Report on Second Survey to Assess CONAFE "Early Childhood Education" Program Impact

Table of Contents

List of Illustrations

1. Report Presentation

Since 2011, the Center for Economic Research and Teaching (CIDE) has executed the Impact Assessment, through financing by the International Initiative For Impact Evaluation (3iE) and assistance from the World Bank, Washington office. This is an impact study for the purpose of estimating the effects of "Early Childhood Education" implemented by the National Council for Educational Development (CONAFE). This investigation seeks to generate solid evidence for decision makers, researchers and public policy makers, mainly in the educational field.

The first report, published in February 2013, contains information on the design of the assessment and the activities to be performed from the start of the project to baseline information gathering. The main discovery in the first run was that the treated and control groups were statistically similar in the majority of the contextual variables, parenting practices and skills development for three and a half year olds. This means the randomization was done satisfactorily and any change in the corresponding variable observed at the completion of the assessment shall be simpler to isolate to identify the potential causal relationship with the Early Childhood Education Program intervention.

This document represents the second data delivery by CIDE and contains results from both the second and first waves of the survey, which means that the information is integrated with data from the second wave and that collected in the baseline phase. Two main objectives are pursued with this report. First, to present an account of the work group activities of this year: from January up until the end of the first follow-up data collection in July and August. Second, to describe the results obtained after the data collection from the second run and compare them with those derived from the baseline in order to have an updated panorama of some of the key measures of interest comprising the study. In this regard, it is worth noting that, at this point of the investigation, the expectation is to identify preliminary changes in the dependent variables in this study, such as parenting practices or children's level of development. Therefore, results in some of the measures of both main constructs are described and discussed with the appropriate caution.

2. Instruments

In the results presented in the first report, it was mentioned that given the main of interest, four types of general instruments were designed and used a fifth type that was developed externally. In this panel run it was necessary to design a second version of each one of the four general questionnaires of the first survey; the differences between the first and the second versions are in response to information derived from the pilot tests and observations made during the pilot survey. In this respect, it is not superfluous to recall that each one of the versions has a pair, according to the group to which it is administered: treatment group or control group. For example, a questionnaire that captures the context of children in treatment communities (if receiving the program) and a questionnaire containing almost exactly the same variables of interest but administered in control communities (not receiving the program). Each one of the questionnaire types shall be briefly described below.

Firstly, the Home questionnaire gathered information about the number and characteristics of household members, along with dwelling conditions and possession of goods, to ascertain the family's level of wellbeing. Secondly, the Caregivers questionnaire contained questions about didactic materials available to the children, and how activities toward children are performed by the closest adult and how they favor the development of certain skills of interest, depending on age range. Thirdly, the Pregnant Women questionnaire scrutinizes the habits of the mother during pregnancy and the initial care of her child.

The fourth questionnaire contains questions related to fathers by which information regarding social conduct and the level of involvement in child education is gathered. Finally, the Ages and Stages Questionnaire assesses the main dependent variables regarding the evaluation study: child skills development. With these questionnaires it is possible to evaluate six development areas according to twenty different age ranges[1]; the areas are: communication, gross motor, fine motor, problem solving and personal-social skills. With regard to ASQ's implementation protocol, note that during the baseline survey questionnaires were administered for an age range of 2 to 42 months, hoping that the majority of the children would remain in the study during the subsequent two years. As such, in the survey performed last July questionnaires included an age range of 48, 54 and 60 months, so it would be possible to document the skills of the older children that comprised the initial sample.

[1] The age ranges in ASQ surveys are established in months using the following increments: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 27, 30, 33, 36, 42, 48, 54, 60.

From another perspective, it is important to mention that data collection considered two potential changes in communities: 1) the integration of new homes to the program (only treatment communities), and 2) the formation of new households in both community types (treatment or control), that is, households that after the baseline survey might have been separated to become two distinct family nuclei. Both possibilities were addressed in the design of so-called "First Interview" questionnaires. The primary objectives of these instruments were to focus on new and comparable cases that receive the service -paragraph 1) - and decrease the likelihood of cases in non-response rate -paragraph 2).

3. Training

The second assessment survey was carried out in five states simultaneously during the month of July, plus one in August, so there was the need for a well-trained team. Data collection was conducted by personnel from the research company *Soluciones Q*, hired through a competitive bidding process to provide those services; the task was performed under the supervision of CIDE field staff. As part of the requirements, personnel hired must had to have previous experience in gathering information and to have at least completed preparatory school. The training program was designed to be completed in 40 hours over five days from July 15 to 19, 2013. The main purpose was for field researchers to properly perform data collection according to a protocol that guarantees data quality and reliability, based on the general comprehension of the Project aims and knowledge of the assessment instruments; the training structure is shown in the table below.

Table 1: Training Activities

Day	Content
Monday *Classroom*	1. General Objective of Project 2. Child Development Concepts 3. ASQ Assessment Instruments
Tuesday *Classroom*	1. Presentation Protocol and Information Gathering Path 2. General Information Gathering Instruments 3. Field Conduct Protocol
Wednesday *Practical*	Practical in the State of Hidalgo. Supervised information gathering in adherence to instrument use and protocol.
Thursday *Practical*	Practical in the State of Hidalgo. Supervised information gathering in adherence to instrument use and protocol.
Friday *Classroom*	Feedback on observations made in the field and response to questions in the classroom.

Eighty-two field researchers attended training within the CIDE facilities. By the end of the second day the first reduction of field researchers was performed so that only the most suitable statistical surveyors would participate in the practical sessions in the State of Hidalgo. The final selection was done on the last day of training, of which 65 surveyors in total remained.

3.1 ASQ-3 Administration Training

The child development program and familiarization with the *Ages & Stages Questionnaire (ASQ)* training was headed by psychologists María del Carmen Hernández and Sandra Martínez, researchers from the National Institute of Perinatology with more than fifteen years experience, specifically in ASQ administration in Mexico.

In one day field researchers were introduced to understanding the basic concepts of child development and assessment. Explained were the nature of ASQ-3, the age ranges for which each instrument was designed and the development areas assessed. After the explanation of concepts, the twenty ASQ assessment instruments were reviewed and videos presentations were shown demonstrating examples of instrument administration and the use of didactic materials[2] employed during assessment.

As part of the training, the type of questions were separated into three codes – A, B and C – and emphasis was placed on the application of code C questions, ones that require interaction between the child, the main caregiver and the field researcher working as an assessor administering the ASQ instrument.

- **Code A**. The question is addressed to the mother or primary caregiver.
- **Code B**. The question implies an action by the child and observation by the field researcher. If the child does not perform the action in front of the researcher the survey will rely on the response of the primary caregiver.
- **Code C**. The child shall perform an action and the researcher makes assessment solely on the basis of the observed response.

[2] Each pair of surveyors was given a closed bag containing two small wooden cubes, eight large wooden cubes, a puzzle (the same one for everyone), a piece of cloth with a button and a zipper, a booklet with plastic coated images, cereal wheels, a small plastic container, a 500ml plastic container, a pair of scissors (child safe), a cup (3.5 inch Peruvian), a cloth doll, a small ball (size 2 sponge), a medium ball (size 10, plastic), a mirror (plastic), six crayons, one shoelace, five perforated bottle caps, a toy car (plastic), a spoon (plastic), a rattle, a Ziploc bag (17x20), a Ziploc bag (27x28), a bag of toys (raffia), a watercolor marker , a storybook, 40 blank sheets of paper, a container of Cloralex handy wipes (to be used to clean the materials after each interview).

Also, field researchers are directed to obtain the exact age of the children, including months and days, as well as make adjustments to age if the child is born prematurely (three weeks or more before the expected date provided and only when the child is less than two years of age when administering the instrument, according to the protocol of the instrument).

When addressing the mother, the protocol instructions are based on respect for the primary caregiver regardless of belief, origin, customs and habits in relation to minor.
An example of the dialogue used to introduce the ASQ instrument to caregivers is as follows:

Instructions for the mother prior to beginning assessment:

1. "I'm going to ask some questions about certain activities generally performed by children. It may be that your child is able to perform some of these activities and others he/she cannot yet do"
2. " We shall carry out some of these activities"
3. "There are 3 ways to respond: YES, SOMETIMES AND NOT YET.
   Respond **YES**, if the child usually performs the activity.
   SOMETIMES, if the activity is performed occasionally, but not all the time, and
   NOT YET, if the child cannot perform the activity"

After reviewing the main instruments of the assessment, the general assessment instruments and implementation protocol shall be presented.

3.2 Context Instruments and General Implementation Protocol Training

Protocol training on the general administration was the responsibility of the CIDE research team members who have experience in the design of quantitative data collection instruments, design of interview protocols and field work. In the introduction to the project it was explained that the program's target population is generally composed of residents of rural and indigenous communities with high and very high levels of marginalization. Also mentioned were the types of sessions included in the program:

- Best parenting practices for caregivers
- Early stimulation for children under four (prior to preschool)
- Sessions on prenatal and initial child care for pregnant women
- Sessions for fathers

Also explained is the number of data collection waves that are to be undertaken during the entire assessment which correspond to the number of times the same household would be

visited. Likewise, regarding the structure used to implement the program, explained was the fact that the sessions are provided on a weekly basis by a promoter who usually resides in the community and in turn, as part of the CONAFE work structure, that CONAFE area supervisors are responsible for the work of up to ten promoters. During training the *impact assessment* and *random experiment* concepts were covered, along with the purpose of the assessment and, in this particular case, the observation of the effects of the CONAFE Early Childhood Education program.

On the data collection instruments, it was mentioned that the assessment of the Early Childhood Education program considers the administration of up to four general instruments and ASQ, so it was necessary to understand the content and the sequence in which they should be applied, both in their group participation and comparison formats. The number of instruments would depend on the number of members in a single participating household or a household that would attend the CONAFE Early Childhood Education sessions, with up to a maximum of three children per home.

Every questionnaire prepared by CIDE includes open – complementary – and closed questions – main measures; questions to record each household visited, number of visits made and implementation schedule were also included. The recording of these data must always be captured as an observation mechanism within the scope of the field researcher's work. As such, the job of the CIDE supervisor shall be to answer questions in the field, and offer general support as a liaison between CONAFE, local authorities and field research staff.

Protocol for field research ethics was considered an important part of the issues to be addressed, especially maintaining respect for the diversity of customs of the people and communities. Similarly, issues such as the interaction there would be between the CIDE team, field researchers and CONAFE staff and commitments they have each to the successful completion of data collection were discussed. Emphasis was placed on the treatment that must be used when dealing with children, language to be used when addressing families and peers, punctuality and clothing required to work in the field in order to facilitate timely and correct work.

3.3 Field Practices and Personnel Selection

For field practices, CONAFE authorities in Hidalgo managed the *Soluciones Q* researchers' visits to family homes currently participating in the Early Childhood Education program but not comprising part of the sample. Therefore, accompanied and guided by CIDE staff and INP psychologists, the researchers had the opportunity to become familiar with the

environment where some program beneficiary families live, understand the logical order of quetionnaire administration, and experience some of the situations that arise when visiting a new home, for example: the family has grown in size compared to the previous visit, the family has been separated, the family has moved away, etc.

Training at CIDE, as previously mentioned, allowed the team of CIDE trainers and supervisors to get to know the potential researchers with whom they will work in the field. Practical work was useful in making the selection of profiles required to work with the study's target population. Adherence to collection protocols, study tools and teamwork, were some of the important criteria that were considered for the final selection of field research staff for this second data collection. All research staff in the field and their distribution by state and institution are shown in the table below.

Table 2: Information gathering personnel

Entity	Surveyors	Company Supervisors	CIDE Supervisors	Total
Chiapas	10	1	2	**13**
State of Mexico	16	1	2	**19**
Oaxaca	4	1	2	**7**
Puebla	12	2	1	**15**
Veracruz	16	1	1	**18**
Queretaro	12	3	4	**19**

4 Indigenous Languages

During the baseline survey, some skills difficulties were experienced with the interpreters hired. Specifically, while some of the interpreters spoke general language, they could not deal with the community variant. Consequently, in the follow up collection, it was decided to request the service directly from the local authorities or by recommendation from the CONAFE promoters and supervisors to ensure the accuracy of the interpretation. In most cases, this meant residents and authorities of the community, or certified interpreters with prior experience in the administration of questionnaires.

All cases where an interpreter was required were in the states of Oaxaca and Chiapas. According to administration protocol, the field researcher always addressed the respondent, even if interpreter gave the message, likewise, when there was express consent by the respondent, the questionnaire was recorded on audio and video to further validate the quality of interpretation and responses. The table below lists the languages and variants for which the support of an interpreter was required during interviews.

Table 3: Indigenous languages and variations by state

Chiapas	Oaxaca
Tzotzil	North-Eastern Mixteco
Tseltal • Eastern • Northern • Southern	Mazateco • From the Presa • South-Eastern • From Eloxochitlan
Chol	Mountain Zapoteco

5 Survey

The survey began in Chiapas, Mexico State, Oaxaca, Puebla and Veracruz this year in the third week of July. In Queretaro, by formal request of CONAFE state authorities, data collection began in August. For this survey, there were details of all households that participated last year, so it was easier to locate homes, define the reasons for no longer attending sessions or know who were out of town, whether temporarily or permanently.

At all times field researchers were kept in communication with CIDE field monitors, and these in turn with the CIDE staff in Mexico City, in order to report any incidents and respond expeditiously to changes in the environment.[3]

6 Supervision

During the course of gathering information, CIDE sent eight field supervisors to accompany the work of the contractor, to observe the execution of the interviews, answer questions arising during implementation and to identify and correct any errors observed. Supervisors sent by the CIDE were a great help to the extent that they also supported the resolution of questions by community authorities, CONAFE staff and field supervisors of the polling firm. The CIDE field staff reported incidents daily to staff CIDE offices located in Mexico City.

7 Response Rate

Once both the field survey and capture and validation processes were completed, the results showed a total of 1,061 completed interviews spread over six states, with a frequency

[3] The work Schedule can be referred to in appendix I.

that generally resulted in a response rate of 75.4% relative to the baseline survey of sample households.[4] Mostly, the rate of non-response cases was due to the absence of the family at the time of the survey; this situation occurred in 15.1% of registered households. This absence is now explained by several factors among which are the following: arrive from work at time not feasible for the implementation of interviews, were temporarily out of town (i.e., on vacation) or had moved away permanently. In 9.5% of the remaining cases of non-response, the rate was as follows: 1.1 % of the interviews could not be completed, 1.5 % of households refused to respond to the interview and 6.9 % registered a number of distinct cases (Tables 4 and 5).

Table 4: Homes and communities by state

State	Homes		Communities	
	Frequency	%	Frequency	%
Chiapas	139	13.1%	18	14.3%
State of Mexico	366	34.5%	36	28.6%
Oaxaca	88	8.3%	10	7.9%
Puebla	79	7.4%	10	7.9%
Queretaro	174	16.4%	26	20.6%
Veracruz	215	20.3%	26	20.6%
Total	1061	100.0%	126	100.0%

In an experimental assessment design such as this, it is important to consider the proportion of the response rate by type of group where the primary units are communities, as in this case. Thus, we can see that the participating communities achieved a lower response rate than control: while the first recorded 71.3% of households in the latter there were about 80%. According to the overall results, we found that in both types of communities the absence of people in the household was the main factor for the non-response rate, 18.4% for participation groups and 11.7% for comparison groups.

[4] The results of this table (table 4) must be taken with caution given that although the field teams visited 126 communities, it does not mean it was possible to carry out interviews in all of them.

Table 5: Interview Results

Interview Results	Frequency	%
Complete interview	1061	75.4
Incomplete or interrupted interview	16	1.1
Participation denied	21	1.5
Person absent	213	15.1
Other	97	6.9
Total	1408	100.0

By order of frequency, the previous factor is followed with 6.8% and 7% for participation and comparison respectively. In both types of communities a similar percentage of households refused to answer 1.5% for participation and 1.4% for comparison, however, one must bear in mind that these percentages are relative because they represent a ratio at different frequencies. Finally, in the control communities, only 0.3% of households interrupted or did not completed the interview while 1.9% of participating households experienced the same situation (see Table 6).

Table 6: Interview Results by Community Type

Interview Results	Type of Community	
	Participation	Control
Complete interview	71.3%	79.6%
Incomplete or interrupted interview	1.9%	0.3%
Participation denied	1.5%	1.4%
Person absent	18.4%	11.7%
Other	6.8%	7.0%
Total	100.0%	100.0%

With respect to the states where the communities are found, the results indicate that Oaxaca and Veracruz are the states with the greatest participation in the interviews, showing 87.1% and 83% respectively where households completed the interviews. Chiapas reached 79.9% of completed interviews, the State of Mexico 70.8%, Puebla 76.7% and finally Queretaro, the state with the lowest participation rate at 68.5% of completed interviews. Absenteeism was the greatest factor affecting the response rate, as expected from the general results: in Queretaro this factor reached 21.3%, 17.4% in the State of Mexico, 15.5% in Chiapas and 11.7% in Puebla; to a lesser extent in Oaxaca with 7.9% and Veracruz 8.5%. There were other significant factors in Queretaro (9.4%) and Puebla (8.7%) and the State of Mexico (7.9%), while in Veracruz, Oaxaca and Chiapas oscillated at a lower range (between 4% and 4.6%).

Table 7: Interview Results by State

Interview Results	State					
	Chiapas	State of Mexico	Oaxaca	Puebla	Queretaro	Veracruz
Complete interview	79.9%	70.8%	87.1%	76.7%	68.5%	83.0%
Incomplete or interrupted interview	0.6%	2.3%	1.0%	0.0%	0.0%	0.8%
Participation denied	0.0%	1.5%	0.0%	2.9%	0.8%	3.1%
Person absent	15.5%	17.4%	7.9%	11.7%	21.3%	8.5%
Other	4%	7.9%	4%	8.7%	9.4%	4.6%
Total	100%	100%	100%	100%	100%	100%

Similarly, it should be added qualitatively that there were difficulties encountered in some communities, as it was impossible to collect data in several homes or even at all. In the State of Mexico, a community participation group received a program similar to the CONAFE Early Childhood Education program implemented by the local state government. Since it is difficult to isolate the effects of both programs, it was decided to treat this community differently as, in terms of the experiment, the community is considered contaminated. Nonetheless, due to the protocol designed for the survey, the field research team sought to implement field interviews but the state program personnel isolated the families, thereby preventing the research team from interviewing them. Thus, all households in this community were registered in the non-response rate.

In the state of Chiapas, meanwhile, two communities were visited that presented other types of problems during data collection. In a comparison community, members of all households made the collective decision not to respond to follow-up interviews, despite having participated during the baseline survey. After making the decision they asked the research team to withdraw from the community. This situation also led to the entire community being registered in the non-response rate. In a participation community, the promoter reported that families stopped attending the sessions. To corroborate the information the research team in charge conducted a follow-up interview with one of families where it was reported that it was the promoter who stopped holding the sessions.

A comparison community in the state of Veracruz suffered a landslide from a nearby hillside as a result of the rainy season, and though it destroyed several homes, fortunately no casualties were reported. Despite this incident, it was possible to carry out most of the interviews. Consequently, at the end of data collection the follow up response rate remained at 123 communities which registered just over a thousand households in the original sample.

It is clear that the reasons why some households did not participate in the survey and comprise part of the non-response rate differ in their source. This situation is not a minor issue. While there were communities, both participation and comparison, that decided not to continue participating in the assessment, other communities simply chose not to participate in the Early Childhood Education program and therefore -but not necessarily- also left the assessment. Something worth adding with respect to the consequences of these situations, while households in communities that chose not to participate in the assessment cannot be interviewed in the third survey, given the research protocol, but those communities where households just chose not to participate in the CONAFE Early Childhood Education program there will be an attempt made to include them in the third survey.

Finally, it is important to note that information was collected from 156 new homes. This, on the one hand, was in order to capture data from homes where the family nucleus had changed and therefore became new homes, an occurrence in both types of communities. Or, on the other hand, households that were added to the Early Childhood Education program in participation communities only. These households were distributed as shown in the table below. It is noted that in the State of Mexico the highest percentage of new households were registered, 33.3%, contrary to Oaxaca which is infrequently in this situation and whose proportion reached 3.8%. While in the states of Queretaro and Veracruz the same percentage, 19.2%, of new households were registered. In the state of Chiapas, 13.5% were reached.

Table 8: First Interview Homes by State

State	Homes	%
Chiapas	21	13.5
State of Mexico	52	33.3
Oaxaca	6	3.8
Puebla	17	10.9
Queretaro	30	19.2
Veracruz	30	19.2
Total	156	100.0

8. Capture and Validation

This time capture and data validation processes were performed exclusively by CIDE with staff hired and properly trained for the occasion. From the end of the survey and during the rest of August and early September, the questionnaires were recorded using a capture mask programmed by CIDE research team staff. The mask was designed expressly for reducing capture errors and making the registration validation process more efficient. Once this process was completed, a review of household and community identifiers, and the most important variables was performed to avoid any errors made during the field survey. Finally, cross-validation was done on 30% of the total questionnaires without finding frequent or systematic errors, thus demonstrating that the whole of the data is reliable.

9. Results

In subsections 9.1 to 9.4 a comparison of the results of the baseline and follow-up to identify possible changes in the context of children is presented. The results shown in these sections contain all households that responded to both the baseline and follow up waves. In the final subsection (9.5) preliminary impact results on early childhood and caregivers measures are exposed. These results derived only from observations that were matched between both waves.

9.1 Home

With respect to the context of the first questionnaire, it is important to remember that in this study the household is considered to be all persons living in the same home and sharing the same food expenses. Thus, a household may contain the nuclear family, extended family and others who meet these two requirements. Around 93% of people who responded to the follow-up questionnaire also did the baseline survey; in most cases, it was the wife of the householder who answered the questions.

Table 9: Is the person being interviewed the same person that was interviewed in the previous survey done by CIDE?

	Percentage
Yes	93.32
No	6.68
Total	100

The results of the baseline survey showed that households had an average of five members, 80% of them were less than three and a half years old, 17% were two, and only 1% three or older. Currently, the follow up data show an average of 3.5 members in the household, 62%

less than five years of age, 32% have two children and 5% have three or more children at preschool age.

Table 10: Children under 60 months in the home

	Percentage
0	0.2
1	62.38
2	32.08
3	5.35
Total	100

Among household members over 5 years of age, we found that almost 95% of them can speak Spanish, while 5% do not or cannot speak it. As a complement, household members older than 10 years were asked if they could read and write a message, either in Spanish or an Indigenous language. Eighty-five percent of household members could do it in Spanish, just under 2% were able to do so in an indigenous language only, while 7% could do it both in Spanish and an indigenous language.

Table 11: Does [NAME] know how to read and write a message? Members 10 years of age or over.

	Percentage
Yes, only in Spanish	85.56
Yes, only in an indigenous language	1.81
Yes, in Spanish and an indigenous language	7.37
No	4.73
Cannot read / write	0.54
Total	100

On the level of education in the household, baseline results showed that just over 80% of those members over 17 had basic education, while 6% do not have any formal education. In preparatory education there were just 11% of the same population and only 2.4% of that reported having attained a bachelor's degree or higher. Follow up results remain almost the same, except in higher education, in which the ratio drops to 1.58%, which in part can be explained by the mobility of household members.

Table 12: Highest level of education completed by members of the household 17 years of age or over.

	Percentage
No formal education	6.17
Preschool	0.94
Primary	42.86
Junior High	37.87
Senior High	10.58
Bachelor	1.54
Postgraduate	0.04
Total	100

The housing conditions did not change much between the baseline and follow up surveys: having an average of three rooms, including the kitchen, and 68% of them had one to three bedrooms. Of these, the average number of rooms used for sleeping is 1.78, and over 85% of households have one or two rooms for this purpose. Flooring materials also have no significant changes: about 9% of households are made of dirt, 80% responded that the main material is cement and the remaining 10% have floors made of wood, tile or other material; baseline percentages were 10%, 80% and 10% respectively.

Table 13: What material is the greatest portion of the floor made of in this home?

	Percentage
Dirt	8.85
Cement or asphalt	80.32
Other flooring (Wood, tile or other)	10.83
Total	100

As for the material of the walls some minor changes were as follows: approximately 75% of households have brick, stone or limestone walls, both in the baseline and follow-up questionnaire. Wood walls represented 14% of total households in baseline, while the follow up percentage decreased to 12.5%. Some 7% of households reported having adobe walls in the baseline survey, while households in the follow up the ratio approximately reached 9%. The remaining materials, including sheet or scrap materials represented 3% in the first survey and 2% in the second.

Table 14: What material is the greatest portion of the walls made of in this home?

	Percentage
Scrap material	0.1
Pegboard	1.29
Asbestos or metal sheeting	0.3
Reeds, bamboo or mud	0.5
Wood	12.55
Adobe	8.67
Brick	76.59
Total	100

Roofing materials also have very similar ratios between the first and second surveys: half of the homes have concrete slab, 32% have sheet metal and 9% have asbestos sheeting. The remaining 7% are distributed among homes with pegboard, metal, tile and wood.

Table 15: What material is the greatest portion of the roof made of in this home?

	Percentage
Pegboard	3.98
Sheet metal	32.17
Asbestos sheeting	8.96
Wood	0.7
Tile	2.39
Cement	51.79
Total	100

Again, the distribution by household is very similar regarding access to electric power. The baseline results indicate that 3.4% of households were without power any day of the week, compared with 3.7% of households in the follow up. One point nine percent of households in the first survey reported having service for four days or less, while in the second survey the percentage increases to 3.3% of households. Finally, the first survey found that almost 92% of households had power all day, but for homes in the follow up survey the percentage drops to 89%.

Table 16: During the past week, how many days were there with power for at least 3 hours?

	Percentage
0	3.77
1	0.3
2	0.7
3	1.09
4	1.19
5	1.29
6	2.68
7	88.98
Total	100

Contrary to what has been seen so far, there are significant differences in how household water is supplied between the surveys. The baseline results indicated that just over two thirds of households had access to water in their home or on their land, a quarter resorted to a natural body of water and two out of twenty households hauled water or were supplied by tanker truck. However, follow up data indicates that almost three quarters of households have a water source on their property, a fifth draws it from a river or lake and only three of every fifty households obtain water from a source off their property or by tanker truck.

Table 17: How does water arrive to this house?

	Percentage
Plumbing within the house	34.7
Plumbing out of doors, but on the property	39.38
Plumbing form public valve	2.89
Plumbing from another house	1.6
Tanker	1.4
Well, river, lake, ravine or other	19.94
Total	100

Waste management shows similarities between surveys, since in both just over 90% of households claimed they had some final destination of waste such as a toilet, latrine or out house. As for the specifics, about one third of households are connected to the public sewage system, another third has a septic tank and just under a third have no sewage system. The rest (8.3%) have pipes connected to an open space such as a ravine or a body of water (river, lake or sea).

Table 18: Does this house have a drainage or sewer system connected to:

	Percentage
Public Sewage	32.3
Septic Tank	32.3
Pipes to a cliff	5.52
Pipes to a river, lake or sea	2.81
No Sewage System	27.08
Total	100

Possession of goods changes between baseline and follow-up survey results, although not substantially since this only occurred in certain cases. For example, possession of cell phones increased from 50% to 53%, while the landline phone went from 10% to 8%. The rest of the property, such as possession of appliances and computer equipment showed no significant differences. The only case where a major difference is observed is the possession of a wood or coal stove: Baseline survey results suggest that 59% of households had a stove of this type, while in the follow up the percentage was raised to 69%. The complete results of the second survey are listed in the table below.

Table 19: Possession of goods and services in the home.

	Percentage of "Yes"
Gas Stove	67.27
Wood Stove	69.58
Water Tank	49.95
Boiler	17.13
Cistern	15.74
Shower	22.99
Power Meter	78.93
Automobile	23.76
Internet	1.69
Computer	3.68
Telephone	8.36
Cellular Phone	53.23
Washer	33.1
Refrigerator	51.84
Television Set	82.87
Radio	58.31

9.2 Caregivers

The primary caregiver of a child under 60 months is a person who spends the most time with him/her and therefore contributes significantly to their upbringing; this may be a parent, sibling, grandparent or another person that may not even be a member of the household. In this regard, it is important to mention that there may be more than one caregiver and hence the interview administration protocol is established so that more than one questionnaire may be applied when there are two caregivers in a given household.

Regarding the questionnaire for caregivers, it is important to mention that the design was adapted given the natural changes in the age ranges of children. Parenting practices have different dimensions depending on two of the three age ranges defined in the Home test: 0 to 35 months and 36 to 71 months of age. In the baseline questionnaire, sections for child caregivers ranged from 0 to 35 months and from 36 to 42 months of age. Currently, the follow-up questionnaire, to indicate the second age range, had to be modified from the limit of between 36 and 60 months of age to be able to capture the children that fall within the 43 to 60 months of age range. While the change is minor, its relevance lies in the impact it could have on the misreading of the responses by the interviewers and consequently in its administration; in fact, this observation was derived from the pilot tests.

Regarding the number of caregivers included in the results, it is important to note that to build the base of caregivers it was decided to give everyone who responded to the questionnaire a second interview, because they were the same ones who were registered in the baseline questionnaire along with caregivers belonging to the new homes in the second interview. An example that can better illustrate such cases is when the primary caregiver from a home in the baseline survey was registered as the child's mother. However, in the second panel run another primary caregiver also attended the program sessions was registered, as the child's grandmother would be, as she in turn happens to spend considerable time with the child while the mother goes to work. Thus, the child receives care from two people although they may or may not be a member of the household but are the people who actually attend the program sessions and are of particular interest for the analysis.

In this second survey the following general characteristics of caregivers were recorded. The primary caregiver was the same person interviewed in the baseline survey in 88 % of cases, while in the remaining 12% it was a new person. As for the relationship with the child, the results are very similar to those of the baseline: just over 90% were the child's mother; in 5% of baseline cases it was a grandparent, while the second survey showed 6%. In the

remaining 4% it was the child's father, brother, uncle and others. On average, the primary caregiver had lived in the community where he/she was interviewed for 19 years. Moreover, in 2012, 73% of caregivers said that there was another person who helped them care for the child, while the follow up survey indicates that the percentage increased to 76%. Regarding the composition of the household, with respect to the caregiver, in the first survey 85% of caregivers responded that the father and mother of the child were living in the same household, while in the second survey this increased to 86%.

Table 20: What relationship do you have with [NAME(s) of child(ren)]?

	Percentage
Mother	90.66
Father	0.87
Brother	0.52
Grandparent	6.38
Other	1.58
Total	100

Regarding government support, on the one hand, we found that in the first survey 45% of caregivers reported that someone in the household was a beneficiary of *Oportunidades*, while in the follow-up questionnaire this percentage increased to just over 50%. As for the *Seguro Popular* affiliation, the percentage increased from 59% in the first year of the study to 70% in the survey this year. Therefore, it is noted that a couple of years ago the National Commission for Social Health Protection (CNPSS) proposed a universal policy for Social Health Protection (SPSS), particularly through its *Seguro Popular*. A major recruitment drive was derived from it and now total *Seguro Popular* members reached almost 54 million individuals (46% of the total population), of which only 1% pay any fees.[5] The *Procampo* program, meanwhile, had coverage very similar to last year's; it went from 6% to 5.8 %.

On the other hand, if we review the contributions to the home by friends or family, baseline results indicate that 6% of caregivers reported receiving some support from people living within the country, and in 3% of cases received support from someone living abroad. Follow up data show a slight change, as 4.8% of caregivers reported some support from people living within the country, and 4.6% from people living abroad.

[5] National Commission for Social Protection in Health, 2013. "System of Social Protection in Health. January-June Report, 2013". Page 10.

Table 21: Is someone in the home of [NAME] currently a beneficiary of the Oportunidades Program?

	Percentage
No	49.78
Yes	50.22
Total	100

As for the educational resources available for child rearing, the baseline survey showed that 68% of children 0 to 35 months do not have children's books, 15% had one or two books and 16% had three or more. In the group of 36 to 42 months of age, 46% did not have children's books, 29% had one or two and 24% had three or more. In the follow up study data is shown in the table below; we found that among children 0 to 35 months of age, the percentage of children with no children's books has dropped to 25%, 23% have one or two books and just over 50% have three or more. In the 36 to 60 months of age group, only 6% did not have any children's book, 28% have one or two and almost two thirds (65%) have three or more.

Table 22: Approximately how many children's books do you have [NAME]?

	0 to 35 months	36 to 60 months
	Percentage	Percentage
Ten or more	30.91	38.1
Three to Nine	19.94	27.02
One or Two	23.34	28.23
None	25.81	6.65
Total	100	100

Information on the use of educational material was also obtained, and the frequency with which the caregiver reads or tells stories to the child. The baseline survey data indicate that nearly half of the caregivers of children 0 to 35 months never did it, 12% did so about once a month, 16% do it once a week, 11% said they read to the child three times a week and about 12% did so every day. In the of 36 to 42 months group, a third of caregivers never read to the child, 22% did so once a month or less, 23% did it once a week, 11.5% three times a week and nearly 10% read stories daily. In the second survey, it was found that 32% of caregivers of children 0 to 35 months never read or told stories, 18% do so once a month or less, 20% read to the child once a week, 15% do it three times a week and almost 13% every day. In contrast, in the 36 to 60 months of age group the follow up survey found that nearly 21% never read, 25% do so once a month or less, another 25% do so once a week, 17 % have story time three times a week and 11% every day.

Table 23: How often are you able to read or tell stories to [NAME]?

	0 to 35 months	36 to 60 months
	Percentage	Percentage
Never	32.61	20.93
Several times per year	7.11	8.85
Several times per month	11.28	16.7
Once per week	20.25	25.35
Three times per week	15.77	17.1
Every day	12.98	11.07
Total	100	100

With respect to the attention the caregiver must provide the child up to 35 months of age, about 90% of those interviewed during the baseline survey responded that it is best to spend time teaching children, either always or most of the time, and only 10% considered the best way for child development is that they learn on their own. However, the results of second survey indicate that 93% of parents prefer to take time to teach their children, while the percentage of those who prefer their children learn on their own fell below 7%.

Table 24: Some parents dedicate time to teaching their children, while other parents think that children learn better on their own. Which of the following statements are nearest to your thoughts?

	Percentage
Parents should always dedicate time to teaching their children	65.07
Parents should always leave their children to learn on their own	4.48
Parents should dedicate time to teaching their children in most cases	28.59
Parents should leave their children to learn on their own in most cases	1.85
Total	100

The use of corporal punishment on children presented some changes. In the baseline questionnaire, when caregivers of children 0 to 35 months were asked about how often they had spanked the child in the past week, 65% said they had not done it, 17% said they had done it once, and 18% said they had had to on two or more occasions. In the group of 36 to 42 months of age, 48% of caregivers said they had not given the child a spanking, 20% had done so only once, and 32% had done so on two or more occasions during the week prior to the interview. In the second survey, in contrast, with respect to caregivers of children 0 to 35

months, 57.5% said they had not resorted to corporal punishment, 20% had done so only once, and the remaining 22% had done so on two or more occasions. Among caregivers of children 36 to 60 months, a reduction in the frequency of acts of corporal punishment is also registered, as almost 52% of respondents had not spanked the child, 24% did so only once, and 24% used physical punishment on two or more occasions.

Table 25: Children are sometimes well behaved and other times they are not. During the past week, about how many times, if at all, did you have to spank [NAME]?

	0 to 35 months	36 to 60 months
	Percentage	Percentage
0	57.56	51.92
1	19.75	24.04
2	14.2	13.13
3 or more	8.48	10.89
Total	100	100

With regard to parenting practices, only small changes occurred with respect to the baseline results. For example, in the first survey, 77% of respondents stated they sing with the child, while in the second survey the percentage rose to nearly 82%. Also, when asked if the child had any musical instrument, whether bought or handmade, 36% of baseline respondents said the child had at least one, while in the follow-up questionnaire this increased to 40%. In the rest of parenting practices, both the caregiver and the father or father figure, the percentages were virtually the same.

Table 26: Now I am going to ask you to please tell me if the following activities or situations occur:

	Percentage
You take [NAME] for a medical checkup even when not sick	58.6
You play with [NAME]	95.46
You sing with [NAME]	81.99
[NAME] has a musical instrument	40.72
The father of [NAME] helps prepare meals for [NAME]	43.58
The father of [NAME] helps dress [NAME]	67.92
The father of [NAME] plays with [NAME]	87.77
The father of [NAME] carries [NAME]	88.44

Finally, on expectations of the child's academic achievement, slightly higher results are presented in the second survey. In the baseline survey, 30% of caregivers expected the child to achieve basic education, 30% thought the child would finish senior high school and 40% felt that the child would achieve a bachelor's degree or post graduate studies. In contrast, 26% of respondents in the second survey expected the child to complete basic education, one-third expected junior high school education and 41% believed the child would attend university.

Table 27: Up to which academic level do you believe [NAME] will study?

	Percentage
Primary	3.73
Junior High	22.2
Senior High	32.76
Bachelor	39.13
Postgraduate	2.18
Total	100

9.3 Pregnant Women

The follow-up questionnaire for pregnant women had to be changed significantly compared with the baseline survey. Naturally, pregnancy at any stage of gestation in the first survey had ended before starting the follow up survey; however, it was possible that the mother became pregnant again. Thus, it was necessary to capture the care given during pregnancy, early parenting practices and the possibility of a new pregnancy. The composition of

households did not change from baseline, as 92% had no members pregnant, while the remaining 8% had one or more. For the second survey, they had spent an average of ten months after pregnancy reported in baseline, and 13% of respondents claimed to be pregnant again, with an average of 4.2 months of gestation. About 28% of pregnant women mentioned a person who helps take care of the newborn child, while 72% have no help.

Table 28: Number of pregnant women in the home

	Percentage
0	91.78
1	7.92
2	0.3
Total	100

Concerning the childcare during pregnancy, the results show slight changes from the baseline data. In the first survey, 4% of respondents said they had been unable to attend a medical appointment, while in the follow-up questionnaire the percentage dropped to 3.8. Among women who reported receiving medical attention in the baseline survey, the average was 4.5% throughout gestation, and a year later, an average of 6.8% respondents received medical attention. It was also asked if the child's father had accompanied her: in the baseline survey, 65% reported that the father had accompanied her to at least one medical visit, while in the follow up data the proportion nearly reached 70%.

Table 29: During the pregnancy, did the father of the child accompany her to a doctor's appointment?

	Percentage
Yes	69.81
No	30.19
Total	100

When reviewing the mother's care during pregnancy, we found that almost 8% had taken some medication that was not prescribed by a doctor, while 92% had not. It was also asked if the pregnant woman had suffered an injury or illness during pregnancy, to which nearly 20% reported an incident of this type and the remaining 80% had experienced no major complications. The table below shows where the mother took some kind of supplement: 84% of mothers said they had taken folic acid and iron during pregnancy, and only 3.7% of them had not taken either.

Table 30: During the pregnancy, was folic acid or iron taken?

	Percentage
Yes, both	84.91
Yes, only iron	3.77
Yes, only folic acid	7.55
Neither	3.77
Total	100

As for their diet during pregnancy, in the baseline survey some 73% of mothers said that their food intake did not change as a result of pregnancy, while in the follow-up questionnaire the percentage dropped to 70%. If we look at the quantity of food in the first survey 23% said they had eaten less during pregnancy, 25% had eaten the same amount, and just over half had eaten more. Conversely, the results from the second survey showed that 22% of mothers had eaten less from the beginning of the pregnancy, 48% had eaten the same quantity and 30% had eaten more.

As for increased risk consumption, the baseline survey established that 97% of mothers had not consumed alcohol during the week and just over 2% had done so at least once, on the other hand, tobacco products had been consumed by only 1% of pregnant women interviewed. Stimulant use was more common, as almost 60% had consumed coffee at least once in the week prior to the interview and 72% had drunk soda. However, the results of follow up survey indicate that 1.9% of mothers to have consumed alcohol in the week prior to the interview, none had used tobacco, 65% had consumed coffee and 82% of them had drunk soda at least once.

Finally, questions were asked about the mother's early child-rearing practices. Almost all respondents in the second survey (96%) believe that babies are aware of what happens around them. When asked if they had considered enrolling their child in preschool, 94% said yes they would, and almost all (90%) said their child would be enrolled at three or four years of age. As for their expectations of academic achievement, 36% believed that their child would achieve some level of basic education, almost 39% thought their child would attend junior high school and 24% believed their child would obtain a bachelor's degree or complete post graduate studies.

Table 31: Up to which academic level do you believe your son/daughter will study?

	Percentage
Primary	8.16
Junior High	28.57
Senior High	38.78
Bachelor	22.45
Postgraduate	2.04
Total	100

9.4 Parents

The model on which the intervention program rests includes fathers as a key participants in the development of children. The sessions provided to parents are meant to make an impact on care and protection, as well as explore parental knowledge on the matter. Therefore, through the questionnaire, parents were asked about parenting practices and social attitudes in the home.

Concerning the general characteristics of the parents, we found the average length of residence in the community to be 27 years, 4 more than that reported in the baseline questionnaire. This difference can be explained by parents who were unable to be found in the second survey having a lower average length of residence in the community compared against the data in the baseline survey (23.03 years); for those who responded in both cases (26.07 years).

In the baseline questionnaire, 84% of parents reported that fathers eat with the child at least once a day, 10% a few times a week, and 5% said they eat with the child once a week or less. However, in the follow-up questionnaire it was found that fathers eat with their children at least once a day in 87% of the cases, 5.4% a few times a week and 7.2% only once a week or less.

Table 32: How often does [NAME] eat with you and his/her mother?

	Percentage
More than once per day	67.27
Once per day	20
Several times per week	5.45
Approximately once per week	5.45
Approximately once or twice per month	1.82
Total	100

As for parenting practices, the first survey indicates that 63.6% of parents take the child in for a checkup even when not ill, while in the second survey the percentage was reduced by one percentage point. The percentage of parents who sang with children increased from 82.4% in the first survey to 94.6% in the second. Affirmative responses of parents that help prepare meals for the child also increased, as they went from 72% in the first survey to 78.5% in the second. As a complement, where previously 89% of parents helped feed the child, in this year's survey responses increased to almost 93%. The frequency with which parents help dress or change the child also increased, as in the baseline survey the proportion was 81% and in the follow-up questionnaire it was 87.5%. Finally, in the first survey 96% of parents reported carrying the child, while in the second frequency was reduced to 92.8%.

Table 33: Now I am going to ask you to please tell me if the following activities or situations currently occur:

	Percentage
Take [NAME] in for a medical checkup even when not sick	62.5
Play with [NAME]	98.21
Sing with [NAME]	94.64
Help preparing meals for [NAME]	78.57
Help feeding [NAME]	92.86
Help dressing [NAME]	87.5
Carry [NAME]	92.86

In regard to social behavior, in the baseline survey 11% of parents surveyed reported smoking either inside or outside of the home, while in the second survey the percentage increased to 16%. As for the economic solvency in the home, the first year 46% of respondents reported having financial problems, while the percentage increased to 55% in the second. The frequency with which the father talked to someone when having problems, feeling sad or angry, remained practically the same.

Table 34: Social Conduct

	Percentage
Smokes	16.07
Smokes in the home	9.09
Decides on household spending without seeking advice	42.86
Has money problems	55.36
Discusses solutions to problems with family	94.64
Speaks to someone when feeling sad	85.71
Speaks to someone when feeling angry	58.93

Finally, the expectations that parents have for the academic achievement of their children show significant changes between the two surveys. In the first survey, 23% of parents expected their child to obtain basic education, compared to 19% in the second. The most significant change was in the expectation of junior high school, as only 20% of parents shared the expectation their child would reach this level in the first year this increased to 48% in the second year. Finally, where 57% of parents in the baseline survey expected their child to attend university, that percentage dropped to 32% in the follow-up survey.

Table 35: Up to which academic level do you believe [NAME] will study?

	Percentage
Primary	3.57
Junior High	16.07
Senior High	48.21
Bachelor	32.14
Total	100

9.5 Preliminary Impact Results: ASQ and Caregivers

We present the preliminary results of the first follow up data collection by child development outcomes, and impacts on caregivers. For child development measures, while the results on the five outcome variables of interest all suggest positive impacts, only the impact on gross motor skill development is significantly significant. For caregivers, we find no pattern of statistically significant impacts in the data self-reported caregiver behaviors. However, four of seven measures of caregiving that were directly observed by enumerators show positive and significant impact on caregiver behavior.

9.5.1 Impacts on Child Development

Using the Ages and Stages Questionnaire (ASQ-3), the study measures the effects on child development by examining five areas of child development, each measured by a normalized index: i) social skills, ii) problem-solving, iii) fine motor, iv) gross motor, and v) communication. As shown in Table 36 and summarized in Figure 1, all of the preliminary effects on child development are positive, however, the only positive and significant effect is on gross motor skills. Our hypothesis, to be tested through the upcoming direct observation of training sessions, is that this reflects a combination of the strengths of the curriculum and potential bias in the delivery of the curriculum. For the first, it may be that communication, gross motor skills, and fine motor skills develop more immediately than problem solving skills.

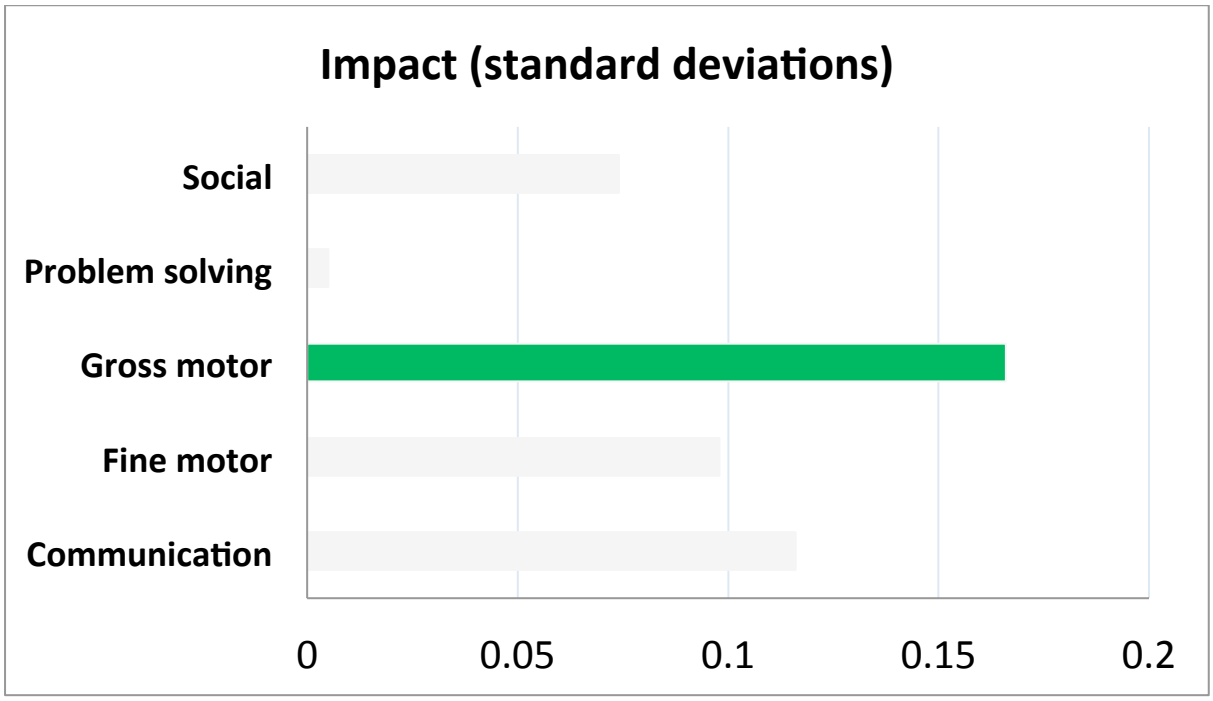
Table 36: Preliminary Impacts on Child Development

	Communication	Fine Motor	Gross Motor	Resolving problems	Social
Treatment * After	0.12	0.10	0.17**	0.01	0.07
	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)
After	0.09*	-0.22***	0.18***	-0.04	0.08
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Constant	-0.06***	0.08***	-0.11***	0.01	-0.05**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
HH Fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	1,357	1,357	1,357	1,357	1,357

Notes: Scores for each area have been converted into z-scores with mean zero and standard deviation one.

For the second, trainers may lead with activities that favor gross motor skill development (e.g., playing with balls), since these activities tend to be easier. Leading with the gross motor skills activities can crowd out other stimulation activities that foster fine motor skills. Activities that foster problem-solving are among the harder activities for trainers to carry out, given their limited preparation in this area.

Figure 1: Effects on Child Development



Note: Results in gray are insignificantly different from zero. Results in green are significantly different from zero.

9.5.2 Impacts on Caregiver Behaviors

With regards to caregivers, the main findings (preliminary) are organized according to self-reported data and data capturing observed behavior of caregivers. For self-reported data, which includes information on behaviors such as reading stories, quality of parent-child interactions, support with homework, family outings, the study finds no pattern of differences between treatment and control (Table 37).

Table 37: Preliminary Effects on Reported Caregiver Behavior (Selected variables)

	(1)	(2)	(3)	(4)	(5)	(6)
	At least one children's book	Mom reads stories at least once a week	Takes child to market 2+ times per week	At least 3 stuffed animals	Parents spend time teaching child	Child goes out at least 4 times a week
Treatment * After	-0.05	0.01	0.06	0.03	-0.01	-0.07*
	(0.033)	(0.040)	(0.043)	(0.036)	(0.029)	(0.041)
After	0.50***	0.12***	-0.06*	0.09***	0.04**	0.05*
	(0.023)	(0.028)	(0.029)	(0.025)	(0.020)	(0.028)
Constant	0.33***	0.37***	0.56***	0.65***	0.89***	0.38***
	(0.010)	(0.011)	(0.011)	(0.009)	(0.007)	(0.011)
Observations	2,729	2,066	2,068	2,053	2,068	2,063
R-squared	0.390	0.039	0.004	0.034	0.008	0.004
Number of hogar1	1,355	1,164	1,163	1,159	1,164	1,162

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

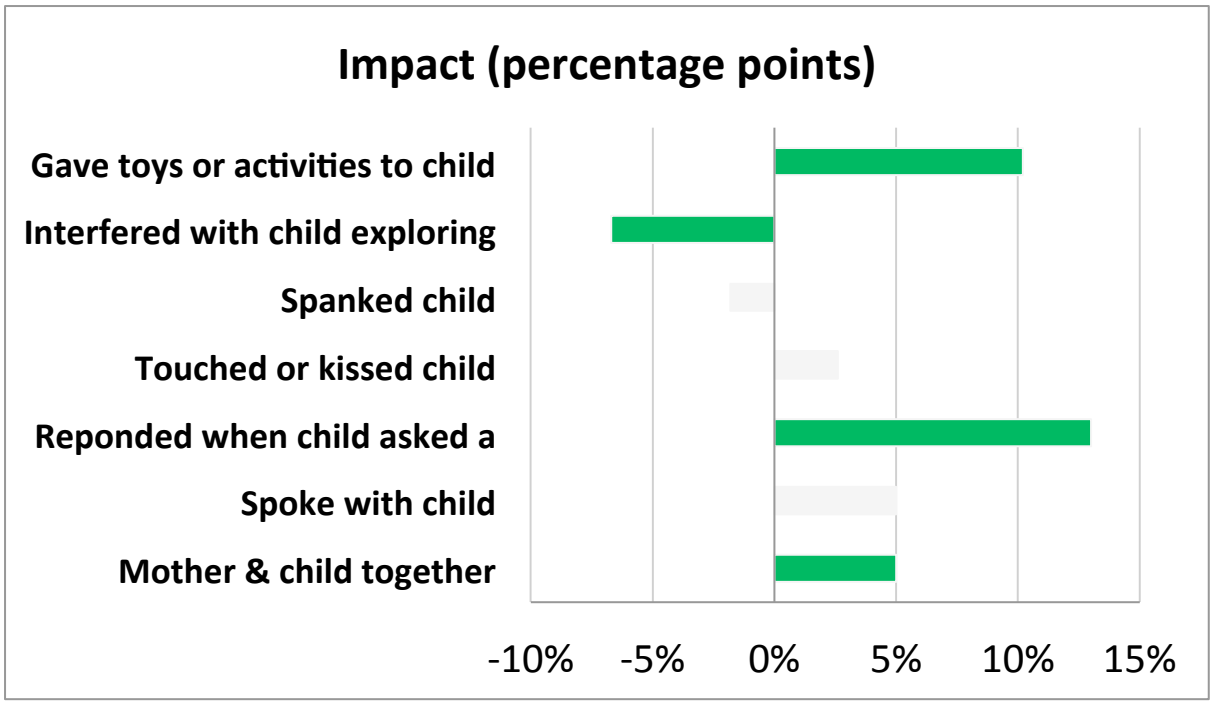
However, when looking at the observed behavior, there are significant differences. The set of questions asked differed slightly for children age 0-35 months and for children age 36-42 months, so these are analyzed separately. For the younger children, we find positive (i.e., desired) and significant differences between treatment and control on four of the eight behaviors (see Table 38 and Figure 2). Note that, for the second variable of interest on the list (interfering with child actions), this refers to caregivers interfering with children exploring their environments. Therefore, a negative score is desirable. Of the four variables that show insignificant effects, all have the desired sign, at least ruling out the possibility of major detrimental impacts of the program. For the older children, the sample is significantly smaller and the measured effect sizes are both insignificant, as well as having point estimates close to zero (Table 39).

Table 38: Preliminary Effects on Observed Caregiver Behavior (Children age 0-35 months)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mom & child together	Mom spoke positively 2x	Responded verbally	Hugged or kissed	Hit or spanked	Stopped child exploring	Gave toys or activities	Kept child in sight
Treatment * After	0.05**	0.05	0.13***	0.03	-0.02	-0.07**	0.10**	0.04
	(0.020)	(0.040)	(0.044)	(0.041)	(0.014)	(0.033)	(0.041)	(0.028)
After	-0.07***	0.08***	0.02	-0.07**	0.02*	-0.00	-0.12***	0.05**
	(0.014)	(0.028)	(0.030)	(0.028)	(0.010)	(0.023)	(0.028)	(0.019)
Constant	0.97***	0.69***	0.69***	0.73***	0.02***	0.17***	0.32***	0.86***
	(0.005)	(0.010)	(0.011)	(0.010)	(0.004)	(0.008)	(0.010)	(0.007)
Observations	2,058	2,050	2,057	2,053	2,053	2,035	2,032	2,035
R-squared	0.032	0.034	0.027	0.009	0.004	0.011	0.020	0.026
Number of children	1,168	1,167	1,168	1,167	1,167	1,165	1,166	1,166

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 2: Effects on Caregivers (Observed)



Our preliminary conclusion is that the program is likely having a modest impact on both children and their parents, particularly for caregivers of younger children. If further analysis bears out these results, then it is likely that this intervention is therefore very cost effective, especially in comparison to other service-delivery options for achieving these changes, such as home visitation.

Table 39: Preliminary Effects on Observed Caregiver Behavior (Children age 36-42 months)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mom & child together	Mom spoke positivel y 2x	Responded verbally	Hugged or kissed	Introduced by name to enumerator	Shook or restrained physically	Hit or spanked	Voice reflect ed positiv e feeling about child
Treatme nt * After	-0.02	0.16	0.10	-0.02	-0.02	-0.04	-0.02	-0.06
	(0.075)	(0.103)	(0.094)	(0.108)	(0.120)	(0.048)	(0.046)	(0.100)
After	-0.03	-0.16**	-0.14**	-0.09	-0.03	0.08**	-0.02	0.10
	(0.051)	(0.070)	(0.063)	(0.073)	(0.081)	(0.032)	(0.031)	(0.068)
Constant	0.94***	0.82***	0.85***	0.65***	0.53***	-0.00	0.04**	0.77***
	(0.029)	(0.040)	(0.037)	(0.042)	(0.047)	(0.019)	(0.018)	(0.039)
Observat ions	657	647	650	658	640	637	654	638
R-squared	0.009	0.036	0.036	0.023	0.004	0.055	0.010	0.018
Number of children	510	506	508	512	502	500	510	501

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

9.5.3 Caveats

One of the main aspects to be considered is that the databases used in this analysis account for a particular design that only aligns to the purpose of performing panel data estimations. In this regard it is important to note that the number of cases included in the impact analysis is not necessarily corresponding with the number used to calculate the results shown in sections 9.1 to 9.4, thus the databases used are not exactly the same. On the one hand we have the panel database that only integrates matched cases: 1) they represent observations found in both the baseline phase and the second wave, and 2) every case registered has its unique pair composed by the following: one observation from the baseline and the other corresponding to the second wave. On the other hand we have the databases used to estimate the aggregated results presented in the sections previously mentioned in which all cases were included no matter if they matched or not.

Another relevant consideration is that the impact results reported in this subsection are extremely preliminary. They represent simple difference-in-differences analysis between children and caregivers at treatment and control villages who were surveyed at both baseline and follow-up. As such, they require significant sensitivity analysis, which will take place in the coming weeks and months. For example, incorporating appropriate clustering will be essential, as will adjusting for imbalanced attrition: Unusually, attrition was significantly higher in the treatment group than in the control group at the first follow-up. For the next follow-up, the team will design a special tracking protocol to overcome the attrition from the first follow-up. (The team is optimistic as to the likely success of this exercise, given that most households can be located and some inducement may be sufficient to encourage their participation.) Furthermore, the team will use propensity score matching to improve the quality of the analysis for the first follow-up, matching each remaining treatment observation with each comparison observation and carrying out an analysis of attrition to quantify – as far as possible – the implications of the short-term attrition for external validity of the sample.

Finally, a more nuanced analysis will reveal whether effects are concentrated in certain ages, or certain states and municipalities. While the study may not have sufficient statistical power to reject equality of effect sizes across sub-groups, such analysis will still be useful to identify whether there are significant positive effects for certain sub-groups that do not exist for the sample as a whole. That said, this preliminary analysis, robust to some initial sensitivity analysis (i.e., dropping a subset of observations with surprising values), is suggestive of positive impacts.

10 Balance between treatment and control groups

As in the previous report, we present a comparison between the characteristics of households in treatment and control groups, as well as child rearing practices by parents and caregivers, and the level of development in children. This is intended to demonstrate that an equitable distribution in the context and exogenous factors affecting the program is maintained. First, the differences that were statistically significant in the baseline survey will be reported and compared with the results of the follow up survey. Subsequently, new differences in the follow up survey data will be presented.

10.1 Home Balance

In the baseline report, the participation group households had a higher average of members compared with the comparison group. For the follow-up questionnaire, however, the difference between groups is not significant. The number of children at an appropriate age to participate in the Early Childhood Education program however, remains significantly different in both cases. In the baseline survey, participation group households had a slightly higher average, but in the second survey, this difference was only significant for a reliability level of 90%.

On housing conditions, in the first report a significant difference in the flooring materials for participation group was found, although this difference is no longer significant in the follow up data. The same applies to the way homes are supplied with water, however the difference persists on waste management in households: nearly 35% of households in the comparison group was connected to the public sewage system versus 24% in the participation group; the difference is considered significant with a 1% margin of error.

In terms of property ownership, baseline results showed that in participation group households the proportion of coal stoves and power meters was higher than the comparison group households. In the follow up survey, the results for comparison group households are greater for possession of wood or coal stoves. This is a significant result with a 95% level of reliability. The difference in the possession of power meters is similar to the last survey, and significant with 99% reliability.

In addition to the differences in the baseline survey, the follow up survey data suggests five differences in asset ownership, all in favor of the participation group: differences in the possession of boilers, showers and computers are significant at a level of 99%reliability. In the case of households that reported having internet service, the difference is significant at 95%, and in the case of a radio, the difference is significant with only 90% reliability.

Table 40: Results from Homes by group

	Comparison	Participation	Difference	
Number of members	4.38	4.474	0.094	
Children aged 60 mos.	1.482	1.391	0.091	*
Pregnant women	0.095	0.069	0.026	
Brick walls	0.789	0.786	0.003	
Spanish speaking members	0.930	0.930	0.000	
Members able to read and write	0.767	0.768	0.001	
Housing conditions				
Concrete roof	0.533	0.562	0.029	
Dirt floor	0.071	0.061	0.010	
Total rooms	2.972	3.155	0.183	
Bedrooms	1.640	1.741	0.100	
Days with electrical power	6.517	6.457	0.060	
Plumbing within the house	0.336	0.365	0.028	
Sewage	0.349	0.245	0.104	***
Assets				
Gas stove	1.317	1.263	0.055	
Wood stove	1.287	1.356	0.069	**
Water tank	1.475	1.491	0.016	
Boiler	1.856	1.767	0.089	***
Cistern	1.834	1.818	0.016	
Shower	1.829	1.714	0.115	***
Power Meter	1.272	1.158	0.115	***
Automobile	1.769	1.751	0.018	
Internet	1.992	1.968	0.024	**
Computer	1.977	1.939	0.039	***
Telephone (land line)	1.930	1.912	0.018	
Cellular phone	1.456	1.413	0.043	
Washing machine	1.678	1.629	0.049	
Refrigerator	1.486	1.452	0.034	
Television	1.153	1.174	0.021	
Radio	1.446	1.385	0.061	*

10.5 Caregiver Balance

The first difference found in the baseline report was that participation households had a higher rate of *Oportunidades* beneficiaries, however, this difference does not hold with the follow up survey data. Secondly, caregivers of children from 0 to 35 months in the participation group reported reading or telling stories to their children more often but, again, this difference does not continue in the second survey. In this same age group, caregivers in the comparison group more frequently allowed children to choose what they wanted to eat, but this difference also no longer holds after a year.

For caregivers with children ages 36 to 42 months, the comparison group had a higher frequency of parents living together, although this difference was reversed in favor of the participation group, with a reliability level of 90%. Finally, the difference that caregivers in the comparison group more frequently took their children of 36 to 42 months to historical or cultural sites was not maintained either.

In addition to these changes, new differences emerged: among caregivers in the Participation Group it was more common for someone to have help with the child, this result is significant at a reliability level of 95%. It was also more common for the participation group respondents to affirm that the household receives financial support from people living within the country, although this difference only holds with 90% reliability.

Among caregivers with children 0 to 35 months, we found a difference of 38% in the possession of children's books for the comparison group, a significant difference at a reliability level of 99%. Among caregivers with children 36 to 60 months, the comparison group more often reported that the child had a device for listening to music, although this result only holds with a 90% reliability level. It was also more often that comparison group caregivers spoke with the child when they were busy with the same reliability level. Likewise, comparison group caregivers reported more frequently that the child feels something or is very close to his father, a significant difference with 95% reliability.

In the group of 36 to 60 months, only two differences in favor of the participation group were presented: it was more common for the child to eat with his father at least once per day, and this difference is maintained with 99% reliability. Finally, the average spanking the caregiver had given the child in the week prior to the survey was lower, and the

difference is significant at a reliability level of 95%. The table below shows the total results for the caregiver questionnaire by study group.

Table 41: Results from caregivers by group

	Comparison	Participation	Difference	
Child's age	2.311	2.292	0.019	
Length of time residing in the community	19.793	18.94	0.853	
Mentioned a second caregiver	0.734	0.786	0.052	**
Beneficiaries of *Oportunidades*	1.58	1.518	0.061	
Beneficiaries of *Seguro Popular*	1.314	1.297	0.017	
Beneficiaries of *Procampo*	1.954	1.982	0.027	
Beneficiaries of scholarships	1.885	1.916	0.031	
Senior Citizen Beneficiaries	1.98	1.983	0.003	
Financial support from people within the country	1.982	1.945	0.037	*
Financial support from people abroad	1.973	1.957	0.016	
Attended a government clinic	1.003	1	0.003	
0 to 36 months				
Children's books	0.944	0.564	0.38	***
Story books	0.738	0.696	0.042	
Soft toys	0.926	0.924	0.002	
Pull toys	0.749	0.724	0.025	
Parents live together	1.13	1.144	0.013	
Frequency with which the child eats with the father	0.783	0.801	0.018	
Frequency with which the child is spoken to	0.725	0.714	0.01	
Closeness of child to the father	0.861	0.854	0.007	
Closeness of child to the mother	0.957	0.939	0.019	
Corporal punishment	0.856	0.93	0.074	
36 to 42 months				
Children's books	0.778	0.803	0.025	
Story books	0.597	0.640	0.044	
Magazines	0.407	0.366	0.041	
Device for listening to music	1.506	1.583	0.077	*
Helps child learn numbers	1.082	1.09	0.008	
Helps child learn letters	1.135	1.098	0.037	
Helps child learn colors	1.102	1.075	0.028	
Helps child learn shapes and sizes	1.201	1.177	0.024	
Child chooses what to eat	0.512	0.518	0.005	
Visits to museums or historical sites	0.172	0.149	0.023	
Parents live together	1.147	1.098	0.049	*

	Comparison	Participation	Difference	
Frequency with which the child eats with the father	0.753	0.851	0.098	***
Frequency with which the child is spoken to	0.745	0.667	0.078	*
Closeness of child to the father	0.891	0.815	0.076	**
Closeness of child to the mother	0.934	0.917	0.017	
Corporal punishment	1.308	0.784	0.524	**
Child rearing practices				
Medical checkups	1.419	1.41	0.009	
Plays with child	1.038	1.06	0.022	
Sings with child	1.196	1.174	0.021	
Musical instrument	1.606	1.589	0.017	
Child's disease	1.687	1.654	0.033	
Expected academic achievement	0.741	0.74	0.001	
Preschool enrollment	1.02	1.017	0.003	
Preschool enrollment age	4.147	3.960	0.186	

10.3 Pregnant women balances

The baseline questionnaire only reported a significant difference in the residence time in the community. However, as this questionnaire could not allow a respondent different to the respondent of the previous questionnaire, this question was omitted for the follow-up questionnaire. For the second survey, two new differences between groups were identified: first, the mothers in the comparison group most frequently mentioned that someone helps them in the care of their child, a difference that is significant with a reliability level of 95%. Furthermore, in the same group of mothers it was more common to exchange tips with other pregnant women, although this result is only maintained with 90% reliability. The other variables of the questionnaire presented no significant differences.

Table 42: Results from pregnant women by group

	Comparison	Participation	Difference	
Presently pregnant	1.92	1.815	0.105	
Months pregnant	2	5.2	3.2	
Mentioned a secondary caregiver	0.84	0.593	0.247	**
Beneficiary of *Oportunidades*	1.786	1.593	0.193	
Financial support				
Beneficiary of *Seguro Popular*	1.5	1.259	0.241	
Beneficiary of *Procampo*	1.818	1.926	0.108	
Beneficiary of a scholarship	1.96	1.852	0.108	
Support from persons living abroad	2	1.963	0.037	
Months since delivery	9.625	10.577	0.952	
Care during pregnancy				
Medical checkups	7.760	6.192	1.568	
Non prescribed medicine	1.913	1.926	0.013	
Father accompanied expectant mother for medical checkup	1.32	1.259	0.061	
Sickness or injury during pregnancy	1.875	1.778	0.097	
Exchange of advice from other expectant mothers	1.44	1.704	0.264	*
Child received a vaccination at birth	1.048	1.167	0.119	
Alcoholic beverages	0.04	0	0.04	
Coffee	0.68	0.593	0.087	
Soda	0.84	0.778	0.062	
Someone smokes inside the house	0.08	0.038	0.042	
Change in eating habits	1.792	1.615	0.176	
Care after pregnancy				
Breastfed child	1.043	1	0.043	
Months of lactation	8.5	10.208	1.708	
Children are aware of their surroundings	1.045	1.037	0.008	
Preschool enrollment	1	1.111	0.111	
Preschool enrollment age	3.818	3.957	0.138	
Expectation of exceeding basic education	0.727	0.577	0.15	

10.4 Father balance

In the baseline questionnaire, two differences between the groups were presented: the parents in the participation group more frequently asserted that they helped feed the child, the same group also responded more frequently that the child felt something or was very close to them. However, neither difference is supported by the follow up survey data.

Besides these, data from the second survey yielded two new differences: 25% of parents in the comparison group said they argued with their partner during the past week, while no one in the participation group reported having done so, and this difference is significant at a reliability level of 95%. Also, all parents in the participation group stated that they discuss their children with other parents, compared to 92% in the comparison group, and this difference holds by a reliability level of 99%.

Table 43: Results from fathers by group

	Comparison	Participation	Difference	
Length of time living in the community	27.825	25.6875	2.1375	
Medical checkups	1.425	1.25	0.175	
Plays with child	1.025	1	0.025	
Sings with child	1.075	1	0.075	
Helps prepare meals for the child	1.2	1.25	0.05	
Helps feed the child	1.075	1.0625	0.0125	
Helps dress the child	1.15	1.0625	0.0875	
Carries the child	1.1	1	0.1	
Closeness to child	0.925	1	0.075	
Expectations of academic achievement	0.775	0.875	0.1	
Preschool enrollment	1.05	1	0.05	
Social Conduct				
Smokes	1.825	1.875	0.05	
Smokes inside the home	1.857143	2	0.142857	
Decides on spending without discussion	1.55	1.625	0.075	
Has financial problems	1.5	1.3125	0.1875	
Speaks with family to resolve problems	1.075	1	0.075	
Speaks with family when feeling sad	1.15	1.125	0.025	
Speaks with family when feeling angry	1.425	1.375	0.05	
Alcoholic beverages	0.3	0.3125	0.0125	
Argued with partner	0.25	0	0.25	**
Discussed his children with other parents	0.925	1	0.075	***

10.5 ASQ balance

As described in the baseline report, ASQ results cannot be summarized in a single measure of performance for each child, but can only show the final score in each assessment area. This is because the cutoffs for each age and each assessment area are different. As we cannot show differences in every case, we will only be able to replicate the same cases that were in the baseline report to verify that they behave as expected.

During the baseline survey, in the comparison group children of 2 months of age had higher scores in the Gross Motor and Problem Solving areas, although the difference is only held by a reliability level of 90%. With the follow-up survey data however, the differences are no longer significant.

Table 44: Results from ASQ 2 months by group

	Comparison	Participation	Difference
Communication	42.7	47.5	4.8
Gross Motor	52.0	53.7	1.7
Fine Motor	50.0	48.4	1.6
Problem Solving	42.5	44.5	2.0
Personal-Social	50.5	50.8	0.3

After reviewing all ASQ cases, we found that the most important changes in results in the baseline survey occurred in the 6 months group. Children at this age last year showed no significant differences in any of the five areas of assessment, however children who are at this age now show a remarkable difference except in the Personal-Social: the comparison group displayed higher scores in all areas of assessment. However, we must be cautious with these results, as in this age group only six children were found to be attending Early Childhood Education sessions.

Table 45: Results from ASQ 6 months by group

	Comparison	Participation	Difference	
Communication	53.4	45.8	7.6	**
Gross Motor	48.1	37.0	11.1	**
Fine Motor	55.4	47.5	7.9	*
Problem Solving	54.4	45.0	9.4	**
Personal-Social	48.1	39.2	9.0	

Among children who were 12 months of age at the time of the baseline survey, the highest score was presented in the comparison group among fine motor activities, although this result does not hold in the follow up data, and none of the other areas presented significant differences.

Table 46: Results from ASQ 12 months by group

	Comparison	Participation	Difference
Communication	49.1	51.7	2.5
Gross Motor	44.3	43.8	0.6
Fine Motor	42.0	37.9	4.0
Problem Solving	40.9	46.7	5.8
Personal-Social	45.9	45.0	0.9