	0.100** (0.0439)	8513
Believes mother should be responsible for children's education	0.730 [0.444]	0.726 [0.446]	0.0440*** (0.0132)	0.0325** (0.0135)	0.0418*** (0.0129)	8358
Identifies home help as reason for children's achievement	0.233 [0.423]	0.210 [0.407]	0.00616 (0.0146)	0.00703 (0.0138)	-0.00463 (0.0143)	8581
B. Expectations/aspirations for child attainment						
Believes child will pass 8th grade	0.833 [0.373]	0.855 [0.352]	0.00299 (0.0134)	0.00154 (0.0127)	0.00351 (0.0123)	7926
Believes child will pass 12th grade	0.622 [0.485]	0.652 [0.476]	0.0118 (0.0181)	0.0159 (0.0181)	0.00929 (0.0172)	7481
Highest grade to which mother aspires for child to study	9.751 [2.768]	10.01 [2.944]	-0.0138 (0.125)	0.0641 (0.130)	0.235* (0.140)	3214
Highest grade to which mother aspires for child to study: child's wish	0.491 [0.500]	0.432 [0.495]	0.0113 (0.0128)	-0.00491 (0.0148)	0.00361 (0.0143)	8484
C. Mother perceptions of child reading/math ability						
Perception of child's language ability (out of 4)	2.073 [1.691]	2.476 [1.583]	-0.0360 (0.0518)	0.153*** (0.0504)	0.122** (0.0517)	7597
Perception of child's math ability (out of 4)	1.985 [1.589]	2.578 [1.608]	0.00394 (0.0549)	0.227*** (0.0551)	0.159*** (0.0579)	7713
Absolute value of difference between language perception and child score	1.485 [1.387]	1.611 [1.366]	-0.0261 (0.0431)	0.107** (0.0451)	0.0383 (0.0439)	7232
Absolute value of difference between math perception and child score	1.243 [1.274]	1.477 [1.263]	-0.0468 (0.0443)	0.149*** (0.0485)	0.0953** (0.0456)	7347

Notes:
Standard deviations in square brackets, standard errors in parentheses.
Columns 3, 4, and 5 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for stratum dummies and baseline values.
"Child" refers to one randomly selected child per household aged 5-8 at baseline.
The regression using "highest grade to which mother aspires for child to study" as an outcome has fewer observations due to mothers indicating that grade attainment would be the child's decision.
Regressions using beliefs of 8th and 12th grade passing and perceptions of child ability have fewer observations due to responses of "don't know" to the survey questions.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Table 12: Child Schooling

	Baseline Mean	Endline Mean	Impact of Treatment in Endline			
	All Obs	Control	ML	CHAMP	ML-CHAMP	*N*
	(1)	(2)	(3)	(4)	(5)	(6)
A. Children aged 3-14 at baseline						
Child is enrolled in school		0.774	0.0133	0.0141	0.0152	25053
		[0.418]	(0.0118)	(0.0117)	(0.0117)	
Child is enrolled or will be enrolled next school year		0.903	-0.00111	0.00683	0.0108	25053
		[0.296]	(0.00751)	(0.00720)	(0.00723)	
Child attends school / aanganwadi	0.835	0.844	-0.00389	0.00436	0.0199*	25053
	[0.371]	[0.363]	(0.0109)	(0.0108)	(0.0106)	
Child attends private school	0.0923	0.0813	0.00556	-0.000765	0.00330	25033
	[0.289]	[0.273]	(0.00846)	(0.00860)	(0.00882)	
B. Target children (aged 5-8 at baseline)						
Hours per day spent in school	3.667	4.072	0.0595	0.0990*	0.0788	8451
	[1.977]	[1.617]	(0.0560)	(0.0546)	(0.0566)	
Days missed last week	2.097	1.200	0.0272	-0.0740	-0.0311	7841
	[2.457]	[1.865]	(0.0635)	(0.0664)	(0.0601)	
Days missed due to work in past month		1.920	-0.0877	-0.00620	-0.126	8357
		[5.120]	(0.180)	(0.189)	(0.172)	
Monthly tuition fees	14.10	20.05	2.726*	3.073	0.440	8403
	[35.91]	[44.04]	(1.608)	(1.878)	(1.551)	
C. Children aged 4-6 at baseline						
Child attends school / aanganwadi	0.764	0.887	-0.00177	0.00190	0.0219*	8110
	[0.425]	[0.317]	(0.0136)	(0.0129)	(0.0129)	
Child attends school	0.620	0.691	0.0390**	0.00673	0.0372**	8110
	[0.485]	[0.462]	(0.0159)	(0.0171)	(0.0173)	

Notes:
Standard deviations in square brackets, standard errors in parentheses.
Columns 3, 4, and 5 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for stratum dummies and baseline values (where available).
In Panels A and C, baseline data were collected for children aged 5-8. Missing values of baseline variables are coded as 0, with additional dummies indicating missing values.
"Target child" refers to one randomly selected child per household aged 5-8 at baseline.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

A Impact Heterogeneity

We use the following estimating equation to examine heterogeneity in treatment effects:

$$\begin{aligned} Y_{1iv} &= \beta_0 + \beta_1 Var_{iv} + \beta_2 ML_v + \beta_3 CHAMP_v + \beta_4 MLCHAMP_v + \beta_5 Var_{iv} * ML_v \\ &\quad + \beta_6 Var_{iv} * CHAMP_v + \beta_7 Var_{iv} * MLCHAMP_v + \beta_8 Y_{0iv} + \delta G_v + \pi H_{iv} + \varepsilon_{iv} \end{aligned}$$

In this equation, Var_{iv} is the interacted variable for individual i in village v, and the remainder of the variables are defined as in Equation (1).

A.1 Mothers

Appendix Table 7 examines heterogeneity in treatment effects of the interventions on mother test scores. We focus on heterogeneity by the state where the intervention took place, the mother's baseline score, mother's age, and mother's education level.

As shown in Column 1 of the table, there is evidence that the ML and ML-CHAMP interventions were more effective in Bihar. The point estimates imply that these interventions resulted in more than double the impact on test scores in Bihar compared to Rajasthan, and the differences between the two states are significant at the 1 percent level for both interventions. The greater effectiveness of the ML and ML-CHAMP interventions in Bihar is consistent with the fact that mothers in the Bihar sample were 13 percentage points more likely to attend a class (see Section 4).

Although we do not observe significant heterogeneity along many other dimensions, we do find that the CHAMP intervention was significantly more effective for mothers with higher initial test scores. As shown in Column 2, the point estimate implies that each standard deviation increase in baseline test scores resulted in an increase in the impact of CHAMP by 0.034 standard deviations. This result is consistent with the finding in Section 4 that CHAMP impacts were spread across higher level skills, particularly in language.We also find that the CHAMP intervention was slightly more effective for younger mothers.

A.2 Children

Appendix Table 8 examines heterogeneity in treatment effects on child test scores. We examine heterogeneity by state, mother baseline test score, mother age, mother education, child age, child baseline test score, and child gender. Overall, there is little evidence of heterogeneity by any of these variables. Consistent with the mother results, point estimates suggest

approximately double the impact of ML and ML-CHAMP in Bihar, but neither interaction is statistically significant. The ML-CHAMP intervention had significantly larger impacts on older children, but there is no similar interaction with either of the individual interventions.

B Location Selection

Because of the slightly different organization of the villages in Rajasthan and Bihar, a different selection procedure was used in each state. The procedure focused on finding distinct hamlets in which the programs could run while limiting spillovers. Hamlet eligibility was therefore determined based on size and distance from other target hamlets. Size and location of hamlets were determined from "Rapid Rural Assessments" conducted in study blocks.

In Ajmer District in Rajasthan, villages are geographically separate, and each village is divided into several smaller hamlets. Hamlets met the size eligibility requirements if they contained between 40 and 100 households. To limit spillovers, one hamlet per village was selected. All villages in the blocks of Kekri and Bhinay were targeted for the intervention. To identify a total of 240 hamlets, the boundaries of Kekri and Bhinay were extended into a third block.

In Purnia District in Bihar, the village boundaries are less distinct, and villages are denser than in Rajasthan. As in Rajasthan, each village is comprised of smaller hamlets. All villages the blocks of Dhamdaha and B. Kothi were targeted for the intervention. Within these villages, hamlets were considered eligible if they contained between 25 and 150 households.[29] To limit spillovers, hamlets in Bihar were selected only if their boundaries were 500 meters or more from the boundaries of other selected hamlets.[30,31]

C Attrition

We analyze attrition across three samples: the mother survey sample (3.5 percent attrition), the mother testing sample (3.4 percent attrition), and the child testing sample (6.0 percent attrition). Mother attrition was primarily driven by a failure to locate the entire household (typically because the household had moved). The additional child attrition arose primarily because of logistical difficulties in testing several children per household during each visit and in finding all of the children present at home. Appendix Table 9 analyzes the levels of attrition

[29]The household criteria differed between Rajasthan and Bihar because the criteria for Rajasthan would not have produced a sufficient number of eligible hamlets were it applied in Bihar. Due to the higher upper bound on number of households, Pratham agreed to hold more than one class in a target hamlet where necessary in Bihar.

[30]GPS coordinates were used to confirm distances between the boundaries of hamlets. Distances were checked between hamlets within villages as well as across villages.

[31]In two cases, target hamlets were eliminated because Pratham determined that adult literacy rates were too high to sustain classes.

across the three treatment groups. In the mother survey and mother testing samples, there is no evidence of differential attrition. However, there is some evidence of differential attrition in the child testing sample: children tested at baseline in the CHAMP and ML-CHAMP groups were about 2 percentage points more likely to exit the sample compared with the control group.

To evaluate the possible effects of attrition on validity of the child learning estimates, Appendix Table 10 analyzes the balance of characteristics across treatment groups in the sample that was tested at endline. Unlike the randomization check in Table 1, this table uses child-level data to mirror the child learning analysis sample in Table 4. Overall, the endline balance is similar to baseline: of the 21 variables analyzed, one is jointly significant across the three treatment groups at the 10 percent level, one more is significant at the 5 percent level, and two more are significant at the 1 percent level. Critically, there is no evidence of imbalance in any of the test score variables. In addition, as shown in Appendix Table 2, inclusion of household-level controls leaves the estimated program effects virtually unchanged.

Appendix Table 11 constructs bounds around the child learning estimates using Lee's (2009) trimming method. The estimation uses only the sample of children tested at baseline and thus follows the specification in Appendix Table 2, Panel B, Columns 3, 6, and 9. We trim the top or bottom test scores in the control, ML, and CHAMP groups such that the generated attrition rates in each group are equal to those of the ML-CHAMP group, in which we observe the highest level of attrition. We then regress endline test scores on dummies for treatment group, baseline test scores, and dummies for stratum. As shown in Appendix Table 11, the small absolute differences in attrition between treatment groups lead to relatively tight bounds. For example, the estimated effect size of the CHAMP treatment on math scores ranges from 0.033 to 0.051 standard deviations. For the ML-CHAMP intervention, the effect ranges from 0.052 to 0.071 standard deviations. There is thus little evidence that differential attrition substantially influences our child learning results.

D Surveyor Demand Effects

Although our main test score outcomes were based on tests conducted by our enumerators, our other outcomes are based on mother self-reports. Because we do not have independent verification of these responses, we cannot completely rule out surveyor demand effects. However, the pattern of effects we find is inconsistent with surveyor demand effects that would generate positive responses to the broad set of questions relating to education, or to only the questions that specifically relate to components of the ML or CHAMP programs.

First, the patterns of impacts do not suggest that the mothers in the intervention groups responded in uniformly favorable ways to the broad set of questions relating to education. Although the interventions did impact our summary indices of participation, empowerment, and

education assets in the home, the impacts on the underlying components are not universally and strongly positive, as would be the case if respondents exhibited surveyor demand effects for all topics relating to education. For example, the ML intervention, while influencing the overall participation index, did not have significant impacts on 5 of the 8 individual components of the index. In addition, as shown in Table 11, we see virtually no detectable impacts of any of the interventions on mother aspirations for child attainment, and limited impacts on perceptions of parental involvement in child education.

A second possibility is that mothers were responding in ways that would be favorable to the specific components related to the programs in which they participated. Although we do observe larger impacts on some of these measures, this would be expected in the absence of surveyor demand effects. It is thus more challenging to separate surveyor demand effects from true program impacts. However, in this case we can also provide evidence inconsistent with systematic systematic surveyor demand effects among these questions.

In the ML program, the three specific outcomes that most closely relate to the program come from the empowerment index: the questions "Do you consider yourself literate?" "Do you count your change?", and "Have you signed your name on official documents?" relate to skills taught in the classes, and we observe positive impacts on all three these outcomes. While it is not possible to fully verify the responses to these questions, we can check the report of signing one's name against the mother's ability to write her name in the endline test. Among mothers who were not able to write their name at endline, about 14 percent indicated they had signed their name on official documents. However, there is no significant difference in this "misreporting" between the villages where writing the mother's name was taught (ML and ML-CHAMP) and villages in which no such instruction was given (Control and CHAMP groups).

We also observe impacts from ML on a large number of indicators, particularly in empowerment and participation, that were not targeted by the program. For example, the ML program did have a specific participation component, yet we do observe impacts on some of the participation measures, as well as the aggregate participation index.

For the CHAMP program, we again see evidence of impacts on components targeted by the programs, particularly the intensive margins of helping children with homework and looking at the child's notebook. However, there is no detectable impact on visiting the child's school, which was also covered in the program. Beyond the participation measures, we also observe impacts on a number of components that were not explicitly part of the program, particularly in the empowerment and education assets categories.

We also note that a number of outcomes, particularly several participation and mother and child time use questions relating to education, increased in the control group between baseline and endline. These differences were likely to have arisen from variation in survey timing and

changes in question wording, rather than surveyor demand effects. The baseline survey was conducted entirely during school vacations before the 2011-2012 school year, while the endline survey partially overlapped with the end of the 2011-2012 school year, and thus one would expect higher participation in school-related activities at endline. A number of younger children were also enrolled in school between baseline and endline. Second, in order to more accurately measure home behaviors that complemented school, in the endline surveyors prompted mothers to respond based on when school was in session. This extra prompt could have also increased reported incidence of education-related behaviors in the endline relative to the baseline, but it would not be expected to cause differential response between treatment groups.

E Cost-Effectiveness

This section presents a discussion of cost-effectiveness of the three interventions in improving mother child test scores. As with most cost-effectiveness comparisons across studies, we note that differences in target population, competencies tested, testing instruments, local prices, and methods of calculating costs may limit comparability. Our cost-effectiveness calculation follows the methodology in Kremer, et. al (2013).

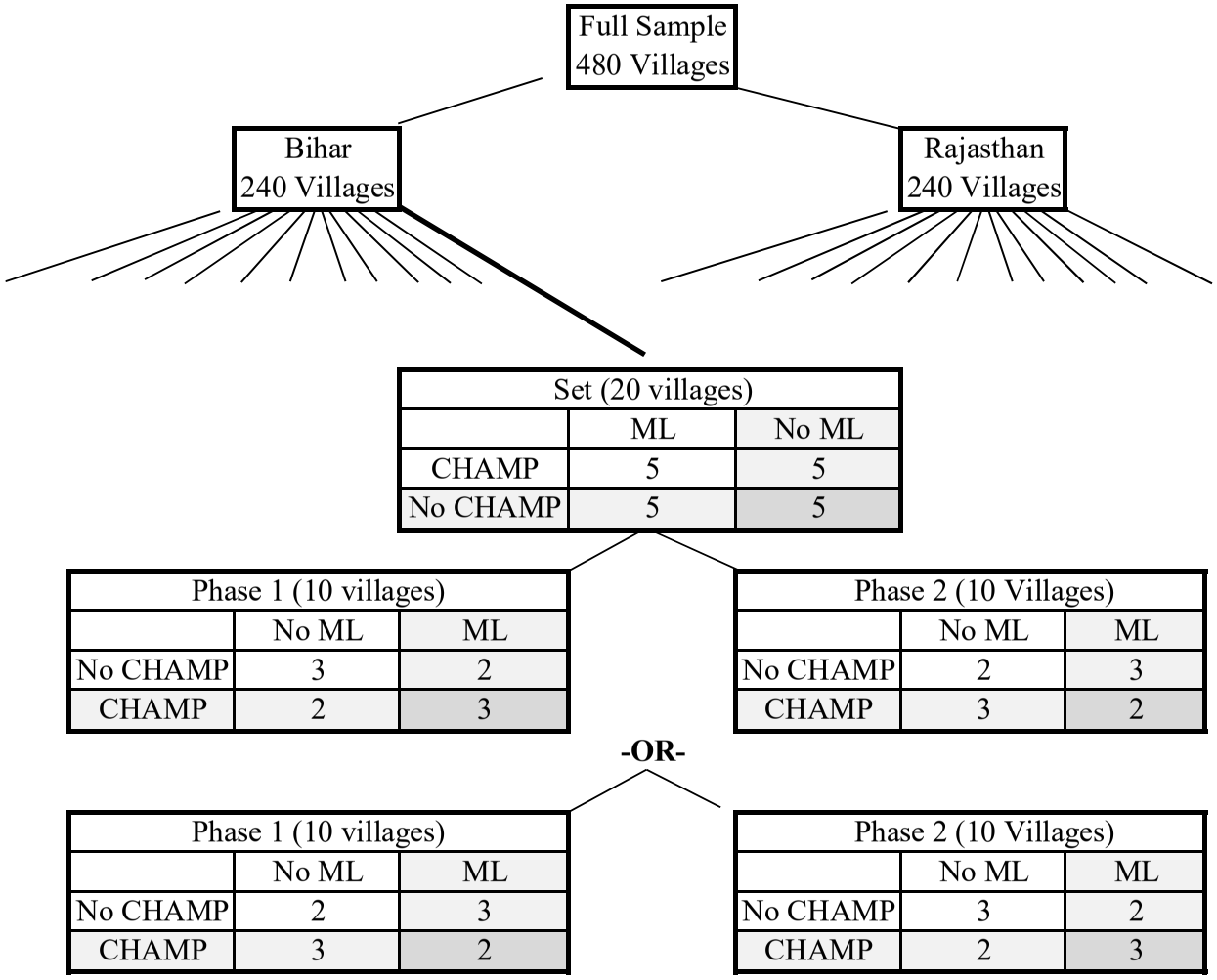
Appendix Table 12 summarizes the costs of the major components of each intervention and the standard deviation improvement per $100 spent for mothers and for children. For mothers, each $100 spent on the ML, CHAMP, and ML-CHAMP interventions resulted in improvements in total test scores of 0.29, 0.27, and 0.25 standard deviations, respective. For children, each $100 spent on ML, CHAMP, and ML-CHAMP resulted in improvements in math scores of 0.22, 0.49, and 0.26 standard deviations, respectively.

Because there are no other cost-effectiveness calculations of program effects on adult learning that we know of, we focus our comparison with the literature on child learning impacts. Our estimates, ranging 0.04 to 0.07 standard deviations, fall below the range of statistically-significant estimates in the Kremer, et al. (2013) study. This latter set of estimates ranges from 0.14 standard deviations to 0.6 standard deviations. Turning to cost-effectiveness, 14 out of the 15 studies in Kremer, et al. (2013) that found statistically significant impacts are also more cost-effective than the interventions studied here.

Even though the interventions studied here may be less cost effective than others in improving child test scores, our study examines and finds impacts on a broad set of outcomes, including mother learning, women's empowerment, and the home learning environment. A full cost-effectiveness analysis would take into account the full set of impacts relative to costs. However, given the limited number of studies that examine these outcomes and the lack of a consistent framework to compare cost-effectiveness along all of these dimensions, we are unable to perform this broader comparison. Absent an explicit comparison with alternative interventions,

we tentatively conclude that the interventions studied here are cost-effective considering the broad set of outcomes affected. Our study should serve as a starting point for future work to expand the evidence base so that more explicit cost-effectiveness comparisons can be made in the future.

Appendix Figure 1: Stratification of Treatment Assignment



Appendix Table 1: Baseline Means of Household-level Variables Split by Rajasthan and Bihar Samples

	Rajasthan	Bihar
	(1)	(2)
First principal component of durables ownership	1.239	-1.190
	[2.280]	[1.384]
Household has electricity	0.808	0.154
	[0.394]	[0.361]
Number of children 0-4	0.819	1.075
	[0.928]	[0.945]
Number of children 5-14	2.383	2.486
	[1.157]	[1.118]
Number of adults 15 and over	3.549	2.973
	[2.007]	[1.386]
Fraction of children 5-8 male	0.523	0.513
	[0.454]	[0.444]
Fraction of children 5-8 in school / aanganwadi	0.846	0.827
	[0.339]	[0.351]
Average child total test score (fraction)	0.114	-0.0730
	[0.886]	[0.803]
Father education	4.819	3.124
	[4.436]	[4.296]
Mother education	0.741	0.922
	[2.173]	[2.480]
Mother age	31.17	33.05
	[5.885]	[7.755]
Mother total test score (fraction)	0.284	0.228
	[0.241]	[0.240]
Total hours worked (in and out of home) per week	73.71	51.47
	[23.85]	[24.17]
Mother has past experience with literacy classes	0.143	0.0743
	[0.350]	[0.262]
Mother is member of self-help group	0.0553	0.490
	[0.229]	[0.500]
N	4334	4523

Notes:
Standard deviations in square brackets.

Appendix Table 2: Mother and Child Learning, Alternative Specifications

	Impact of Treatment in Endline								
	ML			CHAMP			ML-CHAMP		
	Main	Limited Controls	Missing Baseline	Main	Limited Controls	Missing Baseline	Main	Limited Controls	Missing Baseline
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Mothers									
Language	0.0666***	0.0675***	—	0.0241**	0.0299**	—	0.0893***	0.0928***	—
	(0.0125)	(0.0136)		(0.0120)	(0.0140)		(0.0129)	(0.0141)	
Math	0.119***	0.118***	—	0.0616***	0.0638***	—	0.151***	0.151***	—
	(0.0163)	(0.0164)		(0.0149)	(0.0151)		(0.0164)	(0.0167)	
Total	0.0973***	0.0969***	—	0.0450***	0.0491***	—	0.126***	0.127***	—
	(0.0126)	(0.0132)		(0.0114)	(0.0126)		(0.0129)	(0.0136)	
Panel B: Children									
Language	-0.00487	-0.00725	-0.0142	0.0155	0.0205	0.00976	0.0454***	0.0460***	0.0472***
	(0.0158)	(0.0170)	(0.0184)	(0.0146)	(0.0174)	(0.0178)	(0.0146)	(0.0163)	(0.0173)
Math	0.0359**	0.0346**	0.0364*	0.0392**	0.0430**	0.0417**	0.0602***	0.0598***	0.0624***
	(0.0162)	(0.0176)	(0.0191)	(0.0157)	(0.0177)	(0.0185)	(0.0157)	(0.0172)	(0.0182)
Total	0.0180	0.0162	0.0139	0.0295**	0.0340**	0.0282	0.0554***	0.0555***	0.0575***
	(0.0150)	(0.0165)	(0.0174)	(0.0144)	(0.0170)	(0.0171)	(0.0141)	(0.0159)	(0.0164)

Notes:

Standard errors in parentheses.

Columns 1, 4, and 7 display estimated coefficients from Tables 4 and 6 for Panels A and B, respectively.

Columns 2, 5, and 8 use the same specifications as in Tables 4 and 6, without household-level controls.

Panel B, Columns 3, 6, and 9 use the same specification as in Table 6, without household level controls, and including only the sample of children tested at baseline.

* denotes significance at 0.10; ** at 0.05; *** at 0.01

Appendix Table 3: Treatment Effects on Mother ASER Learning Levels

	Baseline	Endline Mean	Impact of Treatment in Endline			
	All Obs	Control	ML	CHAMP	ML-CHAMP	*N*
	(1)	(2)	(3)	(4)	(5)	(6)
Language Competency						
Read letters	0.195	0.173	0.0347***	0.00952*	0.0469***	8552
	[0.355]	[0.338]	(0.00592)	(0.00566)	(0.00604)	
Read words	0.130	0.0926	0.00562	0.00973**	0.0104**	8552
	[0.326]	[0.278]	(0.00438)	(0.00462)	(0.00432)	
Read paragraph	0.0974	0.0778	0.00256	0.00731	0.00382	8552
	[0.297]	[0.268]	(0.00421)	(0.00508)	(0.00445)	
Read story		0.0622	0.00309	0.0147	0.00619	8580
		[0.242]	(0.00967)	(0.0111)	(0.00986)	
Math Competency						
One-digit number recognitior	0.447	0.469	0.0680***	0.0319***	0.108***	8552
	[0.459]	[0.461]	(0.0103)	(0.0101)	(0.0109)	
Two-digit number recognitior	0.102	0.0962	0.0142***	0.00744*	0.0151***	8552
	[0.285]	[0.276]	(0.00442)	(0.00444)	(0.00439)	
Two-digit addition	0.1000	0.0783	0.00196	0.00592	0.0132**	8552
	[0.300]	[0.269]	(0.00549)	(0.00589)	(0.00565)	
Two-digit subtraction	0.0428	0.0356	0.00704	0.00893*	-0.00228	8552
	[0.202]	[0.185]	(0.00464)	(0.00530)	(0.00423)	

Notes:
Standard deviations in square brackets, standard errors in parentheses.
Columns 3, 4, and 5 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for stratum dummies and baseline values (where available
Standard errors are clustered at the village level
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Appendix Table 4: Individual Empowerment Measures

	Baseline Mean	Endline Mean	Impact of Treatment in Endline			
	All Obs	Control	ML	CHAMP	ML-CHAMP	*N*
	(1)	(2)	(3)	(4)	(5)	(6)
Mobility And Networks Index	-0.00686	0.000	0.00124	-0.0407	-0.0174	8581
	[0.970]	[1.000]	(0.0409)	(0.0426)	(0.0439)	
Times left village in the past month	1.409	1.150	0.0420	0.0805	0.0988	8581
	[2.242]	[1.648]	(0.0608)	(0.0623)	(0.0664)	
Left without adult accompaniment	0.147	0.113	0.00341	0.0284***	0.00674	8581
	[0.354]	[0.316]	(0.0101)	(0.0105)	(0.0104)	
Left village without permission	0.0255	0.0168	-0.00381	0.000553	-0.00314	8581
	[0.158]	[0.129]	(0.00368)	(0.00391)	(0.00372)	
Member of self-help group		0.329	-0.00121	-0.0604***	-0.0249	8581
		[0.470]	(0.0207)	(0.0224)	(0.0217)	
Fraction of self-help group meetings attended		0.289	-0.000696	-0.0581***	-0.0262	8578
		[0.441]	(0.0203)	(0.0213)	(0.0215)	
Capability Index	0.000	0.000	0.0699*	0.0377	0.141***	8581
	[0.984]	[1.000]	(0.0366)	(0.0373)	(0.0383)	
Considers self literate		0.233	0.0488***	0.0175	0.0743***	8581
		[0.423]	(0.0163)	(0.0174)	(0.0170)	
Signed name of official documents	0.542	0.561	0.0644***	0.0137	0.0838***	8581
	[0.498]	[0.496]	(0.0139)	(0.0134)	(0.0138)	
Counts change	0.877	0.867	0.0254**	0.0242**	0.0434***	8581
	[0.329]	[0.339]	(0.0116)	(0.0115)	(0.0112)	
Caught mistakes counting change	0.313	0.314	0.0149	-0.00549	0.0254	8581
	[0.464]	[0.464]	(0.0172)	(0.0167)	(0.0172)	
Knows NREGA wage		0.621	-0.0611**	-0.00914	-0.0169	8356
		[0.485]	(0.0249)	(0.0252)	(0.0248)	
Knows self-help group account balance		0.127	-0.00329	-0.0175	-0.0117	8581
		[0.333]	(0.0132)	(0.0133)	(0.0132)	
Fills self-help group passbook herself		0.0135	0.00492	0.00638	0.00325	8332
		[0.116]	(0.00376)	(0.00408)	(0.00385)	
Beliefs and Attitudes Index	0.0171	0.000	0.0641*	0.139***	0.143***	8581
	[1.035]	[1.000]	(0.0375)	(0.0425)	(0.0416)	
Does not believe husband should be more educated than wife	0.344	0.335	0.0412**	0.0128	0.0527***	7936
	[0.475]	[0.472]	(0.0166)	(0.0173)	(0.0175)	
Does not believe girls should be at home or married when 18	0.0528	0.0449	-0.00266	0.00639	0.00274	8581
	[0.224]	[0.207]	(0.00644)	(0.00683)	(0.00682)	
Believes girls should be doing further studies when 18	0.0767	0.280	0.0648***	0.0931***	0.108***	8581
	[0.266]	[0.449]	(0.0176)	(0.0177)	(0.0179)	
Would have wanted to study up to (grade level)		5.799	-0.405**	0.240	-0.269*	8290
		[4.362]	(0.162)	(0.178)	(0.159)	

Appendix Table 4: Individual Empowerment Measures (continued)

	Baseline Mean	Endline Mean	Impact of Treatment in Endline			
	All Obs	Control	ML	CHAMP	ML-CHAMP	*N*
	(1)	(2)	(3)	(4)	(5)	(6)
Decisions Index		0.000	0.0426	0.0266	0.0698*	8581
		[1.000]	(0.0361)	(0.0350)	(0.0373)	
Involved in purchasing utensils		0.544	0.0315*	0.0278	0.0340*	8546
		[0.498]	(0.0177)	(0.0186)	(0.0187)	
Involved in purchasing cot		0.458	0.00362	-0.0143	0.00598	8500
		[0.498]	(0.0177)	(0.0176)	(0.0182)	
Involved in purchasing cycle		0.245	0.00673	-0.0271**	-0.00298	8494
		[0.430]	(0.0141)	(0.0128)	(0.0133)	
Involved in purchasing educational materials		0.482	0.0240	0.0224	0.0434**	8549
		[0.500]	(0.0179)	(0.0178)	(0.0182)	
Value of goods can buy alone		2437.8	117.6	65.20	180.3**	8581
		[2258.8]	(84.06)	(91.72)	(90.98)	
Involved in deciding girl school enrollment		0.493	0.0220	0.0161	0.0305*	7821
		[0.500]	(0.0175)	(0.0177)	(0.0183)	
Involved in deciding girl school type		0.499	0.0190	0.0182	0.0429**	7757
		[0.500]	(0.0172)	(0.0176)	(0.0172)	
Involved in deciding boy school enrollment		0.399	0.00389	0.0139	0.0188	8238
		[0.490]	(0.0164)	(0.0162)	(0.0166)	
Involved in deciding boy school type		0.422	0.00539	0.0219	0.0243	8211
		[0.494]	(0.0165)	(0.0161)	(0.0174)	
Happiness		3.107	0.0480	0.0755	0.0547	8581
		[1.443]	(0.0508)	(0.0524)	(0.0482)	

Notes:
Standard deviations in square brackets, standard errors in parentheses.
Columns 3, 4, and 5 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for stratum dummies and baseline values (where available).
Each index is an normalized average of z-scores of the component variables in the index, using the control group means and standard deviations. Baseline indices only include indicators for which data were collected.
"Knows NREGA wage" indicates whether the mother answered an amount between Rs. 100 and Rs. 150 to the question "What is the per day wage fixed by the government under NREGA?" Official wage rates in Bihar and Rajasthan at the time of the survey were Rs. 122 and Rs. 124, respectively.
"Happiness" is the answer to the question "Last week, how happy were you?" coded on a scale from 1 (very sad) to 5 (very happy).
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Appendix Table 5: Empowerment: Alternative Indices

	Impact of Treatment in Endline			
	ML	CHAMP	ML-CHAMP	*N*
	(1)	(2)	(3)	(4)
Main Empowerment Index (26 components)	0.0690*	0.0498	0.133***	8581
	(0.0386)	(0.0386)	(0.0402)	
Exclude Capabilities (22 components)	0.0398	0.0430	0.0854**	8581
	(0.0384)	(0.0390)	(0.0407)	
Involvement in Purchase Decisions (5 Components)	0.0504	0.0101	0.0676*	8581
	(0.0361)	(0.0355)	(0.0383)	

Notes:
Standard errors in parentheses.
Columns 1, 2, and 3 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for stratum dummies and baseline values (where available).
Each index is an normalized average of z-scores of the component variables in the index, using the control group means and standard deviations.
Baseline indices only include indicators for which data were collected.
"Exclude Capabilities" represents an index that excludes the 4 measures of considering oneself literate, having signed one's name on official documents, counting change, and catching mistakes counting change.
"Involvement in Purchase Decisions" represents an index that includes involvement in the decision to buy educational materials, utencils, a cot, a cycle, as well as the value of goods the mother can buy alone.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Appendix Table 6: Treatment Effects on Child ASER Learning Levels

	Baseline Mean	Endline Mean	Impact of Treatment in Endline			
	All Obs	Control	ML	CHAMP	ML-CHAMP	*N*
	(1)	(2)	(3)	(4)	(5)	(6)
Language Competency						
Read letters	0.345	0.389	0.000890	0.0143	0.0266***	18283
	[0.420]	[0.430]	(0.00890)	(0.00911)	(0.00888)	
Read words	0.215	0.176	0.000243	0.00857	0.0113*	18283
	[0.394]	[0.353]	(0.00646)	(0.00635)	(0.00619)	
Read paragraph	0.126	0.118	-0.00225	0.00133	0.0127**	18283
	[0.332]	[0.322]	(0.00623)	(0.00622)	(0.00610)	
Read story		0.0751	0.00680	0.00829	0.0110	18283
		[0.264]	(0.00876)	(0.00842)	(0.00826)	
Math Competency						
One-digit number recognition	0.528	0.562	0.0193**	0.0368***	0.0390***	18283
	[0.474]	[0.474]	(0.00857)	(0.00881)	(0.00865)	
Two-digit number recognition	0.162	0.201	0.00166	0.00276	0.00667	18283
	[0.339]	[0.368]	(0.00740)	(0.00725)	(0.00704)	
Two-digit addition	0.207	0.198	0.0231**	0.00730	0.00815	18283
	[0.405]	[0.399]	(0.00967)	(0.00930)	(0.00897)	
Two-digit subtraction	0.0621	0.0579	0.0129**	-0.00276	0.00540	18283
	[0.241]	[0.234]	(0.00589)	(0.00567)	(0.00530)	

Notes:
Standard deviations in square brackets, standard errors in parentheses.
Columns 3, 4, and 5 display estimated coefficients of a regression of the outcome in each row on treatment group dummies, controlling for stratum dummies and baseline values (where available).
Missing values of baseline variables are coded as 0, with additional dummies indicating missing values.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Appendix Table 7. Heterogeneity in Impact

Outcome: Mother Total Test Score

	Impact of Treatment in Endline						
	ML	CHAMP	ML-CHAMP	Variable * ML	Variable * CHAMP	Variable * ML-CHAMP	N
Interacted Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
State = Bihar	0.0568***	0.0621***	0.0819***	0.0789***	-0.0342	0.0876***	8552
	(0.0148)	(0.0145)	(0.0163)	(0.0246)	(0.0214)	(0.0251)	
Baseline mother test score	0.0971***	0.0436***	0.126***	-0.00539	0.0313**	-0.00877	8552
	(0.0125)	(0.0111)	(0.0129)	(0.0143)	(0.0129)	(0.0135)	
Mother age	0.0747*	0.146***	0.153***	0.000658	-0.00315**	-0.000876	8552
	(0.0452)	(0.0451)	(0.0471)	(0.00131)	(0.00129)	(0.00137)	
Mother has some education	0.101***	0.0357***	0.135***	-0.0317	0.0571	-0.0604	8528
	(0.0128)	(0.0115)	(0.0131)	(0.0424)	(0.0379)	(0.0409)	

Notes:

Standard errors in parentheses.

Each column displays the results of a regression of the mother's normalized total test score on treatment dummies, the interaction variable indicated, and interactions of the variable and treatment dummies.

Regressions include all control variables used in Table 4.

Mother test score is the mother's total normalized test score.

Standard errors are clustered at the village level.

* denotes significance at 0.10; ** at 0.05; *** at 0.01

Appendix Table 8. Heterogeneity in Impact
Outcome: Child Total Test Score

	Impact of Treatment in Endline						
	ML	CHAMP	ML-CHAMP	Variable * ML	Variable * CHAMP	Variable * ML-CHAMP	N
Interacted Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
State = Bihar	0.0108	0.0249	0.0378*	0.0132	0.00834	0.0324	18283
	(0.0221)	(0.0216)	(0.0200)	(0.0300)	(0.0289)	(0.0277)	
Baseline mother test score	0.0182	0.0288**	0.0562***	0.0103	0.0151	0.0216	18221
	(0.0151)	(0.0144)	(0.0141)	(0.0138)	(0.0143)	(0.0143)	
Mother age	-0.0640	-0.0223	0.0684	0.00255	0.00161	-0.000409	18283
	(0.0545)	(0.0531)	(0.0593)	(0.00160)	(0.00156)	(0.00172)	
Mother has some education	0.0134	0.0239	0.0502***	0.0376	0.0412	0.0359	18235
	(0.0154)	(0.0146)	(0.0144)	(0.0396)	(0.0396)	(0.0404)	
Child age	0.0157	0.0148	-0.00396	0.000501	0.00227	0.00927**	18282
	(0.0266)	(0.0280)	(0.0263)	(0.00459)	(0.00470)	(0.00438)	
Baseline child test score	0.0143	0.0242	0.0564***	0.0103	-0.00228	-0.00549	14576
	(0.0163)	(0.0152)	(0.0151)	(0.0137)	(0.0137)	(0.0129)	
Gender = boy	0.0164	0.0440**	0.0508***	0.00405	-0.0284	0.0100	18283
	(0.0184)	(0.0180)	(0.0176)	(0.0209)	(0.0218)	(0.0217)	

Notes:
Standard errors in parentheses.
Each column displays the results of a regression of the child's normalized composite total score on treatment dummies, the interaction variable indicated, and interactions of the variable and treatment dummies.
Regressions include all control variables used in Table 6.
Mother and child test scores are the mother's and child's total normalized test scores.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Appendix Table 10: Attrition across treatment groups

	Mean	Relative to Control				
	Control	ML	CHAMP	ML-CHAMP	All coeffs 0	N
	(1)	(2)	(3)	(4)	(5)	(6)
Mother attritted: survey	0.0308	0.00700	0.00327	0.00593	0.548	8888
	[0.173]	(0.00530)	(0.00525)	(0.00523)		
Mother attritted: test	0.0313	0.00647	0.00152	0.00647	0.487	8857
	[0.174]	(0.00536)	(0.00517)	(0.00527)		
Child attritted: test	0.0488	0.00879	0.0190***	0.0213***	0.0113	15502
	[0.215]	(0.00650)	(0.00683)	(0.00752)		

Notes:
Standard deviations in square brackets, standard errors in parentheses.
Columns 2, 3, and 4 display the differences in means between each treatment group and the control group.
Column 5 displays the p-value of the F-test that the differences in means between the treatment groups and control group are all zero.
Differences in means are computed by OLS regression, controlling for stratum dummies.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Appendix Table 10: Balance After Attrition: Child-level Data

	Mean	Relative to Control				
	Control	ML	CHAMP	ML-CHAMP	All coeffs 0	N
	(1)	(2)	(3)	(4)	(5)	(6)
First principal component of durables ownership	-0.0726 [2.282]	0.00492 (0.0882)	0.138 (0.104)	0.0417 (0.0958)	0.516	18283
Main source of income: farming	0.481 [0.500]	0.0210 (0.0235)	0.000886 (0.0257)	0.0563** (0.0233)	0.0514	18175
Number of children 0-4	1.074 [0.926]	-0.00342 (0.0355)	0.0672* (0.0407)	0.117*** (0.0354)	0.00179	18283
Number of children 5-8	1.591 [0.668]	-0.0194 (0.0248)	0.0142 (0.0266)	-0.00587 (0.0236)	0.655	18283
Number of children 9-14	1.033 [0.955]	0.0739** (0.0312)	0.0537 (0.0349)	0.0777** (0.0318)	0.0402	18283
Number of children 15-17	0.273 [0.502]	0.00460 (0.0173)	0.0111 (0.0184)	-0.00762 (0.0171)	0.773	18283
Number of adults 18 and over	2.956 [1.608]	0.0281 (0.0659)	0.197** (0.0775)	0.186*** (0.0682)	0.00726	18283
Father education (years)	3.759 [4.376]	-0.182 (0.208)	0.0704 (0.231)	0.199 (0.213)	0.330	17024
Mother education (years)	0.708 [2.154]	-0.00500 (0.0957)	0.136 (0.115)	0.0452 (0.102)	0.566	18235
Mother age	32.35 [6.950]	-0.254 (0.270)	-0.178 (0.299)	-0.143 (0.271)	0.825	18283
Mother has past experience with literacy classes	0.114 [0.309]	-0.00994 (0.0134)	-0.00919 (0.0141)	-0.0228* (0.0133)	0.364	17831
Fraction of household (10 and over) that can read	0.419 [0.412]	-0.00459 (0.0183)	0.0273 (0.0203)	0.0103 (0.0191)	0.424	16236
Fraction of household (10 and over) that can do math	0.278 [0.372]	-0.00442 (0.0165)	0.0153 (0.0190)	0.0126 (0.0169)	0.620	16238
Mother language score (fraction)	0.294 [0.241]	-0.00227 (0.0108)	0.0112 (0.0132)	0.00446 (0.0119)	0.722	18221
Mother math score (fraction)	0.208 [0.234]	0.000837 (0.0109)	0.00881 (0.0132)	0.00464 (0.0120)	0.900	18221
Mother total score (fraction)	0.244 [0.227]	-0.000460 (0.0106)	0.00981 (0.0130)	0.00457 (0.0118)	0.833	18221

Appendix Table 10: Balance After Attrition: Child-level Data (continued)

	Mean	Relative to Control				
	Control	ML	CHAMP	ML-CHAMP	All coeffs 0	N
	(1)	(2)	(3)	(4)	(5)	(6)
Child is male	0.511	-0.00735	-0.00305	-0.0126	0.628	18283
	[0.500]	(0.00979)	(0.0103)	(0.0101)		
Child attends school / aanganwadi	0.822	0.00245	0.0319**	0.0172	0.140	11571
	[0.382]	(0.0167)	(0.0161)	(0.0168)		
Child language score (fraction)	0.348	0.0109	0.0130	0.0134	0.599	14576
	[0.296]	(0.0112)	(0.0116)	(0.0112)		
Child math score (fraction)	0.274	0.0149	0.0178	0.0114	0.447	14576
	[0.302]	(0.0113)	(0.0120)	(0.0112)		
Child total score (fraction)	0.307	0.0131	0.0157	0.0123	0.509	14576
	[0.290]	(0.0110)	(0.0116)	(0.0109)		

Notes:
Standard deviations in square brackets, standard errors in parentheses.
Columns 2, 3, and 4 display the differences in means between each treatment group and the control group.
Column 5 displays the p-value of the F-test that the differences in means between the treatment groups and control group are all zero.
Differences in means are computed by OLS regression, controlling for stratum dummies.
Father education variable has fewer observations due to absence of fathers from households.
Average household reading and math variables have fewer observations due to missing test scores for older non-mother household members.
Child school attendance was measured for children aged 5-8, leading to missing values for some tested children.
Standard errors are clustered at the village level.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Appendix Table 11: Child Learning, using Lee (2009) bounds

	Impact of Treatment in Endline								
	ML			CHAMP			ML-CHAMP		
	Estimate	Lower	Upper	Estimate	Lower	Upper	Estimate	Lower	Upper
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Language	-0.0142	-0.0196	-0.00777	0.00976	-0.000793	0.0245	0.0472**	0.0355*	0.0630***
	(0.0190)	(0.0190)	(0.0190)	(0.0183)	(0.0183)	(0.0183)	(0.0190)	(0.0190)	(0.0190)
Math	0.0364*	0.0329	0.0400**	0.0417**	0.0332*	0.0506***	0.0624***	0.0521***	0.0714***
	(0.0202)	(0.0202)	(0.0202)	(0.0186)	(0.0186)	(0.0186)	(0.0198)	(0.0198)	(0.0198)
Total	0.0139	0.00865	0.0178	0.0282	0.0177	0.0374**	0.0575***	0.0461**	0.0670***
	(0.0188)	(0.0188)	(0.0188)	(0.0189)	(0.0189)	(0.0189)	(0.0183)	(0.0183)	(0.0183)

Notes:
Standard errors in parentheses. Standard errors computed from 500 bootstrap draws, sampling by village.
Sample for regressions is restricted children who took the baseline test.
Columns 1, 4, and 7 display estimates from Appendix Table 2, Panel B, Columns 3, 6, and 9.
Columns 2-3, 5-6, and 8-9 display the bounds of the estimates using Lee's (2009) trimming method.
Number of observations equals 14,576 for main estimates and 14,438 for upper and lower bound estimates.
* denotes significance at 0.10; ** at 0.05; *** at 0.01

Appendix Table 12. Cost Effectiveness of Interventions

	Intervention		
	ML	CHAMP	ML-CHAMP
Panel A: Cost Summary ($)			
Pratham Staff	37,521	30,699	68,219
Volunteer Time	21,653	—	21,653
Training, monitoring, materials	14,861	4,730	19,590
Total	74,035	35,428	109,463
Panel B: Standard Deviation Improvement per $100 spent--Mothers			
Language	0.20	0.14	0.17
Math	0.35	0.37	0.30
Total	0.29	0.27	0.25
Mothers affected	2,173	2,105	2,140
Panel C: Standard Deviation Improvement per $100 spent--Children			
Language	—	—	0.19
Math	0.22	0.49	0.26
Total	—	0.37	0.24
Children affected	4,572	4,448	4,653

Notes:
Costs incurred in rupees converted to dollars using 2011 exchange rate of 46.7 rupees/dollar.
Volunteer time estimated based on average daily wage in non-agricultural occupations.
Mothers and children affected calculated using the same inclusion criteria as Tables 4 and 6, respectively.
Cost-effectiveness estimates in Panels B and C are computed based on the effect sizes reported in Tables 4 and 6, multiplied by the number of mothers or children affected, and divided by the total costs of each program in 100s of dollars.
Cost-effectiveness is only calculated for effects significant at the 10 percent level.