

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

By RUKMINI BANERJI, JAMES BERRY, AND MARC SHOTLAND\*

February 2017

*Using a randomized field experiment in India, we evaluate the effectiveness of adult literacy and parental involvement interventions in improving children's learning. 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 programs. All three interventions had significant but modest impacts on children's math scores. The interventions also increased mothers' test scores in both language and math, as well as a range of other outcomes reflecting greater involvement of mothers in their children's education.*

*JEL: C93, D13, I21, O15*

*Keywords: adult education, education inputs, field experiments*

Improving the quality of education in the developing world remains a crucial issue for researchers and policymakers alike. In many developing countries, despite gains in educational attainment, learning levels are abysmally low, both when compared with developed countries and with national learning standards (Pritchett, 2013). For example, a 2011 survey in India found that less than half of fifth-grade children could read at a second-grade level (ASER Centre, 2012).

In searching for methods to improve levels of learning in developing countries, much of the policy discussion focuses on investments that can be made to improve

\* 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 David Evans, Eric Edmonds, Marcel Fafchamps, Karthik Muralidharan, Hugo Reis, and seminar participants at the Berlin Social Science Center, Cornell University, Duke University, the NBER Education and Children's Meeting, the Northeast Universities Development Consortium Conference, UC Santa Cruz, the University of Buffalo, and the World Bank for useful comments and suggestions. We are incredibly grateful to Vipin Awatramani, Jessica Chan, Diva Dhar, 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. The project was reviewed and approved by the Massachusetts Institute of Technology Committee on the Use of Humans as Experimental Subjects, protocol #1102004358. We acknowledge that Rukmini Banerji is the CEO of Pratham, which represents a potential conflict of interest. She was involved in the design of the program and the evaluation, the implementation of the program, and the final paper review, but not with the data collection and the analysis of the data. All errors are our own.

schools, such as improving school infrastructure, providing additional learning materials, changing pedagogy, or improving teacher selection or incentives. Indeed, there is a growing evidence base of such cost-effective in-school interventions that can improve learning outcomes in developing countries (Kremer et al., 2013).

Beyond investments that governments can make in schools, evidence suggests that parents, and in particular mothers, also play an important role in children’s learning—through activities at home as well as through decisions that interact with the formal schooling system. The literature on intergenerational transmission of human capital shows that more educated (or, in some cases, more literate) mothers make decisions that improve their children’s learning (Andrabi, Das, and Kwajha, 2012; Behrman, Foster, and Rosensweig, 1999; Guryan, Hurst, and Kearney, 2008). The direct policy implications from this research imply that investing in the education of today’s children, particularly girls, should have positive spillovers on their children in the future. This literature does not, however, reveal obvious policy solutions for today’s children who live with today’s uneducated parents. There are nearly one billion adults worldwide who are currently illiterate (UNESCO, 2015).

To address low levels of adult literacy, a number of developing countries have launched ambitious adult literacy education campaigns.<sup>1</sup> For example, in 2009 the government of India launched the Saakshar Bharat adult literacy campaign, with a goal of bringing literacy skills to 70 million adults and reducing the gender gap in literacy. In South Africa, the Kha Re Gude Mass Literacy Programme was launched in 2008, aiming to reach 4.7 million adults by 2012. These campaigns often focus on women in order to address gender inequality in literacy and to promote women’s empowerment (UNESCO, 2015).

Although the primary motivation of initiatives like these has been to improve the livelihoods of adult beneficiaries, also listed among the instrumental benefits are improvements in the education of children (Abadzi, 2003; UNESCO, 2005b). Specifically, many of the intermediate outcomes through which parents’ educational attainment leads to children’s learning could also be activated by these types of programs. For example, such programs could provide parents with the skills to facilitate children’s learning at home and foster an interest in children’s education. This could lead to increased participation in learning at home, more effective participation, increased availability of educational assets at home, or greater interest in formal schooling. By targeting women, the programs could also influence child education by increasing decision making power of mothers in the household, under the assumption that child education is a higher priority for mothers than for fathers.

However, while adult literacy education may provide parents with the skills and interest to participate more in their children’s education, for those who were never formally educated, these programs may not be sufficient to provide them

<sup>1</sup>With respect to adult literacy programs, “literacy” typically encompasses reading, writing, and basic mathematics. See UNESCO (2005a) for a discussion of the definition of literacy.

with knowledge of how specifically they can participate. If a key constraint to facilitating children’s learning at home is parents’ lack of understanding of how they can participate, parental “participation training” that provides direct instruction on how they can be involved may prove more effective. 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 their children’s learning.

Although adult literacy and participation programs could serve as important tools to improve children’s 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 their children’s learning. Whether these programs are effective in improving children’s education outcomes is thus an empirical question.

To address this question, we conducted a randomized evaluation of three interventions in rural India designed to improve maternal literacy—defined to include reading, writing, and basic mathematics—and mothers’ involvement in children’s education. Villages were randomly assigned to one of four groups. In the first group, mothers in the village were offered the Maternal Literacy (ML) program, consisting of daily language and math classes. In the second, mothers were given the Child Home Activities and Materials Packet (CHAMP) program: 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 programs (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 find that the programs had small positive impacts on children’s learning levels: the ML, CHAMP, and ML-CHAMP interventions resulted in statistically-significant increases in children’s math scores by 0.035, 0.032, and 0.056 standard deviations, respectively. The only significant impacts on language scores came from the combined ML-CHAMP intervention.

We also provide evidence that the programs improved mothers’ learning outcomes and indicators of participation and investments associated with improved children’s learning. The three interventions increased mothers’ overall test scores by 0.096, 0.043, and 0.12 standard deviations, respectively. All three programs had impacts on mothers’ participation in their children’s education and on availability of education assets in the household, while the ML and ML-CHAMP programs had impacts on formal school attendance. By contrast, we find no

significant impacts on mothers' time spent directly helping the children with homework. The programs also had limited impacts on mothers' involvement in household decision making, as the only statistically significant impact arose from the combined ML-CHAMP intervention.

The evidence supports our hypothesis that the home learning environment can be improved through these programs and provides suggestive evidence that doing so resulted in increased learning of children. However, a key caveat to this interpretation is the possibility that the programs affected children directly. Indeed, we observed both attendance of children in ML classes and direct participation of children in CHAMP activities.

While we cannot rule out the possibility of these direct impacts, we present suggestive evidence that in the case of ML, the impacts we observe on children's test scores were largely driven by mothers' participation in ML classes. We also show that changes in mothers' learning levels, participation in their children's education, presence of educational assets in the home, and children's school attendance account for almost all of the impacts of ML on children's learning. In the case of CHAMP, these intermediate outcomes account for a smaller fraction of the impacts, and therefore direct participation of children could have played a greater role.

From a standpoint of policy, we note that the magnitudes of the effects we find on children's learning are small, and basic cost-benefit calculations suggest that other interventions that target children directly are more cost effective. Thus, the interventions as implemented may not be the most effective policies to improve children's learning. In the case of ML, 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. In addition, while the magnitudes of the impacts may not justify the programs on the basis of children's learning alone, the effects we find on mothers' learning suggest that these programs could be promising tools for policy makers interested in influencing adult literacy as well as children's learning.

Our study contributes to several strands of literature. First, it is one of only two randomized evaluations of developing-country adult literacy programs that we know of and the first to study spillovers on children. Early studies find significant effects of adult literacy programs on learning of adults, but these impacts are identified 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 participants' writing and math scores. However, there is no evaluation of the adult literacy program, per se. Most recently, Aker and Ksoll (2015) conduct an evaluation of an adult literacy program in Niger that included both a standard adult literacy program and adult literacy plus additional monitoring. The researchers find impacts of approximately 0.2 standard deviations on reading and

math scores of adults, with significantly larger effects when monitoring was included. Research on the effects of adult literacy programs on children’s outcomes is even more sparse, and existing studies rely on retrospective selection of comparison groups (Aoki, 2005; Abadzi, 2003).<sup>2</sup>

Second, our study contributes to the limited literature evaluating parental participation programs in developing countries, where parents have little or no experience with formal education. Kagitcibasi, Sunar, and Bekman (2001) conduct a randomized evaluation of a small-scale program in Turkey that included both a home learning component and training in general parenting skills for mothers of preschool-aged children. The authors find positive impacts on primary school enrollment and academic performance in subsequent years. Several recent studies evaluate programs to encourage parental participation in management of schools, with mixed results on effectiveness (Banerjee et al., 2010; Duflo, Dupas, and Kremer, 2015; Gertler, Patrinos, and Rubio-Codina, 2008).<sup>3</sup> 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 test scores (Avvisati et al., 2014).

The remainder of our paper is structured as follows. Section I discusses the context and programs. Section II covers the evaluation design and data collection. Section III analyzes program take-up. Section IV describes the results. We provide discussion of the results in Section V and conclude in Section VI.

## I. Program Description

### A. Context

The programs in this study were designed and implemented by Pratham, a non-governmental organization (NGO) specializing in child literacy and numeracy.<sup>4</sup> 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 2011 Census of India, these states have the lowest female literacy rates in India (Government of India, 2011). The intervention districts within each state were selected because of existing Pratham programs and

<sup>2</sup>Although there are numerous evaluations of adult literacy programs in the U.S., much of the research also suffers from methodological limitations (Beder, 1999).

<sup>3</sup>Our study also relates to the small but growing literature evaluating early-childhood stimulation interventions in developing countries which typically focus on development of children aged 0 to 3 (see Baker-Henningham and Lopez Boo, 2010 for a review). The interventions we study here differ in that they focus on households with school-aged children.

<sup>4</sup>For more information, see <http://www.pratham.org>.

infrastructure in those areas. Within the intervention districts, the study blocks were selected because they did not have any pre-existing Pratham programs.

In online Appendix Table A1 we use data from the 2011 Census of India and 2011 Annual Status of Education Report (ASER) surveys to compare the intervention states, districts, and blocks with our evaluation data (Government of India, 2011; ASER Centre, 2012). According to the Census, households in Rajasthan are more likely to have electricity than those in Bihar (58 percent compared with 10 percent) and have higher ownership rates of costly durable goods such as televisions, mobile phones, and motorcycles. Literacy rates, children’s school attendance, and children’s learning levels are broadly similar between the two states. Within each state, the intervention districts follow similar patterns to the state-wide averages, with slightly lower rates of female literacy and children’s schooling outcomes in both the Rajasthan and Bihar intervention districts compared with the state-wide averages. Characteristics in the intervention blocks are largely similar to those of the broader districts.<sup>5</sup>

As in the census data, households in our sample displayed generally higher asset ownership in Rajasthan than in Bihar and substantially higher likelihood of having electricity (81 percent compared with 15 percent). Households in our sample have higher levels of asset ownership than those in the blocks as a whole, which could be due to our selection criteria leading to larger households. We also observe lower levels of school attendance in our sample than the district-level ASER data indicate. This difference could be due to a difference in survey timing: the ASER surveys took place in the fall of 2011, while the surveys for this evaluation took place during the summer vacation period just after the beginning of the school year.

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.<sup>6</sup>

### *B. Intervention Descriptions*

Each of the three interventions was implemented 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. The mothers of these children (referred to as “target

<sup>5</sup>These comparisons suggest that Pratham may have been targeting districts with worse education outcomes within states for its programs in general, but the intervention blocks are largely representative of these districts.

<sup>6</sup>In online Appendix A we explore heterogeneity in the program effects on children’s and mothers’ learning by state.

mothers”) were identified from a census of communities selected to receive the programs (see Section II).

The Maternal 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 villages 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 a 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 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, government programs, and to a lesser extent, the value of educating girls.<sup>7</sup>

The direct objective 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.

Our primary hypothesis was that the ML intervention would improve children’s learning outcomes. This could arise if the skills learned, exposure to the learning process, and opportunity for discussion about education increased mothers’ interest and participation in their children’s education. The intervention could thereby change the type and quality 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 34 weekly discussion topics related to the benefits of educating girls, there were no specific program com-

<sup>7</sup>Of 34 topics, two related to girls’ education.

ponents providing instruction on participation in children’s learning. Although mothers were the sole focus of class activities, children were not prohibited from attending class and listening passively or following along with their mothers, and could have thereby directly benefited from attending classes themselves. As we show in Section III below, some study children attended both with their mothers and by themselves. We discuss the potential for direct effects on children in Section V.A.

The Child and Mother Activities Packet (CHAMP) intervention was designed to directly encourage mothers to be engaged with their children’s learning, focusing on activities with 5 to 8-year-old children in the household. Unlike the volunteers who implemented the ML intervention, CHAMP implementers were paid Pratham staff who often had prior experience as Pratham volunteers or supervisors.<sup>8</sup> Pratham decided to use paid implementers primarily because each implementer was responsible for conducting CHAMP in specific households in five separate villages, and Pratham believed that unpaid volunteers would not be willing to take on these responsibilities. These implementers were given about one week of training by senior Pratham staff and met together several times per month to review upcoming activities.<sup>9</sup>

Once per week during the intervention period, the CHAMP implementer was to visit each target mother and give her a worksheet to help her child complete and instruction on how to be involved in her child’s learning.<sup>10</sup> Each worksheet contained several written language or math activities that the mother was instructed to help the child complete in the following week. The activities focused on first-grade language and math skills. For example, some worksheets instructed the child to write the letters of the alphabet or a sequence of numbers. In about half of the sessions, the mother received direct instruction on 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 to 20 minutes. When the mother was not available, the CHAMP session was conducted with another adult member of the household when possible.

The primary objective of CHAMP was to improve children’s learning by involving the mother in the process. Mothers were shown how to participate in their children’s learning and were encouraged to spend more time on educational activities with their children. Beyond increasing participation of the mother, the

<sup>8</sup>Hiring staff with prior Pratham experience was not strictly necessary for the program, but Pratham’s practices for recruiting paid staff for new projects often consist of promoting successful volunteers or reallocating existing staff to from previous projects.

<sup>9</sup>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.

<sup>10</sup>When the mother had more than one child aged 5 to 8, she was given multiple worksheets and instructed to do the activities with each child in this age group.



program’s encouragement of home learning could also result in an increase in education assets available at home. Finally, the experience of participating in children’s learning could foster a broader interest in child education, resulting in increased aspirations for children and increased 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 exposure to children’s education and interactions with CHAMP implementers could result in increased learning of mothers. Empowerment of mothers could also increase as a result of the program, through increases in learning, or through the experience of participating in their children’s education.

As with ML, there was a possibility of direct effects of CHAMP on children beyond the involvement of the mother. CHAMP sessions were typically held at the mothers’ homes, and children could be present during these sessions. In addition, the worksheets provided at each session could have been completed without the involvement of the mother. We present evidence on children’s direct participation in CHAMP in Section III below.

The combined intervention (ML-CHAMP) included both the ML and CHAMP programs. 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, ML classes were conducted by volunteers, and CHAMP visits were conducted by separate Pratham staff. The ML-CHAMP intervention was designed to test whether both mothers’ exposure to learning from ML and instruction on how to facilitate learning at home from CHAMP would be necessary for mothers to engage with their children and improve children’s 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.

Although the programs were scheduled to be held from September 2011 through June 2012, the agricultural cycle led to lapses in implementation during this period. Programs were run between September 2011 and February 2012, upon which time the harvest season began in both states. In Bihar, both ML and CHAMP continued with lower intensity in March and April, upon which time they stopped running. In Rajasthan, programs were halted in March and April and began again in May and June.

## II. Research Design

### A. 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 maternal 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 maternal 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.<sup>11</sup> 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 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. 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.<sup>12</sup> These set-by-phase groups of 10 villages formed the strata for the randomization.<sup>13</sup> Online Appendix Figure A1 provides a graphic representation of the stratification procedure.

### *B. Data collection*

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 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.

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 children’s learning (ASER Centre, 2012). At baseline, the tests were administered to target mothers, children aged 5 to 8 in their households, other children in grades 1 to 4 outside of that age range, and

<sup>11</sup>In online Appendix B we provide more detail on how hamlets were selected in each state.

<sup>12</sup>Apart from the small difference in implementation timing, Phase 1 villages were identical to Phase 2 villages in program administration and content.

<sup>13</sup>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.

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 baseline. The mothers' test contained the same questions as the children's 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.

The household surveys were administered to target mothers. These surveys were designed to measure the mothers' involvement in household decisions, mothers' interest and participation in their children's education, and children's schooling status. Mothers' interest and participation were measured through modules on participation in learning at home, time use of the mother and child, and presence of educational materials in the household. A number of questions referred to a single child aged 5 to 8 in the household. If the household had more than one child in that age range, one was randomly selected to keep the length of the survey manageable.<sup>14</sup> The surveys also measured schooling status of all children in the household.<sup>15</sup> While most questions appeared in both survey rounds, some additions were made to the endline survey based on observations made during implementation.<sup>16</sup>

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 mothers were relatively low: the tests focused primarily on first-grade competencies, yet the average test score in the control group was 30 percent in the language section and 22 percent in the math section.

Columns 2, 3, and 4 of Table 1 present differences in means of the baseline variables between each treatment group and the control group. Two of the 21 variables are 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, four are significant at the 10 percent level, and six more are significant at the 5 percent level—slightly more than what would be expected by pure chance, particularly in the comparisons between the ML-CHAMP and control groups. However, many of

<sup>14</sup>The questions that focused on a single child included the questions on mothers' and children's time use, as well as several questions in the sections on participation in children's education and children's schooling behavior. We do not find evidence that the focus of the surveys on a single child had impacts on that child's test scores. See online Appendix C for details.

<sup>15</sup>In the baseline survey schooling status was measured for children aged 5 to 8. The age range was expanded to 3 to 14 in the endline survey.

<sup>16</sup>We also administered short questionnaires to fathers at baseline and endline. These questionnaires contained questions on participation in children's 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.

TABLE 1—RANDOMIZATION CHECK

	Relative to Control					
	Mean				P-value:	
	Control	ML	CHAMP	ML-CHAMP	Coeffs 0	N
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Household-level Variables</i>						
First principal component of durables ownership <sup>a</sup>	-0.0341 [2.25]	0.00735 (0.0859)	0.0939 (0.0944)	0.0306 (0.0899)	0.747	8888
Main source of income: farming	0.494 [0.500]	0.0137 (0.0222)	0.00796 (0.0243)	0.0526 (0.0218)	0.064	8885
Number of children 0-4	0.929 [0.900]	-0.0112 (0.0298)	0.0179 (0.0332)	0.0675 (0.0297)	0.037	8888
Number of children 5-8	1.44 [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	2.91 [1.50]	0.0263 (0.0569)	0.127 (0.0610)	0.117 (0.0562)	0.071	8888
Fraction of household 15 and over that can read <sup>b</sup>	0.389 [0.426]	-0.0137 (0.0173)	0.0189 (0.0189)	0.00887 (0.0183)	0.365	7576
Fraction of household 15 and over that can do math	0.246 [0.376]	0.00415 (0.0154)	0.0284 (0.0181)	0.0254 (0.0160)	0.220	7579
Mother's education (years)	0.764 [2.28]	0.0475 (0.102)	0.152 (0.118)	0.0694 (0.103)	0.625	8864
Father's education (years) <sup>c</sup>	3.88 [4.44]	-0.150 (0.203)	0.133 (0.226)	0.234 (0.213)	0.290	8181
Mother's age	32.3 [7.10]	-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
Mother's language score (fraction)	0.299 [0.247]	0.00442 (0.0110)	0.0134 (0.0131)	0.00548 (0.0115)	0.785	8857
Mother's math score (fraction)	0.215 [0.241]	0.00639 (0.0110)	0.0106 (0.0129)	0.00598 (0.0115)	0.860	8857
Mother's total score (fraction)	0.250 [0.234]	0.00557 (0.0108)	0.0117 (0.0128)	0.00577 (0.0113)	0.830	8857
<i>B. Child-level Variables</i>						
Child is male	0.520 [0.500]	-0.0132 (0.0108)	-0.00781 (0.0110)	-0.0213 (0.0106)	0.231	15502
Child attends school / aanganwadi <sup>d</sup>	0.801 [0.399]	0.00766 (0.0136)	0.0250 (0.0135)	0.0149 (0.0137)	0.277	15501
Child's language score (fraction)	0.350 [0.296]	0.00753 (0.0110)	0.00912 (0.0113)	0.0103 (0.0109)	0.786	15502
Child's math score (fraction)	0.277 [0.303]	0.0114 (0.0111)	0.0131 (0.0117)	0.00857 (0.0109)	0.669	15502
Child's total score (fraction)	0.310 [0.291]	0.00971 (0.0108)	0.0113 (0.0113)	0.00934 (0.0106)	0.726	15502

TABLE 1—RANDOMIZATION CHECK (CONTINUED)

*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. Standard errors are clustered at the village level.

<sup>a</sup> First principal component of durables ownership is computed based on a set of variables indicating ownership of table/chair, watch/clock, mobile phone, landline, bullock cart, radio, generator, refrigerator, television, electric fan, car/truck, scooter/motorcycle, bicycle, jugaad (a basic motorized transport vehicle), plough, thresher, tractor, and harvester; the number of cows/buffaloes, oxen/bullocks, sheep/goats, pigs, hens/ducks, other birds, and other animals owned; an indicator of a cement/stone/metal/beam/plastic roof; and the number of rooms in the house.

<sup>b</sup> Fraction of household that can read or do math is constructed as the average number of non-mother members 15 years old and over who passed the basic reading or math test, as described in Section II.B of the text, ignoring missing values. These variables have fewer observations due to missing test scores for some household members.

<sup>c</sup> Father’s education variable has fewer observations due to absence of fathers from some households.

<sup>d</sup> Aanganwadi is community-based kindergarten.

these significant differences are small in magnitude. For example, there are 0.055 more children aged 9 to 14 in ML-CHAMP compared with the control group mean of 0.98, and the fraction of male children in ML-CHAMP is 0.021 less than the control group mean of 0.52.<sup>17</sup> Critically, there are no significant differences between any of the treatment groups and the control group in any of the children’s or mothers’ test score variables, our primary outcomes of interest.<sup>18</sup>

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 online Appendix D we present a formal analysis of attrition, including the construction of bounds on the children’s 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.

### III. Program Take-up

Participation in the ML and CHAMP programs is analyzed in Table 2. For ML, we use data from both the endline survey and Pratham administrative records.

<sup>17</sup>Online Appendix Table A2 explores robustness of ML-CHAMP impacts on key outcomes to several different combinations of control variables. The results are virtually unchanged when these different sets of controls are included.

<sup>18</sup>Online Appendix Table A3 presents tests of equality of means between the ML, CHAMP, and ML-CHAMP treatment groups. The number of significant differences at the 10 percent level or less is one for ML vs. CHAMP, three for ML-CHAMP vs. ML, and one for ML-CHAMP vs. CHAMP.

Because the endline survey did not ask questions about CHAMP, we rely on Pratham records for measures of CHAMP participation.

TABLE 2—TAKE-UP OF ML AND CHAMP

	Mean				Difference
	Control	ML	CHAMP	ML-CHAMP	ML-CHAMP - ML / CHAMP
	(1)	(2)	(3)	(4)	(5)
<i>Panel A. Maternal Literacy</i>					
<i>A.1 Endline Survey</i>					
Mother knew about ML classes	0.220 [0.414]	0.628 [0.483]	0.214 [0.410]	0.675 [0.469]	0.0438 (0.0218)
Mother attended ML class	0.0731 [0.260]	0.401 [0.490]	0.0703 [0.256]	0.445 [0.497]	0.0440 (0.0221)
Child attended ML class (with mother or alone) <sup>a</sup>	0.0523 [0.223]	0.268 [0.443]	0.0427 [0.202]	0.324 [0.468]	0.0546 (0.0203)
Child attended with mother	0.0267 [0.161]	0.190 [0.392]	0.0265 [0.161]	0.245 [0.430]	0.0521 (0.0174)
Child attended with mother: always	0.00701 [0.0834]	0.0580 [0.234]	0.00521 [0.0720]	0.0913 [0.288]	0.0309 (0.00849)
Child attended with mother: sometimes / rarely	0.0182 [0.134]	0.132 [0.339]	0.0190 [0.136]	0.151 [0.359]	0.0204 (0.0128)
Child attended alone	0.0376 [0.190]	0.181 [0.385]	0.0311 [0.174]	0.211 [0.408]	0.0316 (0.0162)
<i>A.2 Attendance Records</i>					
Mother attended ML class	–	0.764 [0.425]	–	0.844 [0.363]	0.0674 (0.0265)
Days attended	–	24.7 [31.3]	–	27.3 [30.7]	2.68 (1.92)
Percent of classes attended	–	0.248 [0.259]	–	0.255 [0.241]	0.00752 (0.0121)
<i>Panel B. CHAMP Attendance Records</i>					
Household participated	–	–	0.990 [0.0976]	0.995 [0.0700]	0.00640 (0.00273)
Number of successful visits	–	–	15.8 [5.28]	15.5 [5.22]	-0.302 (0.301)
Percent of visits successful	–	–	0.843 [0.212]	0.849 [0.197]	0.00726 (0.00721)

*Notes:* Columns 1-4 present raw means and standard deviations (in square brackets) of the indicated participation measure in each group. Column 5 presents regression-adjusted differences and standard errors (in parentheses) between the mean in the ML-CHAMP treatment and the ML treatment (Panel A) or CHAMP treatment (Panel B). Regressions in Column 5 control for stratum, and standard errors are clustered at the village level.

<sup>a</sup> *Child* refers to one randomly selected child per household aged 5 to 8 at baseline.

Panel A.1 presents measures of ML attendance, using self reports from the endline survey. About two-thirds of mothers in both the ML and ML-CHAMP recall an adult literacy class running in the village in the past year. We note that this is somewhat low because according to Pratham, all mothers were approached at least once to be recruited for the classes. Mothers who initially refused to

attend or attended only at the beginning may not have recalled these classes, particularly in villages with low attendance among other mothers.<sup>19</sup>

Self-reported take-up of ML classes is also relatively low: 40 percent of mothers in ML and 45 percent of mothers in ML-CHAMP reported having attended ML classes in the past year.<sup>20,21</sup> Children attended the classes as well. The endline survey asked mothers whether and how often the selected child had attended with the mother and whether the child ever attended alone. Nineteen percent of selected children in ML villages and 25 percent of selected children in ML-CHAMP villages were reported to have attended with the mother. Of those, about one third always attended when the mother attended. In addition, about 20 percent of children in ML and ML-CHAMP villages attended alone at least once.

Panel A.2 reports ML attendance collected by Pratham volunteers. These data show take-up rates of 76 percent in ML and 84 percent in ML-CHAMP, substantially higher rates than those self-reported by the mothers. There are two potential reasons for this discrepancy. First, because the administrative data were collected by Pratham volunteers and were not verified, they may overstate actual attendance. Second, self-reported attendance may be subject to recall errors. To avoid prompting, enumerators in the endline survey were not told the treatment status of a given village and therefore did not probe for recall of ML attendance in the treatment villages.<sup>22</sup> Thus, the survey and administrative data likely provide upper and lower bounds on take-up. To avoid these discrepancies, future research examining these types of programs would benefit from measures of attendance observed directly by the research team. In spite of these differences, however, the correlates of take-up are similar across self-reported and administrative measures (see online Appendix F).

Although the administrative data do show a relatively high percentage of mothers attending at least one class, the number of classes attended was low. The

<sup>19</sup>Indeed, among mothers who did not attend, mothers' knowledge of the classes is strongly related to average take-up in the village: 10 percentage points higher take-up in the village is associated with 2.7 percentage points higher likelihood of a non-attending mother knowing about the adult literacy classes (p-value < 0.01, not shown).

<sup>20</sup>The surveys collected self-reports on any participation but not the intensity of participation because it was determined to be logistically infeasible to obtain accurate recall of the number of classes the mother attended over an entire year. It could also lead to surveyors learning the treatment status of the village, which could lead to differential surveyor prompting by treatment status.

<sup>21</sup>We note that 7 percent of mothers in the CHAMP and control groups reported attending classes. 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 (see online Appendix E). 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 in Bihar during the spring of 2012. While virtually no Saakshar Bharat classes were actually held after the initial mobilization, 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.

<sup>22</sup>There is some evidence supporting recall errors. Among those recorded in the administrative data as having attended, recorded attendance is 69 percent higher (41.2 days vs 24.3) for those who also reported attending in the endline survey. Within this group, the number of months since the last recorded attendance is also significantly related to a mother not reporting attending in the survey (results not shown).

average targeted mother in ML and ML-CHAMP attended 25 and 27 classes over the course of the year, respectively. Attendance was also sporadic throughout the year: among those who took up the classes, attendance between September 2011 and February 2012 averaged between 3 and 5 days per month. As noted above, classes were held less consistently from March through the end of the program in June due to the onset of the harvest.

Taken together, these measures indicate relatively low participation in ML classes. According to Pratham staff, much of the low take-up was driven by mothers' interest and availability, although availability of volunteers did impact whether classes were held. While classes were scheduled to be held six days per week (about 24 days per month), the administrative data indicate that 12.5 classes were held on average per month in ML and ML-CHAMP between September 2011 and February 2012. Some of the "missing classes" were likely holidays: the peak festival period in India falls between October and December. However, many could be the result of a lack of volunteer availability or effort (which could itself be an endogenous response to low interest of mothers). Nonetheless, when classes were held, the average selected mother only attended about 25 percent of the time, indicating that the low participation was primarily due to choices of mothers.

As shown in Table 2, there is slightly higher likelihood of attendance in the ML-CHAMP group compared with ML: self-reported take-up was 4.4 percentage points higher, and take-up in the administrative data was 6.7 percentage points higher (both significant at the 5 percent level).<sup>23</sup> However, we note that the magnitudes of the differences are relatively small, representing 11 and 9 percent higher attendance in ML-CHAMP for the self-reported and administrative data, respectively. In addition, the administrative data do not indicate a significant difference in the number of classes attended between the two groups.

CHAMP was a door-to-door intervention where Pratham staff conducted weekly visits with each mother at home. If the mother was not at home, then the CHAMP session was to be conducted with another adult in the household. Panel B of Table 2 presents take-up of CHAMP according to Pratham administrative data. Ninety-nine percent of target households were visited at least once during the intervention period. Households were successfully visited about 16 times on average. As with ML classes, the majority of CHAMP visits were held between September 2011 and February 2012, with fewer visits taking place between March and the end of the program in June.

During the second half of the CHAMP program, attendance data in both Rajasthan and Bihar included information on the mother's and child's participation in CHAMP activities. Within this subsample, mothers were present during 81

<sup>23</sup>When we split the sample by state, only Rajasthan exhibits significantly higher take-up in ML-CHAMP (results not shown). We speculate that this could be driven in part by the lower overall participation of mothers in ML classes in Rajasthan, as described below. Mothers in Rajasthan could have been closer to the margin of participating and were therefore more responsive to the additional contact with Pratham staff.



percent of successful CHAMP visits. In about half of cases, implementers also recorded children’s presence during CHAMP sessions and self-reports of who in the household helped the child with the previous week’s worksheet. These data show that a target-aged child was present during 58 percent of CHAMP visits. The mother helped the child with the CHAMP worksheet 53 percent of the time. In 19 percent of cases another household member helped with the worksheet, while the child completed the worksheet alone 28 percent of the time.

In online Appendix F we present a detailed analysis of the correlates of mothers’ participation in ML and CHAMP. For the ML intervention, several relationships suggest factors that could explain the low take-up discussed above. First, take-up is positively associated with slightly higher baseline test scores of mothers and a small amount of formal education. This highlights a potential challenge in targeting of adult literacy programs: even though they may be meant to reach adults with no education, a small amount of education may signal both experience with classroom learning and an interest in learning more. Second, experience with literacy classes in the past and experience with meeting in groups (as measured through self-help group membership)<sup>24</sup> are also important determinants of the take-up decision. Thus, some mothers may not be comfortable in the group learning environments typical in many adult literacy programs. Third, we observe higher self-reported take-up in Bihar, where Pratham staff reported that mothers were more available to attend than in Rajasthan. This result is echoed by our qualitative interviews, which highlighted a lack of free time as a key constraint to attendance.<sup>25</sup>

Almost all households participated in CHAMP, and hence there is little variation in the extensive margin of participation. Examining the intensity of participation, a higher fraction of children in school and higher baseline child test scores are strong determinants of the percentage of sessions attended. This could reflect a complementarity between the CHAMP material and the school curriculum. In contrast with the determinants of ML participation, intensity of participation in CHAMP has a small negative relationship with mothers’ education and no significant relationship with mothers’ test scores.

#### IV. 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.<sup>26</sup> The PAP includes the weights used in aggregating test scores, the estimating equation for test scores, and the list of intermediate outcomes to be

<sup>24</sup>A self-help group is a group of villagers that pools savings and provides loans to members of the group.

<sup>25</sup>Time spent in household and market work was 43 percent higher Rajasthan than in Bihar. We note, however, that conditional on state, we do not find a significant relationship between time spent on household and market work and participation. On the other hand, availability may also be driven by the opportunity cost of that time, or the value of leisure.

<sup>26</sup><https://www.socialsciceregistry.org/trials/65>

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.

### A. Estimating Equation

Throughout the analysis we utilize the following estimating equation:

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

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 II.A, while  $H_{iv}$  represents a vector of household-level characteristics.  $\varepsilon_{iv}$  is the individual error term, clustered by village, the level of randomization. Equation (1) is estimated using ordinary least squares.

For all impact results using child-level data, we use all variables listed in in Table 1 as controls. When the mother is the unit of observation, we use all household-level variables in Table 1, Panel A, and household-level averages of the child-level variables in Panel B. This includes two variables that were not originally listed in the PAP: child gender and child schooling status. We include these because we find some imbalance in these variables across treatment groups, as shown in Table 1, and because child schooling status is a key outcome of interest. Excluding these variables from the set of controls does not substantively affect any of the main results presented below (not shown).

We present our results in three parts. First, we present effects on children’s test scores. Second, we examine mothers’ learning and involvement in household decisions. Finally, we estimate program impacts on mothers’ participation in their children’s education, presence of education assets in the home, and children’s schooling behavior.

### B. Children’s Test Scores

Table 3 presents the effects of the treatment groups on children’s language, math, and total (math plus language) test scores. We estimate Equation (1), controlling for baseline math and language test scores and the full set of variables presented in Table 1. 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. 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 control variables as zero and include dummy variables to indicate missing values.

TABLE 3—CHILDREN’S TEST SCORES

	Impact of Treatment at Endline			P-Values			
	ML	CHAMP	ML-CHAMP	ML=	All	Additive	N
	(1)	(2)	(3)	(4)	(5)		
Language	-0.00684 (0.0158)	0.0118 (0.0146)	0.0424 (0.0144)	0.225	0.006	0.084	18283
Math	0.0351 (0.0162)	0.0321 (0.0155)	0.0558 (0.0152)	0.850	0.235	0.611	18283
Total	0.0166 (0.0150)	0.0237 (0.0142)	0.0515 (0.0137)	0.624	0.031	0.585	18283

*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 all variables in Table 1. Missing values of control variables are coded as 0, with additional dummies indicating missing values. Column 4 displays the p-value of the F-test that the impacts of ML and CHAMP are equal. Column 5 displays the p-value of the F-test that the ML, CHAMP, and ML-CHAMP impacts are equal. Column 6 displays the p-value of the F-test that the impact of ML-CHAMP equals the sum of the ML and CHAMP impacts. 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.

Column 1 of Table 3 displays the impacts of the ML intervention on language, math, and combined test scores. The intervention had an impact of 0.035 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. As shown in Column 2, the CHAMP intervention had a similar impact on math scores of 0.032 standard deviations (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 0.024 and is statistically significant at the 10 percent level. In Column 3 we display the impacts of the combined ML-CHAMP intervention. The estimated effects are 0.042, 0.056, and 0.052 standard deviations for language, math, and total test scores (respectively), and all three estimates are significant at the 1 percent level.<sup>27</sup> For each category of test score, the ML-CHAMP impacts are larger than the individual ML or CHAMP impacts, and we only reject that the sum of the ML and CHAMP impacts equals the impacts of ML-CHAMP in the case of language, where ML-CHAMP had an impact 0.037 standard deviations larger than the sum of the ML and CHAMP interventions (p-value = 0.084). We note, however, that these tests are relatively low powered: for example, the sum of the ML and CHAMP impacts on math scores has a 95 confidence interval between 0.013 and 0.12.

To understand how the interventions impacted specific skills, online Appendix Table A5 presents effects on a subset of the competencies covered on the test.

<sup>27</sup>In online Appendix Table A4 we evaluate the robustness of the children’s test score 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.

With few exceptions, these competencies correspond to those tested in the ASER national assessments of reading and language (ASER Centre, 2012).<sup>28</sup> We focus on math, where the main results show significant impacts of all three interventions. Among the individual math skills, the ML intervention had 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.

### C. Mothers' Test Scores and Empowerment

Table 4 presents the effect of the interventions on mothers' normalized math, language, and total test scores. We again estimate Equation (1), controlling for baseline language and math scores as well as the control variables listed in Table 1.

All three programs had statistically significant impacts on mothers' language, math, and total test scores. The ML program improved mothers' test scores by 0.066 standard deviations in language, 0.12 standard deviations in math, and 0.096 standard deviations overall. All three estimates are significant at the 1 percent level. As shown in Column 2, The CHAMP program improved mothers' test scores by 0.023 standard deviations in language (significant at the 10 percent level), 0.059 standard deviations in math (significant at the 1 percent level), and 0.043 standard deviations overall (significant at the 1 percent level). As shown in Column 3, the ML-CHAMP intervention had impacts of 0.088, 0.15, and 0.12 standard deviations on language, math, and total test scores, respectively (each significant at the 1 percent level).<sup>29</sup> As with the results for children, the impacts of the combined intervention on mothers' test scores are larger than those of either individual intervention. Indeed, we cannot reject that the effects are additive for any of the three test score categories. In addition, these tests are comparatively well-powered: for example, the 95 percent confidence interval on the sum of the ML and CHAMP math impacts is between 0.12 and 0.23 standard deviations, well above the impact of either individual intervention and relatively close to the point estimate of the ML-CHAMP impact.

In online Appendix Table A6 we examine the effects on a subset of competencies on the test, similar to the exercise for children in online Appendix Table A5. In language, ML and ML-CHAMP had the largest impact on letter recognition, the most basic skill tested. CHAMP, by contrast, had impacts of similar magnitude across a broader range of competencies, although these impacts are considerably smaller than the ML and ML-CHAMP impacts on letter recognition. In math, all three interventions had the largest impacts on one-digit number recognition.

<sup>28</sup>The primary difference between the competencies examined and those on the ASER test is that we include two-digit addition, while the ASER test includes division.

<sup>29</sup>Online Appendix Table A4 examines robustness of these results to exclusion of household-level control variables. As in the case of the results on children, running the impact regression controlling only for baseline test scores leaves the results virtually unchanged.

TABLE 4—MOTHERS’ TEST SCORES

	Impact of Treatment at Endline			P-Values			
	ML	CHAMP	ML-CHAMP	ML=	All	Additive	N
	(1)	(2)	(3)	(4)	(5)		
Language	0.0660 (0.0125)	0.0231 (0.0118)	0.0881 (0.0130)	0.001	0.000	0.956	8580
Math	0.117 (0.0160)	0.0592 (0.0147)	0.150 (0.0163)	0.000	0.000	0.259	8580
Total	0.0955 (0.0124)	0.0432 (0.0113)	0.124 (0.0130)	0.000	0.000	0.440	8580

*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 all variables in Table 1, using household-level averages for child-level variables. Missing values of control variables are coded as 0, with additional dummies indicating missing values. Column 4 displays the p-value of the F-test that the impacts of ML and CHAMP are equal. Column 5 displays the p-value of the F-test that the ML, CHAMP, and ML-CHAMP impacts are equal. Column 6 displays the p-value of F-test that the impact of ML-CHAMP equals the sum of the ML and CHAMP impacts. 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.

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 I, either the ML or CHAMP intervention could promote empowerment by providing mothers with the ability and confidence to make decisions for themselves or for their households. Our main empowerment analysis focuses on the mother’s involvement in household decisions vis-à-vis her husband.<sup>30,31</sup> This dimension is among the most commonly used in studies of empowerment (Malhotra et al., 2002), and we hypothesized that it could be a channel through which the programs could impact children’s outcomes. We include nine indicators of the mother’s involvement in various types of household purchases, as well as involvement in child schooling decisions. These measures are aggregated into an index using the method specified in Kling, Liebman, and Katz (2007).

<sup>30</sup>Enumerators were instructed to isolate the mother from other household members when asking questions on empowerment. This is a similar procedure to that followed in the Demographic and Health Surveys, a set of large-scale international surveys commonly used in studies of gender (ICF International, 2016).

<sup>31</sup>We note that empowerment and the intermediate outcomes in the analysis are based on self-reports from the household 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 online Appendix G.

To construct the index, we compute the normalized value of each variable based on the control group mean and standard deviation at baseline. 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 of one in the control group.<sup>32</sup>

Table 5 displays the impact estimates of the interventions on the index and the component variables. Among the three treatment groups, only ML-CHAMP had a significant impact on the full set of measures, with a point estimate of 0.091 standard deviations (significant at the 1 percent level). The estimates for both ML and CHAMP are positive, smaller in magnitude and not statistically significant.<sup>33</sup> One interpretation of the pattern of estimates is that ML and CHAMP indeed had small impacts on mothers' involvement in household decisions, as the sum of the point estimates of the individual interventions is almost exactly equal to the point estimate of the combined intervention.<sup>34</sup> However, this is speculative as we cannot reject that the impact of either individual program is zero.

Thus, while women's empowerment has been a policy outcome of interest for adult literacy programs in particular, we find limited evidence for impacts on our measure of involvement in household decisions. This case is similar to that of group-based microcredit: while these programs have been credited with increasing women's empowerment, recent rigorous evaluations have found either small positive results on similar measures of involvement in household decisions (Angelucci et al., 2015) or no evidence for impacts (Banerjee et al., 2015; Crépon et al., 2015; Tarozzi et al., 2015).

#### *D. Mothers' Participation, Home Inputs, and Child Schooling*

As described in Section I, we hypothesized that the programs would impact mothers' interest and involvement in their children's education. In this section we present results on mothers' perceptions of their role in child education, their participation in child education, mothers' and children's time use, education assets in the home, and children's school participation.

We measured mothers' perception of their role in childhood education by asking

<sup>32</sup>The variables representing mother involvement in household decisions are a subset of variables originally listed in the PAP. We have narrowed the original list of 26 variables in the PAP to reflect those most frequently used in the literature. We analyze the full set of variables from the PAP in online Appendix H. The additional variables include mothers' mobility, capabilities to perform market transactions, beliefs and attitudes about female education, and happiness. When all 26 variables are included in an aggregate empowerment index, we find impacts of ML of 0.071 standard deviations (significant at the 5 percent level) and impacts of 0.14 standard deviations (significant at the 1 percent level).

<sup>33</sup>The results are qualitatively similar if the components of the index are restricted to the 5 variables not related to children's schooling decisions (not shown).

<sup>34</sup>Results from the endline father questionnaire suggest that these 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 intervention groups compared with the control group (not shown). We note, however, that the response rate of this questionnaire was only 57 percent.

TABLE 5—MOTHERS' INVOLVEMENT IN HOUSEHOLD DECISIONS

	Impact of Treatment at Endline				P-Values		
	Endline Mean Control (1)	ML (2)	CHAMP (3)	ML-CHAMP (4)	ML=CHAMP (5)	Equal (6)	Additive N (8)
Decisions Index <sup>a</sup>	-0.000 [1.00]	0.0440 (0.0340)	0.0378 (0.0316)	0.0906 (0.0347)	0.854	0.285	8581
Involved in purchasing utensils	0.544 [0.498]	0.0325 (0.0171)	0.0330 (0.0177)	0.0425 (0.0182)	0.974	0.835	8546
Involved in purchasing cot	0.458 [0.498]	0.00468 (0.0172)	-0.00827 (0.0167)	0.0133 (0.0175)	0.423	0.426	8500
Involved in purchasing cycle	0.245 [0.430]	0.00572 (0.0138)	-0.0230 (0.0124)	0.00255 (0.0130)	0.031	0.045	8494
Involved in purchasing educational materials	0.482 [0.500]	0.0241 (0.0171)	0.0281 (0.0168)	0.0522 (0.0174)	0.812	0.237	8549
Value of goods can buy alone	2437.8 [2258.8]	112.1 (71.7)	36.0 (72.2)	178.7 (73.9)	0.294	0.176	8581
Involved in deciding girl school enrollment	0.493 [0.500]	0.0229 (0.0170)	0.0203 (0.0163)	0.0400 (0.0175)	0.883	0.510	7821
Involved in deciding girl school type	0.499 [0.500]	0.0199 (0.0167)	0.0217 (0.0166)	0.0509 (0.0164)	0.916	0.142	7757
Involved in deciding boy school enrollment	0.399 [0.490]	0.00415 (0.0153)	0.0186 (0.0149)	0.0266 (0.0155)	0.373	0.420	8238
Involved in deciding boy school type	0.422 [0.494]	0.00562 (0.0157)	0.0265 (0.0150)	0.0323 (0.0165)	0.179	0.242	8211

*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 all variables in Table 1, using household-level averages for child-level variables. Missing values of control variables are coded as 0, with additional dummies indicating missing values. Column 5 displays the p-value of the F-test that the impacts of ML and CHAMP are equal. Column 6 displays the p-value of the F-test that the ML, CHAMP, and ML-CHAMP impacts are equal. Column 7 displays the p-value of the F-test that the impact of ML-CHAMP equals the sum of the ML and CHAMP impacts. Standard errors are clustered at the village level.

<sup>a</sup> The Decisions Index is a normalized average of z-scores of the other variables in the table, using the control group means and standard deviations.

whether they believe they should be responsible for child education at all (relative to fathers), and what activities they could do to assist their children. A detailed analysis is presented in online Appendix I. While most mothers believed they have a role to play in child education, mothers in the ML, CHAMP, and ML-CHAMP villages were 4.1, 3.0, and 4.0 percentage points more likely to report being responsible for child education, respectively (each significant at 5 percent or less). We also asked an open-ended question on what the mothers could do to help children in their studies, and we counted the total number of responses.<sup>35</sup> Relative to a control group mean of 1.9 activities, mothers in ML and ML-CHAMP listed 0.069 and 0.094 more activities (significant at the 10 and 5 levels, respectively).

Turning to actual participation in children’s education, our measures consist of survey questions on mothers’ involvement in the education of the selected 5 to 8-year-old child, focusing on monitoring and supervision of the child’s learning. We include nine measures of 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 children’s education.<sup>36</sup> As with the empowerment measures, we construct an index of mothers’ participation using the nine measures, following the aggregation procedure described above.

As shown in Table 6, we find positive and statistically significant impacts of each program on the mothers’ participation index. The magnitudes are 0.074 standard deviations for ML (significant at the 5 percent level), 0.12 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 examining their children’s notebooks, talking to their children about school, and talking to others about their children’s studies, CHAMP and ML-CHAMP also had significant impacts on mothers helping with homework.<sup>37</sup>

The impact of CHAMP on the participation index is 62 percent higher than the impact of ML (although the difference is not statistically significant). The larger impact of CHAMP is likely due to the CHAMP program’s direct encouragement of participation. In addition, the sum of the ML and CHAMP effects is almost twice as large as the ML-CHAMP impact, and in contrast with the results on test scores, we can reject the null hypothesis that the program effects are additive ( $p$ -value = 0.049). In this case, the programs may have influenced the participation of similar groups of mothers in similar ways, and one program simply crowded

<sup>35</sup>Lists of potential activities were given to the surveyor, but not shared with mothers. If a mother included the activity in her response, the surveyor checked it off from that list.

<sup>36</sup>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 the mother’s learning and is therefore not appropriate for the index. Second, the PAP included educational assets in the participation measures, and we have chosen to analyze impacts on these assets separately.

<sup>37</sup>We note that the recall period of the survey questions overlapped with the end of the CHAMP intervention in Rajasthan, and thus some of the CHAMP results could be driven by activities undertaken by the mother during the intervention. However, the impact of the CHAMP program on the participation index in Bihar only is 0.092 standard deviations and is significant at the 10 percent level (not shown).



TABLE 6—MOTHERS' PARTICIPATION IN CHILDREN'S EDUCATION

	Impact of Treatment at Endline					P-Values	
	Endline Mean			ML=	All	Additive	N
	Control	ML	CHAMP	ML-CHAMP	CHAMP		
(1)	(2)	(3)	(4)	(5)	(6)	(8)	
Mothers' Participation Index <sup>a</sup>	0.000 [1.00]	0.0740 (0.0309)	0.120 (0.0325)	0.106 (0.0329)	0.124	0.281	8573
Takes child to school (times per week) <sup>b</sup>	0.321 [1.13]	-0.0378 (0.0375)	0.0340 (0.0416)	-0.00299 (0.0356)	0.057	0.159	8480
Visited school in past 3 months	0.157 [0.364]	0.00998 (0.0120)	0.0131 (0.0124)	0.0160 (0.0122)	0.805	0.894	8480
Reason for most recent visit was child's learning <sup>c</sup>	0.0598 [0.237]	0.00638 (0.00833)	0.0147 (0.00819)	0.00838 (0.00790)	0.373	0.647	8480
Knows how often child receives homework	0.762 [0.426]	0.00938 (0.0108)	0.00624 (0.0107)	0.00504 (0.0112)	0.755	0.912	8480
Helps children with homework	0.789 [0.408]	0.00427 (0.0124)	0.0287 (0.0123)	0.0219 (0.0120)	0.042	0.112	8572
Opened child's notebook in past 2 months	0.218 [0.413]	0.0325 (0.0133)	0.0618 (0.0132)	0.0466 (0.0130)	0.030	0.094	8572
Talks to children about school (times per week)	2.99 [3.05]	0.203 (0.0925)	0.204 (0.0936)	0.262 (0.0961)	0.990	0.747	8525
Talks to friends or family about child's education (times per week)	1.50 [2.25]	0.254 (0.0721)	0.169 (0.0722)	0.246 (0.0762)	0.220	0.407	8520

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, baseline values (where available), and all variables in Table 1, using household-level averages for child-level variables. Missing values of control variables are coded as 0, with additional dummies indicating missing values. Column 5 displays the p-value of the F-test that the impacts of ML and CHAMP are equal. Column 6 displays the p-value of the F-test that the ML, CHAMP, and ML-CHAMP impacts are equal. Column 7 displays the p-value of the F-test that the impact of ML-CHAMP equals the sum of the ML and CHAMP impacts. Standard errors are clustered at the village level.

<sup>a</sup> The Mothers' 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.

<sup>b</sup> *Child* refers to one randomly selected child per household aged 5 to 8 at baseline.

<sup>c</sup> This variable indicates whether the most recent visit to the child's school was to discuss child studies with school staff, attend a parent-teacher meeting, or to inquire or complain about lack of teacher effort.

out the other when they were combined.

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 children’s learning at home, we hypothesized that they would increase mothers’ time spent directly helping children learn and child time spent learning at home. However, we find limited evidence of these effects. Panel A of Table 7 presents impact estimates of the programs on mothers’ weekly time spent with children and working inside and outside of the household.<sup>38</sup> Across all measures, we see little evidence that the programs impacted time use. 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. Panel B of Table 7 presents the estimated impacts of the programs on weekly time use of the randomly-selected 5 to 8-year-old child in the household. As with the results for mothers, 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 was 0.15 hours per week for CHAMP and 0.20 hours per week for ML-CHAMP, but neither is statistically significant at conventional levels.<sup>39</sup>

Table 8 presents estimated treatment effects on the presence of education assets in the home, as reported by the mother.<sup>40</sup> These assets include pencils, school books, other books, newspapers and magazines, and slates. The first row of Table 8 displays the impacts on an index of the five types of assets, where the index is constructed in the manner described above. The ML, CHAMP, and ML-CHAMP interventions had significant impacts on the index, with magnitudes of 0.057, 0.081, and 0.090 standard deviations. Among the individual asset types, the CHAMP intervention had statistically significant effects on the presence of pencils 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.

We note that although neither intervention provided any of these materials for home use, the ML program did provide small workbooks to mothers. Because the presence of the education assets was self reported we cannot rule out the possibility that some mothers confused these workbooks for school books. When

<sup>38</sup>The measures presented here are a subset of those listed in the PAP. Online Appendix Tables A7 and A8 present impacts on all of the time use variables listed in the PAP for mothers and children, respectively.

<sup>39</sup>The time-use results also provide suggestive evidence that ML and ML-CHAMP increased work hours by both the mother and child. ML-CHAMP increased paid, agricultural and livestock work of the mother by 1.6 hours per week (significant at the 10 percent level), and effect of ML is 1.2 hours per week, falling just short of statistical significance at conventional levels. ML-CHAMP also increased the child’s time spent working in a household business by 0.30 hours (significant at the 5 percent level). We speculate that the skills taught and positive messages about employment in the maternal 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 increase in child skills. However, given the lack of statistical significance of many of these estimates and the large number of coefficients estimated on time use, these results may also be due to chance.

<sup>40</sup>When the mother was unsure that a particular item fell into one of the categories, the surveyor verified it visually.

TABLE 7—MOTHERS' AND CHILDREN'S WEEKLY TIME USE

	Impact of Treatment at Endline				P-Values			
	Endline Mean Control (1)	ML (2)	CHAMP (3)	ML-CHAMP (4)	ML=CHAMP (5)	All Equal (6)	Additive (7)	N (8)
<i>Panel A. Mothers</i>								
Help children with homework	2.31 [2.70]	0.0935 (0.0899)	0.109 (0.0931)	0.0648 (0.0894)	0.861	0.875	0.272	8533
Read with children	0.320 [1.37]	-0.00389 (0.0391)	-0.0377 (0.0356)	-0.00477 (0.0401)	0.350	0.559	0.506	8454
Paid / agricultural / livestock work	47.8 [25.1]	1.23 (0.810)	-0.0243 (0.772)	1.62 (0.831)	0.121	0.111	0.724	8555
Housework	19.0 [8.01]	0.262 (0.299)	0.113 (0.297)	0.246 (0.288)	0.618	0.859	0.761	8581
Look after children	4.65 [3.68]	-0.174 (0.105)	-0.125 (0.107)	0.114 (0.104)	0.624	0.009	0.005	8581
<i>Panel B. Children</i>								
Homework	3.80 [4.31]	-0.0509 (0.136)	0.150 (0.127)	0.203 (0.128)	0.130	0.154	0.585	8404
Housework	3.60 [4.11]	0.0850 (0.133)	0.144 (0.131)	0.175 (0.133)	0.659	0.800	0.774	8449
Household business	1.89 [4.15]	0.207 (0.136)	0.0350 (0.131)	0.299 (0.137)	0.229	0.163	0.770	8447

*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, baseline values (where available), and variables listed in Table 1. missing values of control variables are coded as 0, with additional dummies indicating missing values. Column 5 displays the p-value of the F-test that the impacts of ML and CHAMP are equal. Column 6 displays the p-value of the F-test that the impacts of ML, CHAMP, and ML-CHAMP are equal. Column 7 displays the p-value of the F-test that the impact of ML-CHAMP equals the sum of the ML and CHAMP impacts. Time spent in each activity measured in hours per week. Children's time use was measured for one randomly-selected child aged 5 to 8 in the household. Standard errors are clustered at the village level. \* denotes significance at 0.10; \*\* at 0.05; \*\*\* at 0.01.

TABLE 8—HOME EDUCATION ASSETS

	Impact of Treatment at Endline					P-Values	
	Endline Mean Control	ML (2)	CHAMP (3)	ML-CHAMP (4)	ML=CHAMP (5)	All Equal (6)	N (8)
Education Asset Index <sup>a</sup>	0.000 [1.00]	0.0573 (0.0311)	0.0811 (0.0315)	0.0895 (0.0323)	0.431	0.564	8581
Has pen / pencil	0.945 [0.227]	-0.000317 (0.00794)	0.0126 (0.00680)	0.00783 (0.00746)	0.051	0.149	8581
Has school books	0.906 [0.292]	0.0169 (0.00845)	0.0113 (0.00825)	0.0201 (0.00863)	0.473	0.544	8581
Has other books / comics	0.251 [0.434]	0.0132 (0.0131)	0.0235 (0.0150)	0.0285 (0.0139)	0.483	0.529	8581
Has newspaper / magazine	0.0561 [0.230]	0.00851 (0.00736)	0.0243 (0.00809)	0.00346 (0.00736)	0.050	0.036	8581
Has slate	0.889 [0.314]	0.0148 (0.00997)	-0.00645 (0.0101)	0.0227 (0.00955)	0.037	0.013	8581

*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, baseline values (where available), and all variables in Table 1, using household-level averages for child-level variables. Missing values of control variables are coded as 0, with additional dummies indicating missing values. Column 5 displays the p-value of the F-test that the impacts of ML and CHAMP are equal. Column 6 displays the p-value of the F-test that the impact of ML-CHAMP equals the sum of the ML and CHAMP impacts. Standard errors are clustered at the village level.

<sup>a</sup>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.

school books are excluded from the index, the impact of ML falls to 0.046 and is no longer statistically significant ( $p$ -value = 0.14). CHAMP provided mothers with individual worksheets, which are not easily confused with the asset items. CHAMP also did not directly encourage purchase of any of these items.

Finally, Table 9 presents the impacts of the programs on school participation. Panel A presents impacts on school enrollment and attendance for the sample of children tested at endline.<sup>41</sup> ML-CHAMP increased current or planned enrollment in the upcoming school year by 1.3 percentage points (significant at the 5 percent level) and attendance in school or aanganwadi (community-based kindergarten) by 2.0 percentage points (significant at the 10 percent level).<sup>42</sup> We also find impacts of about 2 percentage points of ML and ML-CHAMP on current attendance in formal school (significant at the 5 and 10 percent levels, respectively). 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. Indeed, as shown in Panel B, the impacts on formal school attendance are nearly double for ML and ML-CHAMP when the sample is restricted to children around the age of initial enrollment in school (ages 4 to 6 at baseline). By contrast, we find little evidence for impacts on private school attendance, school days missed, or monthly tuition expenditures. These results are supported by our qualitative interviews, in which 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.

## V. Discussion

### A. Children’s Attendance in Maternal Literacy Classes

As discussed in Section I, the theory of change of the ML intervention was that mothers’ attendance in the classes would lead to changes in mothers’ decisions and behaviors, ultimately leading to improved children’s learning. However, as we document in Section III, children attended the ML classes in addition to mothers. In this section we present suggestive evidence that attendance of mothers may have served as a more important mediator of impacts compared with children’s attendance. We focus on impacts on math scores, where we find statistically significant impacts of the ML intervention.<sup>43</sup>

<sup>41</sup>Panel A of Table 9 differs from the PAP in two ways. First, in order to more closely follow the test score results we have restricted our sample to children tested at endline (the majority of whom were aged 3 to 8 at baseline), while the PAP measured school participation for all children aged 3 to 14. We find no evidence for school participation on this broader group (results not shown). Second, we include participation in formal school (excluding kindergarten), which was not included in the PAP.

<sup>42</sup>School attendance and aanganwadi attendance were measured through the questions “Does the child go to school or aanganwadi?” and “What grade does the child attend?”

<sup>43</sup>As shown in section III, some children also participated directly in CHAMP. Because we have limited data on the extent and nature of this direct participation, however, we focus the analysis on children’s

TABLE 9—CHILDREN’S SCHOOLING

	Endline Mean Control	Impact of Treatment at Endline				P-Values		
		ML (2)	CHAMP (3)	ML-CHAMP (4)	ML=CHAMP (5)	All Equal (6)	Additive (7)	N (8)
<i>Panel A. Children tested at endline</i>								
Child is enrolled or will be enrolled next school year	0.941 [0.235]	0.00684 (0.00550)	0.00388 (0.00535)	0.0131 (0.00527)	0.564	0.185	0.758	18080
Child attends school / aanganwadia	0.882 [0.323]	-0.000257 (0.0110)	-0.00112 (0.0105)	0.0204 (0.0106)	0.932	0.056	0.145	18283
Child attends school	0.728 [0.445]	0.0249 (0.0102)	0.00623 (0.0106)	0.0199 (0.0111)	0.075	0.191	0.473	18283
Child attends private school	0.0790 [0.270]	0.00849 (0.00843)	-0.00225 (0.00834)	0.00257 (0.00863)	0.207	0.449	0.766	18065
<i>Panel B. Children aged 4-6 at baseline</i>								
Child attends school / aanganwadi	0.888 [0.315]	-0.00110 (0.0134)	0.00182 (0.0128)	0.0205 (0.0126)	0.815	0.151	0.271	8066
Child attends school	0.692 [0.462]	0.0426 (0.0151)	0.00710 (0.0165)	0.0350 (0.0168)	0.023	0.067	0.521	8066
<i>Panel C. Selected children<sup>b</sup></i>								
Hours per day spent in school	4.07 [1.62]	0.0612 (0.0536)	0.0885 (0.0516)	0.0796 (0.0548)	0.583	0.857	0.361	8480
Days of school missed in past week <sup>c</sup>	1.38 [1.94]	0.0176 (0.0698)	-0.0729 (0.0737)	-0.0384 (0.0676)	0.213	0.445	0.871	6993
Days missed due to housework or other work in past month	2.19 [5.42]	-0.136 (0.203)	-0.0305 (0.208)	-0.177 (0.191)	0.634	0.788	0.971	7393
Monthly tuition fees	19.9 [43.9]	2.83 (1.58)	2.55 (1.75)	0.291 (1.47)	0.867	0.192	0.032	8457

*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, baseline values (where available), all variables in Table 1. Missing values of control variables are coded as 0, with additional dummies indicating missing values. Column 5 displays the p-value of the F-test that the impacts of ML and CHAMP are equal. Column 6 displays the p-value of the F-test that the ML, CHAMP, and ML-CHAMP impacts are equal. Column 7 displays the p-value of the F-test that the impact of ML-CHAMP equals the sum of the ML and CHAMP impacts. Standard errors are clustered at the village level.

<sup>a</sup> School and aanganwadi attendance were measured through the mother’s answer to the questions *Does the child go to school or aanganwadi?* and *What grade does the child attend?*

<sup>b</sup> *Selected child* refers to one randomly selected child per household aged 5 to 8 at baseline.

<sup>c</sup> Days of school missed are coded as missing for children not in school.

We first analyze the potential impacts of mothers’ and children’s attendance by exploiting the design of the experiment. As described in Section II, the treatment assignment was stratified geographically for ease of implementation. Sets of 20 geographically proximate villages were selected, and these sets were randomly split into two phases of 10 villages each to determine the order of program rollout. The randomization was stratified by these 48 set-by-phase groups of 10 villages. We thus have 48 “separate” randomized experiments within the larger experiment of 480 villages. In addition to geographic differences and slight differences in the timing of program initiation, each Pratham supervisor was assigned to a set of 20 villages, potentially giving rise to implementation differences across strata. These implementation differences—such as in the timing of classes and in the of recruitment of mothers—could give rise to different patterns of take-up of mothers and children in each strata. To the extent that these take-up patterns influenced program impacts, we can examine the relationship between mothers’ or children’s take-up in each stratum and the learning impacts on children.

For each stratum, we estimate 48 separate ML treatment effects on mothers’ attendance, children’s attendance, and children’s test scores using the specification in Equation (1).<sup>44</sup> For a consistent measures of attendance for both mothers and children across treatment and control groups, we rely on self-reported attendance from the endline survey. We then correlate the treatment effects on children’s math scores with the effects on mother’s or children’s attendance. The results are plotted in online Appendix Figure A2. There is a clear positive relationship between the stratum-wise ML treatment effects on children’s math scores and mothers’ attendance ( $\rho = 0.31$ , p-value = 0.03), but not children’s attendance ( $\rho = -0.084$ , p-value = 0.57). These relationships are, of course, correlations, but they do suggest that attendance of mothers, rather than attendance of children, is most closely related to the impacts on children’s test scores.

We next analyze the potential effect of attendance by including attendance of mothers and children as regressors in the main children’s learning specification from Section IV.B. Inclusion of endogenous intermediate outcome variables in an impact estimation can provide suggestive evidence for the influence of these variables as mediators of impacts (Baron and Kinney, 1986). When the intermediate outcome variable is included, a decrease in the coefficient of the treatment variable and a positive coefficient on the intermediate variable suggests that this variable is a mediator of the treatment effect. However, this rests on several strong assumptions (Green et al., 2010; Imai et al., 2011). In particular, the analysis assumes that the intermediate variable is uncorrelated with other factors that influence the final outcome variable. If, for example, the intermediate variable is positively correlated with these unobserved factors, then the variable may appear to be a mediator due to omitted variable bias. Alternatively, if the intermediate variable is measured with error, then the analysis will tend to underestimate the

attendance in ML classes.

<sup>44</sup>This is implemented by interacting the treatment groups with dummies for the 48 strata.

effect of mediation.<sup>45</sup>

Table 10 presents the results of the analysis, including combinations of mothers' and children's attendance variables in the regression of children's test scores on the treatment groups. Again, we rely on self-reported attendance from the endline survey, as it provides a consistent measure across treatment and control groups. Because we only have attendance data for one 5 to 8-year-old child per household, we restrict the sample to those children, and hence the impacts using the main specification from Section IV.B are noisier and fall just short of statistical significance at conventional levels (p-value = 0.12). Columns 2 and 3 include dummies for attendance of either the mother or child in the regressions. In each specification, the coefficient on the ML treatment decreases by about one-third, and the coefficient on the attendance dummy is about 0.04 (both significant at the 5 percent level). As shown in Column 4, when we include both variables in the same regression, the ML treatment variable decreases by 43 percent. There is a slightly larger coefficient on attendance of the child compared with that of the mother (0.028 vs 0.019), but neither is statistically significant. Finally, Column 5 includes both attendance of the mother and two richer measures of attendance of the child—attendance with the mother weighted by frequency<sup>46</sup> and attendance without the mother. In this regression, we observe a 39 percent decrease in the coefficient on ML. The coefficient on the mother's attendance (0.027) is now larger than that of the child's attendance with the mother (0.019) but again neither is statistically significant. The coefficient on the variable indicating the child's attendance without the mother is effectively zero. In slight contrast with the stratum-wise analysis above, the results in this table suggest that attendance of the mother and the child could have had similar mediating effects on the children's learning results.<sup>47</sup>

The slightly different results of the two analyses presented in this section are likely due to the use of different sources of attendance in each one. When we isolate variation in attendance due to geographic and implementation factors, there is no evidence that children's attendance mediated the impacts. When we use variation at the individual level by including attendance measures in our learning regressions, the analysis suggests that both could have served as mediators. Nonetheless, the results suggest that while we cannot rule out the possibility that children's participation in ML influenced the learning results, much of the impact appears to be mediated by participation of the mothers.

<sup>45</sup>See Green et al. (2010) and Imai et al. (2011) for a more formal discussion of these issues.

<sup>46</sup>The weighted variable equals 1 for "always", 0.67 for "sometimes," and 0.33 for "rarely."

<sup>47</sup>It is also important to note that attendance of the child only appears effective when the child attended with the mother rather than alone, suggesting that to the extent that children did learn through attendance in ML, it only occurred alongside their mothers.



TABLE 10—CHILDREN’S MATH IMPACTS: CONTROLLING FOR ML ATTENDANCE

	Dependent Variable: Child’s Math Score				
	Included Attendance Variable(s)				
	Main	Mother Attended	Child Attended	Mother Attended and Child Attended	Mother Attended and Child Attended
	(1)	(2)	(3)	(4)	(5)
ML	0.0337 (0.0214)	0.0213 (0.0219)	0.0227 (0.0217)	0.0191 (0.0219)	0.0206 (0.0219)
CHAMP	0.0500 (0.0207)	0.0492 (0.0207)	0.0493 (0.0207)	0.0492 (0.0207)	0.0489 (0.0207)
ML-CHAMP	0.0546 (0.0197)	0.0404 (0.0209)	0.0417 (0.0202)	0.0380 (0.0209)	0.0395 (0.0209)
Mother attended		0.0354 (0.0162)		0.0192 (0.0191)	0.0272 (0.0195)
Child attended			0.0406 (0.0171)	0.0284 (0.0200)	
Child attended with mother (weighted) <sup>a</sup>					0.0194 (0.0324)
Child attended alone					-0.00439 (0.0223)
R-Squared	0.709	0.709	0.709	0.709	0.709
N	8367	8311	8262	8262	8241

*Notes:* This table presents regressions using the specification from Table 3, including self-reported measures of mother’s and child’s attendance measured in the endline survey as regressors. The sample consists of one randomly selected child aged 5 to 8 in the household.

<sup>a</sup> Child’s attendance with mother is weighted by frequency of attendance: *Always* = 1, *Sometimes* = 0.67, and *Rarely* = 0.33.

### B. Mediation of Impacts Through Intermediate Outcomes

The ML, CHAMP, and combined ML-CHAMP programs influenced a number of intermediate outcomes that could have led to the observed impacts on children’s learning. In the case of ML, we observe significant impacts on mothers’ average learning levels, increased participation in their children’s education as measured through our participation index, but not an increase in time spent directly helping their children with homework. We also find impacts on education assets available in the home and on attendance in formal schooling. CHAMP had smaller impacts on mothers’ learning, but larger impacts on mothers’ participation and education assets.

In Table 11 we present suggestive evidence of the importance of these intermediate outcomes by including these measures as regressors in the estimation of impacts on children’s test scores, as was done in the previous section with measures of attendance. As before, we focus on math outcomes, where we observed significant impacts from all three interventions. We also restrict our sample to the one selected 5 to 8-year-old child per household to parallel the analysis in the previous section, and because some of the survey questions in the participation

index focused on those children.

As in Table 10, each column of Table 11 includes a separate variable or variables representing the intermediate outcomes, with the first column displaying the results using the specification from Section IV.B for reference. In all cases, we include baseline values of the intermediate outcomes as controls, and thus the displayed coefficients on the endline variables can be interpreted as their influence *relative* to baseline. Column 2 includes endline mothers' math and language scores. Inclusion of these variables decreases coefficient on the ML treatment dummy by 43 percent and the impact of CHAMP by 12 percent. One standard deviation higher math (language) score of the mother is associated with a higher child's test score of 0.073 (0.053) standard deviations (both significant at the 1 percent level). As shown in Column 3, inclusion of the participation index reduces the coefficient on the ML treatment by 13 percent and the CHAMP coefficient by 11 percent. We observe similar reductions in the coefficients when the education asset index is included. When we include formal school attendance, the coefficient on ML drops 34 percent, and the coefficient on CHAMP drops 3 percent. As shown in Column 6, when all intermediate variables are included in the regression, the coefficient on ML decreases 82 percent, and the coefficient on CHAMP decreases 22 percent. In this regression, the mothers' test score variables, the asset index, and the school attendance variables have about the same magnitude as in the individual regressions and remain statistically significant, while the coefficient on the participation index becomes closer to zero and is not statistically significant. Finally, in Column 7 we include attendance of the mother and child in the ML classes to the specification in Column 6. Inclusion of these variables has virtually no influence on the coefficients on the other intermediate outcomes.

Although we emphasize that this analysis rests on strong assumptions and must be taken as suggestive, the patterns from the analysis highlights several potential pathways for impact. First, our index of mothers' participation in children's learning appears to be a modest mediator of the learning impacts, while mothers' learning appears much stronger, both in the regressions where each measure is included individually and in the regressions where all mediators are included together. This could suggest that mothers' learning could be influencing the *quality* of interactions beyond what is measured in the participation index, which primarily measures the different types of involvement of the mother. Alternatively, it could suggest that the participation index is measured with error, in that it does not fully reflect all of the potential activities the mother could do with her child. Second, the analysis suggests that formal school attendance is a strong mediator of impacts for the ML intervention. This implies that ML could have increased interest in formal schooling, and mothers responded by ensuring attendance in school—a margin they considered particularly feasible, according to our qualitative interviews. Third, when all intermediate outcomes are included together, they account for almost all of the ML impact, but only about a quarter of CHAMP impact. This could imply more potential for direct impacts of CHAMP

TABLE 11—CHILDREN’S MATH IMPACTS: CONTROLLING FOR INTERMEDIATE OUTCOMES

	Dependent Variable: Child’s Math Score						
	Included Intermediate Variable(s)						
	Main (1)	Mother’s Learning (2)	Participation Index (3)	Education Asset Index (4)	Attends School (5)	All Intermediate (6)	Intermediate With Attendance (7)
ML	0.0337 (0.0214)	0.0192 (0.0214)	0.0294 (0.0214)	0.0288 (0.0212)	0.0223 (0.0209)	0.00608 (0.0208)	0.000551 (0.0214)
CHAMP	0.0500 (0.0207)	0.0439 (0.0208)	0.0446 (0.0207)	0.0437 (0.0206)	0.0484 (0.0202)	0.0390 (0.0201)	0.0393 (0.0201)
ML-CHAMP	0.0546 (0.0197)	0.0379 (0.0198)	0.0495 (0.0196)	0.0483 (0.0195)	0.0538 (0.0188)	0.0352 (0.0188)	0.0288 (0.0198)
Mother’s language score		0.0530 (0.0178)				0.0475 (0.0176)	0.0476 (0.0177)
Mother’s math score		0.0729 (0.0135)				0.0666 (0.0136)	0.0654 (0.0137)
Participation Index			0.0240 (0.00587)			-0.00390 (0.00606)	-0.00319 (0.00611)
Education Asset Index				0.0522 (0.00672)		0.0343 (0.00685)	0.0338 (0.00688)
Child attends school					0.241 (0.0174)	0.220 (0.0183)	0.220 (0.0184)
Mother attended ML							-0.00118 (0.0192)
Child attended ML							0.0236 (0.0197)
R-Squared	0.709	0.710	0.710	0.711	0.721	0.724	0.724
N	8367	8307	8310	8311	8367	8298	8249

Notes: This table presents regressions using the specification from Table 3, including intermediate outcomes listed and binary measures of self-reported attendance in ML classes as regressors. The sample consists of one randomly selected child aged 5 to 8 in the household. Participation and Education Asset Indices are defined in Tables 6 and 8, respectively.

on the child beyond the involvement of the mother, but it is also possible that our measures are simply not picking up the precise changes in the home environment that led to increased children’s learning.

In addition, the specification of Column 7 shows that the influence of these intermediate outcomes is not confounded by children’s attendance in ML, as the coefficients do not change when the attendance of the child is included in the regression. This provides further evidence to the finding of the previous section that the children’s learning results in ML were largely driven by mothers’ attendance. Nonetheless, we emphasize these analyses presented should be viewed as suggestive, and research designs explicitly testing the influence of specific intermediate outcomes as well as mothers’ and children’s attendance in these programs would be a fruitful area for future work.

### *C. Cost Effectiveness*

This section presents a discussion of cost effectiveness of the three interventions in improving children’s and mothers’ 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-benefit calculation follows the methodology in Kremer, et. al (2013).

Details of the cost-benefit calculation are provided in online Appendix Table A9. While both interventions included a similar amount of Pratham staff time—either supervising the volunteers and classes in ML or running the sessions in CHAMP—the cost of the ML intervention also included the opportunity costs of volunteer time. This difference leads to much higher cost per beneficiary household in ML (\$34 compared with \$17 per household in CHAMP).

For children, each \$100 spent on ML, CHAMP, and ML-CHAMP resulted in improvements in math scores of 0.22, 0.40, and 0.24 standard deviations, respectively. The greater cost-effectiveness of CHAMP is driven by the lower costs of the intervention, as both ML and CHAMP had similar treatment effects on children. For mothers, each \$100 spent on the ML, CHAMP, and ML-CHAMP interventions resulted in improvements in mothers’ total test scores of 0.28, 0.26, and 0.24 standard deviations, respectively. The similarity in cost effectiveness between the ML and CHAMP interventions is due to the fact that while ML was about twice as expensive, the effects of ML on mothers were twice as large as the effects of CHAMP.

Because there are no other cost-benefit calculations of program effects on adult learning that we know of, we focus our comparison with the literature on learning impacts on children. Our impact estimates, ranging from 0.03 to 0.06 standard deviations, fall below the range of statistically-significant estimates in the Kremer et al. (2013) study. This latter set of estimates ranges from 0.13 standard deviations to 0.59 standard deviations. Out of the 15 studies in Kremer et al.

(2013) that found statistically significant impacts, 14 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 mothers' learning—a key policy objective of adult literacy programs. However, given the lack of data for other adult literacy programs, we are unable to perform comparisons along these additional dimensions. 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.

## VI. Conclusion

In developing countries, adult literacy programs have been credited with improving children's education outcomes by encouraging parental investments in education and an improved home learning environment. Programs that encourage parental involvement have also been cited as a means of improving children's education through similar channels. In spite of the potential of these programs to achieve these objectives, there has been little rigorous evidence on whether they are actually effective.

Our paper evaluates the impacts of 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 each of these three programs had modest impacts on children's learning levels, particularly math scores.

The programs also influenced several intermediate outcomes related to decisions of mothers and the home learning environment that could in turn have affected children's learning: all three had impacts on mothers' learning levels, interest in participation in child education, reported levels of participation, and the presence of educational materials at home. By contrast, we find no significant impacts on time spent directly helping children with homework. This suggests that child outcomes could have been influenced by the quality of time spent, or frequency or type of interactions, even if there was no increase in time spent helping with homework. We also find that the maternal literacy and combined programs increased children's attendance in formal schooling.

A more direct mechanism could be that children learned by participating directly in the programs, independent of how the mother was affected. Indeed, children did attend ML classes and were often present during CHAMP sessions. If children's participation was found to be the sole mechanism through which children learned, it would call in to question a key assumption in the programs' theory of change—that the decisions of mothers and the home environment were integral inputs into their children's learning. Based on the wishes of the partner NGO, it was not feasible to prevent children from attending ML classes or to prevent direct involvement of children in CHAMP. Our experimental design thus does not allow us to rule out the influence of direct participation of children on the impacts on children's test scores.

However, in the case of ML classes, we provide suggestive evidence that attendance of children was likely not the primary mechanism for the children’s learning impacts. We also show that for the ML intervention, improvements in mothers’ learning levels, participation in their children’s education, presence of educational materials at home, and children’s school attendance can almost entirely account for the impacts we observe on children’s learning. Of these, improvement in mothers’ learning itself appears to be the largest mediator of impacts. For the CHAMP intervention, the analysis shows that the intermediate outcomes were smaller mediators of children’s learning, suggesting a greater role of direct children’s participation or household involvement that was not measured in our surveys.

We conclude that although we cannot fully isolate the specific channels through which the programs influenced children’s learning, the evidence is generally consistent with our hypothesis that the programs would influence children’s learning through changes in mothers’ decisions and the home learning environment—mechanisms previously identified in studies of the intergenerational transmission of human capital. However, we note that while we are able to show that similar mechanisms *can* be activated by adult literacy and participation programs, we cannot speak directly to the precise mechanisms at play in the broader literature.

Because of the relatively low magnitude of the impacts on children’s learning, the interventions we study are not cost effective relative to interventions that have been studied elsewhere. The low effect sizes were likely a result of low take-up. Most mothers reported never attending the classes, and those who did, attended infrequently. Our experience with low attendance of mothers 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 suggests scope for larger impacts with more intensive interventions. In our key results on children’s and mothers’ test scores, the point estimates for ML-CHAMP are approximately equal to the sum of the ML and CHAMP impacts. 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. Evaluation of such strategies would be 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

- Abadzi, Helen.** 2003. “Adult Literacy: A Review of Implementation Experience.” *The World Bank Operations Evaluation Department*.
- Aker, Jenny C., and Christopher Ksoll.** 2015. “Call Me Educated: Evidence from a Mobile Monitoring Experiment in Niger.” *Working Paper, Tufts University*.

- Aker, Jenny C., Christopher Ksoll, and Travis J. Lybbert.** 2012. "Can Mobile Phones Improve Learning? Evidence from a Field Experiment in Niger." *American Economic Journal: Applied Economics*, 4(4): 94–120.
- Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja.** 2012. "What Did you Do All Day? Maternal Education and Child Outcomes." *Journal of Human Resources*, 47(4): 873–912.
- Angelucci, Manuela, Dean Karlan, and Jonathan Zinman.** 2015. "Micro-credit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco." *American Economic Journal: Applied Economics*, 7(1): 151–182.
- Aoki, Aya.** 2005. "Assessing learning achievements and development impact: Ghana's national functional literacy program." *Australian Journal of Adult Learning*, 45(1): 63–81.
- ASER Centre.** 2012. "Annual Status of Education Report 2011." Accessed August 17, 2016.
- Avvisati, Francesco, Bruno Besbas, and Nina Guyon.** 2010. "Parental Involvement in School: A Literature Review." *Economie De L'Education*, 120(5): 759–778.
- Avvisati, Francesco, Marc Gurgand, Nina Guyon, and Eric Maurin.** 2014. "Getting Parents Involved: A Field Experiment in Deprived Schools." *The Review of Economic Studies*, 81(1): 57–83.
- Baker-Henningham, Helen, and Florencia Lopez Boo.** 2010. "Early Childhood Stimulation Interventions in Developing Countries: A Comprehensive Literature Review." *IZA Discussion Paper No. 5282*.
- Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan.** 2015. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." *American Economic Journal: Applied Economics*, 7(1): 22–53.
- Banerjee, Abhijit V., Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani.** 2010. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy*, 2(1): 1–30.
- Baron, Reuben M., and David A. Kinney.** 1986. "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology*, 51(6): 1173–1182.
- Beder, Hal.** 1999. "The outcomes and impacts of adult literacy education in the United States." *National Center for the Study of Adult Learning and Literacy Report 6*.
- Behrman, Jere, Andrew Foster, Mark Rosenzweig, and Prem Vashishtha.** 1999. "Women's Schooling, Home Teaching, and Economic Growth." *The Journal of Political Economy*, 107(4): 682–714.

- Carron, G.** 1990. "The Functioning and Effects of the Kenya Literacy Program." *African Studies Review*, 33(3): 97–120.
- Crépon, Bruno, Florencia Devoto, Esther Duflo, and William Parienté.** 2015. "Estimating the Impact of Microcredit on Those who take it up: Evidence from a randomized experiment in Morocco." 1(7): 123–150.
- Duflo, Esther, Pascaline Dupas, and Michael Kremer.** 2015. "School governance, teacher incentives and pupil-teacher ratios: Experimental Evidence from Kenyan Primary schools." *Journal of Public Economics*, 123: 92–110.
- Gertler, Paul, Harry Anthony Patrinos, and Rubio-Codina.** 2008. "Empowering Parents to Improve Education: Evidence from Rural Mexico." *World Bank Policy Research Working Paper No. 3935*.
- Government of India.** 2011. "Census of India 2011: Population Enumeration Data." Accessed July 5, 2016.
- Green, Donald P., Shang E. Ha, and John G. Bullock.** 2010. "Enough already about "black box" experiments: Studying mediation is more difficult than most scholars suppose." *Annals of the American Academic of Political and Social Science*, 628: 200–208.
- Guryan, Jonathan, Erik Hurst, and Melissa Kearney.** 2008. "Parental Education and Parental Time with Children." *Journal of Economic Perspectives*, 22(3): 23–46.
- ICF International.** 2016. "The DHS Program - Protecting the Privacy of DHS Survey Respondents." Accessed June 20, 2016.
- Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto.** 2011. "Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies." *American Political Science Review*, 105(4): 765–789.
- Kagitcibaci, Cigdem, Diane Sunar, and Sevda Bekman.** 2001. "Long-term effects of early intervention: Turkish low-income mothers and children." *Applied Developmental Psychology*, 22(4): 333–361.
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz.** 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica*, 75(1): 83–119.
- Kremer, Michael, Conner Brannen, and Rachel Glennerster.** 2013. "The Challenge of Education and Learning in the Developing World." *Science*, 340: 297–300.
- Lee, David S.** 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies*, 76(3): 1071–1102.
- Ortega, Daniel, and Francisco Rodriguez.** 2008. "Freed from Illiteracy? A Closer Look at Venezuela's Mision Robinson Literacy Campaign." *Economic Development and Cultural Change*, 57(1): 1–30.
- Pritchett, Lant.** 2013. *The Rebirth of Education: Schooling Ain't Learning*. Center for Global Development.



- Tarozzi, Alessandro, Jaikishan Desai, and Kristin Johnson.** 2015. "The Impacts of Microcredit: Evidence from Ethiopia." *American Economic Journal: Applied Economics*, 7(1): 54–89.
- UNESCO.** 2005a. *Aspects of Literacy Assessment: Topics and Issues from the UNESCO Expert Meeting.*
- UNESCO.** 2005b. *Education for All Global Monitoring Report 2006: Literacy for Life.*
- UNESCO.** 2015. *Education for All Global Monitoring Report 2015: Education for All 2000-2015: Achievements and Challenges.*
- Wagner, Daniel A.** 2000. "EFA 2000 Thematic Study on Literacy and Adult Education." *International Literacy Institute.*