



**Evaluation of USDOL
Services to Prevent Child
Labor in Bolivia:
Randomized Trial to
Evaluate *Desarrollo y
Autogestión's Ñaupaqman
Puriy-Kereimba-Chic'k'y
Wawita* (Combating
Exploitive Child Labor in
Bolivia) Extended Hours
School Program, 2011**

Baseline Report

September 2012



Submitted to:

U.S. Department of Labor

Office of Child Labor, Forced Labor and Human Trafficking

**Evaluation of USDOL Services
to Prevent Child Labor in
Bolivia: Randomized Trial
to Evaluate *Desarrollo y
Autogestión's Ñaupaqman
Puriy-Kereimba-Chic'k'y
Wawita* (Combating
Exploitive Child Labor in
Bolivia) Extended Hours
School Program, 2011**

Baseline Report

September 2012

Submitted to:

U.S. Department of Labor
Office of Child Labor, Forced Labor and Human Trafficking
Frances Perkins Building
200 Constitution Avenue NW
Washington, DC 20210

Submitted by:



11785 Beltsville Drive
Calverton, MD 20705

REPORT PREPARATION

Authors

Stacey S. Merola
Sarah Edith Jones
Matthew C. Holtman

ACKNOWLEDGEMENTS

This report is based on evaluation research undertaken by ICF International under its contract with the U.S. Department of Labor's Office of Child Labor, Forced Labor, and Human Trafficking (OCFT). The authors wish to thank the staffs of OCFT and of *Desarrollo y Autogestión* for their support and collaboration in the design and execution of the study.



Funding for this evaluation was provided by the United States Department of Labor under Task Order number DOLB109K31094. Points of view or opinions in this evaluation report do not necessarily reflect the views or policies of the United States Department of Labor, nor does the mention of trade names, commercial products, or organizations imply endorsement by the United States Government.

CONTENTS

Report Preparation	ii
Acknowledgements	iii
List of Acronyms	v
Map of Bolivia and Project Sites	vi
I Introduction	1
1.1 The Evaluation	1
II Methods	3
2.1 Participants	3
2.2 Interventions	3
2.2.1 The Extended Hours Program for Children Under 14	4
2.2.2 The Extended Hours Program for Teens	4
2.3 Objectives	6
2.4 Data Collection	6
III Sample	9
3.1 Power Analysis	9
3.2 Recruitment	10
3.3 Randomization	11
3.3.1 Sequence Generation	11
3.3.2 Allocation Concealment	11
3.4 Blinding	12
3.5 Statistical Methods	12
IV Results	13
4.1 Participant Flow	13
4.2 Analysis	16
4.2.1 Rural Schools	16
4.2.2 Urban Schools	22
V Discussion	28
5.1 Interpretation	28
5.2 Generalizability	28
5.3 Other Limitations	29
5.4 Next Steps	30

LIST OF ACRONYMS

DyA	Desarrollo y Autogestión
ILAB	Bureau of International Labor Affairs
IRB	Institutional Review Board
NGO	Non-governmental organization
ÑPKCW	Ñaupaqman Puriy–Kereimba-Chic'k'y Wawita
OCFT	Office of Child Labor, Forced Labor, and Human Trafficking
RCT	Randomized Controlled Trial
UNICEF	United Nations Children's Fund
USDOL	United States Department of Labor

MAP OF BOLIVIA AND PROJECT SITES



I INTRODUCTION

The Office of Child Labor, Forced Labor, and Human Trafficking (OCFT) is part of the United States Department of Labor's Bureau of International Labor Affairs (ILAB). The office was created in 1993 in response to a request from Congress to investigate and report on child labor around the world. OCFT has provided hundreds of millions of dollars to support programs to reduce or eliminate child labor around the world. In 2009 it awarded \$6 million to Desarrollo y Autogestión (DyA) to implement the Ñaupaqman Puriy-Kereimba-Chic'k'y Wawita (ÑPKCW) program in Bolivia starting in 2010. DyA is a nongovernmental organization (NGO) based in Quito, Ecuador, with offices in Bolivia and significant experience developing and managing programs to combat child labor.

The magnitude of child labor is significant in Bolivia. The United Nations Children's Fund (UNICEF) estimates that more than 850,000 children and adolescents work, which represents 21% of the population between ages 5 and 14. This is higher than the average prevalence of child labor in the region, which is estimated at 16%. Children in Bolivia are most commonly found working in the worst forms of child labor in sugarcane and chestnut harvesting, mining, domestic service, and construction.¹

In the rural areas of Bolivia, child labor exists primarily in the agricultural and domestic service sectors. Such labor can include work on family farms, in commercial agricultural enterprises, and in private homes. Forced labor of families exists on private ranches through debt servitude or similar mechanisms. In urban areas of Bolivia (particularly in the Plan 3000 community of Santa Cruz and the neighborhood of El Alto in La Paz), child laborers work in a variety of sectors, including domestic work, manufacturing, street vending, recycling, construction, restaurants, and transportation.²

ÑPKCW is an integrated program with many different interventions designed to work together to improve the situation of child labor in Bolivia. ÑPKCW includes national-level policy strengthening work, awareness raising, livelihood services, and educational services.

The ÑPKCW project aims to reduce the number of children working in child labor in Bolivia by increasing school enrollment, reducing the number of hours children work, and removing them from hazardous and exploitive work situations. The project provides direct educational services to enhance and strengthen the existing educational system. It also strengthens child labor and education policies by working directly with government agencies to develop and enact policy changes. It provides services to improve opportunities for youth employment and alternative income generation, as well as vocational training. Finally, it works with indigenous organizations, parents, and teachers to raise awareness of the problems associated with child labor and the benefits of education for children.

1.1 THE EVALUATION

ICF International, a diversified professional services firm, has more than 40 years of experience leading complex research projects in developing countries. Under task order DOLB109K31094, ICF International is providing a variety of technical assistance and evaluation services to programs funded by OCFT. In April 2011, an impact evaluation was started under ICF's guidance to evaluate a specific element of the ÑPKCW project, the Extended Hours school program.

¹ Flores, E. (2008). *La Problemática del Trabajo Infantil en los Pueblos Indígenas de Bolivia: Estudio Preliminar*. Lima, Peru: Centro de Estudios Jurídicos e Investigación Social.

² *Ibid.*

The goal of the evaluation is to measure the Extended Hours program's impact on children's participation in school and work. The study design was approved by ICF International's Institutional Review Board (IRB).

An RCT always entails some ethical concerns. One question that has been raised by the IRB and DyA staff is whether it is ethically acceptable to provide services to some children and not others. In designing the evaluation, ICF staff made the argument that USDOL-funded programs, like all public benefit programs, have finite resources and cannot cover all the children who would benefit from them. There are typically more families and children who would like to receive DyA's services than can be accommodated given the resource levels available. Under these circumstances, random assignment is usually the fairest approach to allocating services. However, perceptions of unfairness may persist in particular because intervention and control children live in the same communities and, in many cases, come from the same families. A reasonable solution to this problem, when it is feasible, is to randomize at the community rather than individual level. As will be discussed in the Other Limitations section below, community-level randomization was not possible under the resource constraints of this program.

II METHODS

2.1 PARTICIPANTS

The ÑPKCW Extended Hours program is targeted to children in certain indigenous communities of Bolivia who are engaged in or at risk of engaging in child labor. These children come predominantly from poor families in both rural and urban areas of Bolivia.

The DyA project administrators believe that educational problems and child labor are closely connected. Many children who work either miss school or perform poorly in school in part because work prevents them from attending consistently. The Extended Hours program is designed to increase school persistence by helping children who are struggling in their studies, while at the same time reducing the number of hours children work by keeping them engaged in educational support activities during hours when children typically work.

The urban areas reached by the project include El Alto, a neighborhood of the capital city La Paz; San Julian, a city to the northeast of industrial center Santa Cruz; and Plan 3000, an urban settlement within Santa Cruz. The rural communities are the predominantly ethnic Guaraní villages in the Chaco region. In the urban areas, there are separate youth and adolescent Extended Hours programs, whereas in rural communities a multi-grade approach is used.

Because the urban and rural programs are different, it was decided to run essentially two separate impact evaluations in parallel: one for the urban, Extended Hours program divided into two age groups, and one for the multi-grade rural program. Separate analyses will be presented for the two programs.

2.2 INTERVENTIONS

Extended Hours services are provided to children who are at risk of abandoning school in order to work. The Extended Hours program provides an additional 3.5 hours of play, sports, and informal instruction per day, at the end of the regular school day. The Extended Hours program takes place four days a week for primary school students and two days per week for secondary school students. It is designed to shift the balance of children's time away from work activities and toward study and to reinforce and strengthen the children's academic skills. The program uses existing school infrastructure and provides training and remuneration to teachers to support these activities. It also involves community members as tutors. The Extended Hours program runs for two years.

As noted above, the Extended Hours program in rural areas is a multi-grade program that groups children and adolescents together in the same classrooms.

The Extended Hours program in urban areas is conducted in two separate groups based on beneficiaries' ages. In both groups, there are different (but overlapping) sets of interventions based on the participants' ages. The first intervention is for children under 14 years and consists of play activities, music, and other activities. The second group is for children and adolescents who are 14 to 17 years old and consists of tutoring, sports, and livelihood services, such as learning to plant a garden. These intervention elements are described in more detail in the section that follows.

2.2.1 The Extended Hours Program for Children Under 14

The aim of the Extended Hours program is twofold: first, to reduce the number of working hours of children attending schools in areas of project intervention; and second, to reinforce children's educational achievement through extracurricular activities that are aimed at—

- Providing support to children in areas identified by their teachers;
- Promoting self-esteem and social integration;
- Strengthening language proficiency in Spanish; and
- Introducing content to prevent hazardous child labor.

The extracurricular activities are conducted by special tutors and they take place in the afternoon following completion of the regular school schedule. The program is supported by a special curriculum, tutor training, and other pedagogical tools designed to foster a climate favoring creativity, with a flexible approach giving priority to games as a central learning tool. The curriculum includes play, recreation, and informal education activities, and it seeks to develop children's skills in reading and writing, listening, and speaking.

In order to develop their interest in reading, children are introduced to literature selected to be appropriate to the group's reading level. An important contribution of ÑPKCW was the development of a collection of reading material that includes stories based on personal histories of working children to illuminate many of the problems of child labor. This reading work is intended to help students develop their interest in literature so that they will start reading on their own.

The program addresses students' writing skills through the creation and production of texts. Art is also used, including media such as painting, drawing, puppets, theater, social dramas, and puzzles.

Tutors use a curriculum planning tool and guides to structure their daily work based on the children's age group.

The Extended Hours program also has a training component for tutors, delivered through a series of three workshops. The training is aimed at ensuring that tutors can effectively use the program's methodology in the classroom.

2.2.2 The Extended Hours Program for Teens

The Extended Hours program for teens is also conducted in the afternoon following the regular school day. This program provides recreational and cultural activities for teenagers to reduce their participation in child labor and to prevent problems, such as drug and alcohol use and teen pregnancy.

The teens have access to a course called Applied Technical Training and work on developing the skills they need to make better decisions on matters that have a direct impact on their lives and welfare. Resources such as reading material, media, computers, and musical instruments are also provided.

The curriculum has three themes:

1. Applied technical training
2. Youth issues
3. Arts and sports

Applied technical training is a curriculum designed to train teens to analyze community problems and work with their communities to develop and manage development proposals. It is based on the development of job skills that involve both technical knowledge and self-efficacy.

The second theme provides information on topics such as sexual and reproductive health, drugs, and alcohol. Work issues are linked to leadership and self-esteem through creative activities that engage the interest of teenagers.

The third component uses sports to create a healthy space for training and reinforcement of values such as teamwork and the importance of following rules. The program also promotes artistic activities (music, pottery, dance, and painting) to stimulate teens' creativity and offer them an alternative way to express their thoughts and emotions. There are also exchanges with teenagers from other communities and schools.

The program has a duration of 16 months, spread over 2 academic years. The program is carried out by trained tutors whose work will be complemented by specialists from partner institutions (e.g., local universities and health centers). Tutors are given training through three workshops per year to develop their ability to manage the program's methodology. "Clinical" training is also provided to help the tutors enrich their classroom practice.

The program is supported by an educational mediator present in each local office of the project. The mediator supports tutors in the development of planning activities and ensures the provision of required materials. School boards work with parents to coordinate activities. District directorates monitor schools that have Extended Hours programs and promote participation in community activities, such as fairs and sports events, that are aimed at increasing parents' awareness of the problems of child labor.

Analysis of the baseline data shows that there were only 35 children enrolled in the evaluation who met the age requirement to be included in the adolescent group. This is not a large enough group to merit separate analysis. Furthermore, those 35 children are distributed across 5 different schools. It seems unlikely that teachers are providing a completely separate curriculum to those children. If the instruction in the urban program is not, in fact, strongly differentiated by age, then it is appropriate to group the adolescents and children in the analysis. Therefore, rather than drop those children's data from the study, we elected to include them. This provides greater comparability with the rural program, which also has a number of children age 14 and above.

2.3 OBJECTIVES

It is hypothesized that the Extended Hours program will increase school attendance, reduce school abandonment, and reduce the proportion of children who work. Among children who do report working, it is also hypothesized that those children who participate in the Extended Hours program will work fewer hours and will be less likely to work under hazardous conditions.³

Improving academic achievement, including basic reading, writing, and mathematics skills, is a goal of the Extended Hours program but is not a focus of the evaluation. This is because we do not expect to see measurable improvements in skills over the course of a single year even if the program does have a beneficial effect on these skills over the long term. In the evaluation, we are looking at attendance and enrollment, which are important educational outcomes in themselves that should also be correlated with later improvements in academic performance.

The objective of the present report is to describe the randomization and baseline survey process and show whether randomization was effective in balancing the intervention and control groups. The baseline data will also be used in post-intervention analyses to provide statistical control in order to improve the precision of impact estimate measures. Future data collection efforts will measure additional outcomes, such as exposure to hazards at work, which are important for identifying children working in child labor.

2.4 DATA COLLECTION

The study dataset was generated through a household survey. The study focused on 910 children who were randomly assigned to either an intervention or control condition. The intervention condition was participation in the Extended Hours school program, along with exposure to adjunct community activities (such as awareness-raising fairs) that were implemented by project staff in the villages where the children lived, and exposure to a small educational packet that included a few books and some writing materials. The control condition was exposed to community activities, the educational packet, and a small food ration (received once) without enrollment in the Extended Hours program. The food ration was provided to the control group children as compensation for their participation in the study. It was small enough that it was not expected to have a substantial impact on household economic behavior. Thus, we measured the additional impact of the Extended Hours program above and beyond whatever impact the awareness-raising, curriculum development, teacher training, and other community-level interventions were having on children's educational and child labor outcomes.

Furthermore, because children in the intervention and control groups were recruited from the same communities, and because both groups are exposed to community-level awareness-raising activities, there was a strong possibility that "spillover" effects will occur. This is discussed in greater detail in the Other Limitations section below.

The presence of community-level interventions in both the intervention and the control groups, combined with the possibility of spillover effects, may result in impact estimates that are lower than they would be in the context of a comparison across communities. An alternative community-level cluster randomization

³ Information on exposure to hazards at work was not collected in the baseline survey, but will be collected in the follow-up survey because discussion of the baseline results has identified work hazards as an important outcome. Hazards include such things as exposure to smoke or chemicals, working in the dark, working with sharp tools, etc. With this information we will not be able to explore changes in exposure risk at the individual level, but will be able to show whether control and intervention group children are exposed to hazards at different rates.

design was considered but was felt to be infeasible because of the resource requirements for identifying and screening villages for eligibility—and for political reasons, since program staff felt that it would be very difficult to enroll villages in the study if they did not receive the intervention program.

The evaluators, implementing agency, and sponsors of the program weighed these tradeoffs in arriving at the current evaluation design.

Enumerators conducted one household survey for every child who was enrolled in the study. Many of the children in the study lived within the same households. There were up to four children per household who were assigned an intervention status (either intervention or control). Because a lottery was used for assignment, a group of siblings within the same household could receive the same or different intervention assignments on a random basis. In principle, this would allow analysis of dose-response effects at the household level, though in practice the number of cells with different assignment patterns were likely to be too small to afford sufficient statistical power to support more than exploratory analyses.

In addition to collecting information on children who participated in the random assignment process, the enumerators conducted interviews to collect data on employment status of all adult household members. In cases where there was more than one participating child in a household, the household data were only collected once. Detailed data on time allocation and school expenditures were collected for study children and their siblings, and time allocation data were also collected for study children's parents. Again, if more than one child in a family participated in the evaluation, parents' time allocation data and school expenditure were only collected once. In cases where the children were under 11 years of age, the head of household was the primary respondent and provided the information on both adults and children. In cases where children participating in the evaluation were 11 or older, they were interviewed directly to obtain time allocation data.

It was also possible for more than one family to be living in a house together. Therefore, data were also collected on family clusters within each household. This information is used primarily to supply supplemental information to help understand family structure and relationships within the household and to ensure that analyses are being conducted appropriately.

The study was designed to estimate the impact of the urban and rural Extended Hours programs on children's school attendance, working status, and quantity of work.

School attendance was assessed with a question asking whether the child was currently in school. Work hours were measured using a special retrospective time record that was developed for the project based on tools that project staff have been using successfully for several years to assess and monitor program beneficiaries' work participation. The time record is a simplified version of the clock that the project staff used, and it asks about children's allocation of time in one-hour increments throughout the day for a typical weekday, a typical Saturday, and a typical Sunday.

The time record was filled out by enumerators who were trained to ask respondents to tell them what they typically do during specific hours of the day. When children were not able to remember the exact timing of their activities, enumerators were trained to ask about the sequence of events. For example, a child might be asked, "After you have breakfast, what you do?" For children under the age of 11, an older family member generally provided the response. Any gaps in the timing of recorded activities were filled by probing, and the enumerators worked with the children to estimate the amount of time spent in each activity. Enumerators recorded the codes corresponding to the children's responses using the list of codes supplied on the instrument.

The survey instrument was piloted between March 15 and April 10, 2011. The main data collection period occurred simultaneously in all study regions, and ran from April 11, 2011 through April 20, 2011. Over 95% of the survey forms were completed in the 6 days from April 11 through April 16.

During analysis, the various activities were grouped into categories that include schoolwork (homework and classes), eating, resting, play, and working. The reported number of hours per day spent in each category was calculated.

A follow-up survey will be conducted in September 2012 in order to collect parallel information on all of these variables post-intervention. More detail on the planned follow-up is provided in the final section of this report.

The present baseline report is limited to describing the study's initial conditions and examining the randomization process. No program impacts are examined at this stage, since the baseline survey was conducted before the start of program services.

III SAMPLE

3.1 POWER ANALYSIS

A power analysis was used to determine the sample size required to detect effects of the program in the context of the impact evaluation using standard statistical techniques, as described below. Two major categories of effects were considered: educational participation and labor participation for children in the study.

Educational participation was measured by survey items asking whether the child was currently in school or had gone to school during the past year and if not, whether the child had ever gone to school. Additional educational participation questions that address enrollment and attendance in more detail will be asked during the follow-up survey. Educational participation rates will be compared for children in the intervention and the control groups after the follow-up survey is completed.

Power analysis for the educational outcomes was done assuming a chi-squared analysis of a two-by-two contingency table. This is a robust approach that does not depend upon strong distributional assumptions for the analysis variables.

It was felt that educational participation among children in the impact evaluation was likely to be high, since the intervention was school-based. The intervention was being carried out among children in communities with access to a functioning school as a prerequisite to being able to provide the extended hours program; enrollment rates were therefore expected to be high. As children age, they become somewhat more likely to drop out of school, though we did not anticipate high dropout rates among the study population. If we assume that school participation is 99% among children in the intervention group and 90% among children in the control group, this translates to an effect size w of approximately 0.197. Using an alpha level of 0.05 and a typical power level of 0.80, the required sample size is approximately 200 children in total (including both the intervention and the control groups).

Labor participation was measured using the time record instrument, which measures the allocation of time of children in the study, their siblings, and their parents. The time record broke a typical week up into Saturdays, Sundays, and weekdays (Mondays through Fridays). Weekly hours of work could then be calculated by adding up those hours of a typical week that were devoted to activities such as paid work, work in family business or on family farms, and household chores. Average reported hours of work per week were compared for children in the intervention and control groups. Power analysis for this comparison was done assuming use of an F-test in the context of a linear regression framework.

Assuming that the standard deviation of the average number of hours of work per week is 15, a reduction of 6 hours per week on average in the intervention versus the control group would correspond to an effect size of 0.4.⁴ Assuming an alpha of 0.05 and power of 0.80, the required sample size is 200 children, assuming no attrition.

⁴ The assumption of a standard deviation of 15 is loosely based upon an informal review of available data from various child labor surveys, such as the 2007 Encuesta Nacional de Trabajo Infantil in Peru (International Labor Organization and Instituto Nacional de Estadística e Información).

The minimum of 200 participants was set as a recruitment goal for each major segment (urban and rural) of the impact evaluation, of which 100 would be assigned to the intervention and 100 to the control group in each setting.

3.2 RECRUITMENT

Recruitment and site selection were carried out by project staff using a procedure that was agreed upon with the research team. The procedure mirrored the one normally used by project staff in identifying which communities to target for the ÑPKCW program, except that we specifically targeted oversubscribed communities. (Please see last bullet point below).

Site selection was conducted as follows:

Pre-selected communities were those with high incidence of child labor or with children at risk of child labor. These criteria were met when—

- A community developed a productive activity that frequently used child labor (e.g., farms, poultry, or brickmaking);
- Children or adolescents were left alone for long periods because of established patterns of parental involvement in migrant labor;
- Poverty rates were higher than in other communities;
- There were cultural practices that involved children in risky work (such as brickmaking); and
- There were more children than could be accommodated within available classroom facilities.

The oversubscription condition, addressed in the last bullet point above, was a prerequisite to enrolling a community in the impact evaluation so that children could be randomized into intervention and control groups. The ÑPKCW program normally includes some oversubscribed communities in its projects, and we took advantage of this situation to introduce randomization rather than have project staff choose beneficiaries as they normally would.

Groups of children who were currently participating or at risk of participating in child labor were then identified within the selected communities. This was done with the help of local teachers in the classroom. In some cases, teachers applied a version of the time record to confirm children's work hours. In other cases, having a sibling who was working was considered a risk factor for becoming involved in child labor. A total of between 50 and 60 children were identified in each community, depending on the number of children who met the eligibility requirements.

Within selected sites, recruitment was done by holding an initial meeting of parents of preselected children to explain the research process in detail. During this meeting, parents were informed that because the program was oversubscribed, a lottery would be used to assign children to the program. Parents who consented to have their children participate signed an informed consent form.

3.3 RANDOMIZATION

3.3.1 Sequence Generation

A block randomization process was used. Blocking is a sampling technique that helps to ensure balance across variables that are likely to influence outcomes independently of the intervention being studied. In the post-intervention analysis of outcomes, statistical controls were introduced to account for the blocking; this should improve the precision of impact estimates. Blocks consisted of 14 school catchment areas, corresponding to the communities in which children were recruited for the study. Particularly in the rural program, communities tended to be geographically dispersed and the schools represented an important community anchor. In some cases, multiple hamlets or clusters of homes constituted a community, together with the school to which they were attached. As described above, a group informed consent process was conducted by project staff to explain the purpose of the research and to obtain consent from parents for their children to participate in the lottery to assign services and to participate in surveys and other data collection activities. After informed consent was obtained, a lottery was used to assign an equal number of children to control and intervention groups within groupings of age and sex. The age categories included children ages 5 to 13 and ages 14 to 17.

Separate lotteries for each age group took place, corresponding to the age cut-offs for the child and adolescent programs, respectively. The lottery took place at an event attended by school boards and school authorities. The procedure for the lottery was as follows:

1. Write down the names of 50 to 60 preselected children on slips of paper, one name per slip.
2. Put the papers in a box. Mix the slips thoroughly.
3. Remove half of the names and read publicly.
4. Write down each name in a report to be signed by those present.
5. The names remaining in the box are assigned to the control group.

The lottery for random assignment to intervention groups was conducted by DyA project staff.

3.3.2 Allocation Concealment

The group informed consent process was carried out and immediately followed by the public lottery for randomization. These activities were carried out in an open forum before conducting the baseline survey. Though it would have been ideal to randomize participants after the baseline survey, randomization preceded data collection because project staff felt strongly that it was necessary to do it this way in order to gain community members' support and approval. Therefore, it was not possible to conceal intervention assignments from participating children or their families. This might have created a threat to internal validity if participants' knowledge of intervention assignments influenced their outcomes. In addition, the Extended Hours program itself could not be concealed from participants or other members of the community, so there was no way in principal to conceal participants' intervention assignments during the course of program implementation.

3.4 BLINDING

Enumerators were not told the intervention assignments while conducting the baseline survey. However, it was likely that survey respondents did know the intervention assignments of the children in their households and might have disclosed those assignments inadvertently to the enumerators.

3.5 STATISTICAL METHODS

For this baseline report, we present descriptive statistics for the main variables characterizing the intervention and control groups, including gender, and hours spent on schoolwork, homework, agricultural work, and other paid work. Rural and urban schools were analyzed separately. We do not emphasize tests of statistical significance in the present report. Since the groups were randomly assigned, all differences in the groups would be due to chance. Significance tests are designed to assess the likelihood that differences across groups are due to chance, so they are redundant in this case.⁵ However, at OCFT's request we are including significance tests to help the reader identify differences between the intervention and control groups that are unexpectedly large relative to sample variability.

⁵ Pocock, S. J., Assmann, S. E., Enos, L. E., & Kasten, L. E. (2002). Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Statistics in Medicine*, 21(19), 2917–2930.

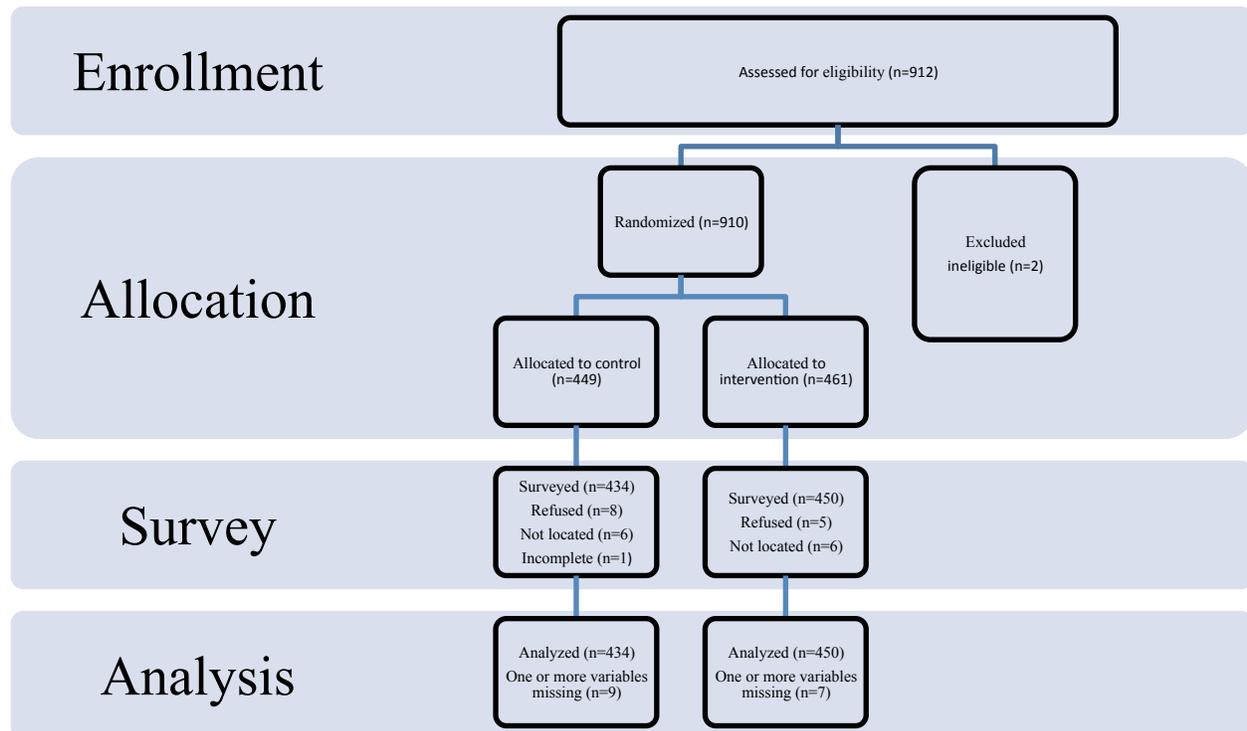
IV RESULTS

4.1 PARTICIPANT FLOW

The flow of participants through the various stages of the study is shown in Figures 1, 2, and 3. Figure 1 presents the overall flow of participants, Figure 2 presents the flow of rural participants, and Figure 3 presents the flow of urban participants.

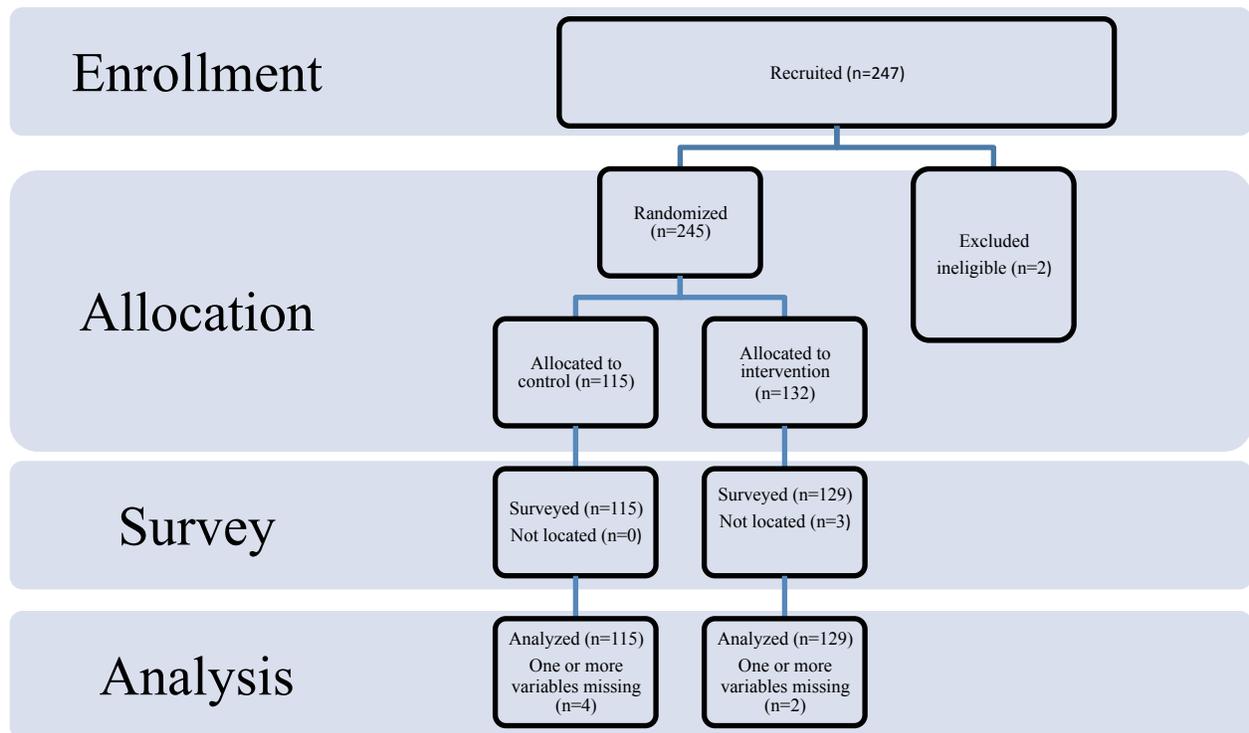
Overall, 912 children were assessed for eligibility and 910 were subsequently randomized (Figure 1). Of the 461 students allocated to the intervention group, 450 were surveyed and had data included in the analyses presented in this report. Some children's records were missing data on demographic characteristics or on time allocation variables because of enumerator errors or refusals; there were 7 children missing one or more of the variables used in the analyses that are presented. The control group was slightly smaller than the intervention group, with 449 students allocated through randomization, of which 434 were surveyed and included in the baseline analyses. Among those, there were 9 children missing data on one or more of the main analysis variables presented here.

Figure 1: Overall Sample Participant Flow



As discussed above, due to differences in demographics and in program structure, the rural and urban samples were considered separately in the analyses and can be treated as two separate but parallel studies. Figures 2 and 3 present participant flow charts for the rural and urban samples, respectively. The rural sample was smaller than the urban sample with 247 students assessed for eligibility (Figure 2). Of those students, 2 were excluded because they were older than 17; the remaining 245 were randomized, with 115 allocated to the control and 132 allocated to the intervention condition. All 115 in the control group were surveyed and had their data subsequently analyzed, while 129 of those in the intervention group were surveyed. Of those, two had one or more missing values among the variables analyzed. Three children in the intervention group could not be located.

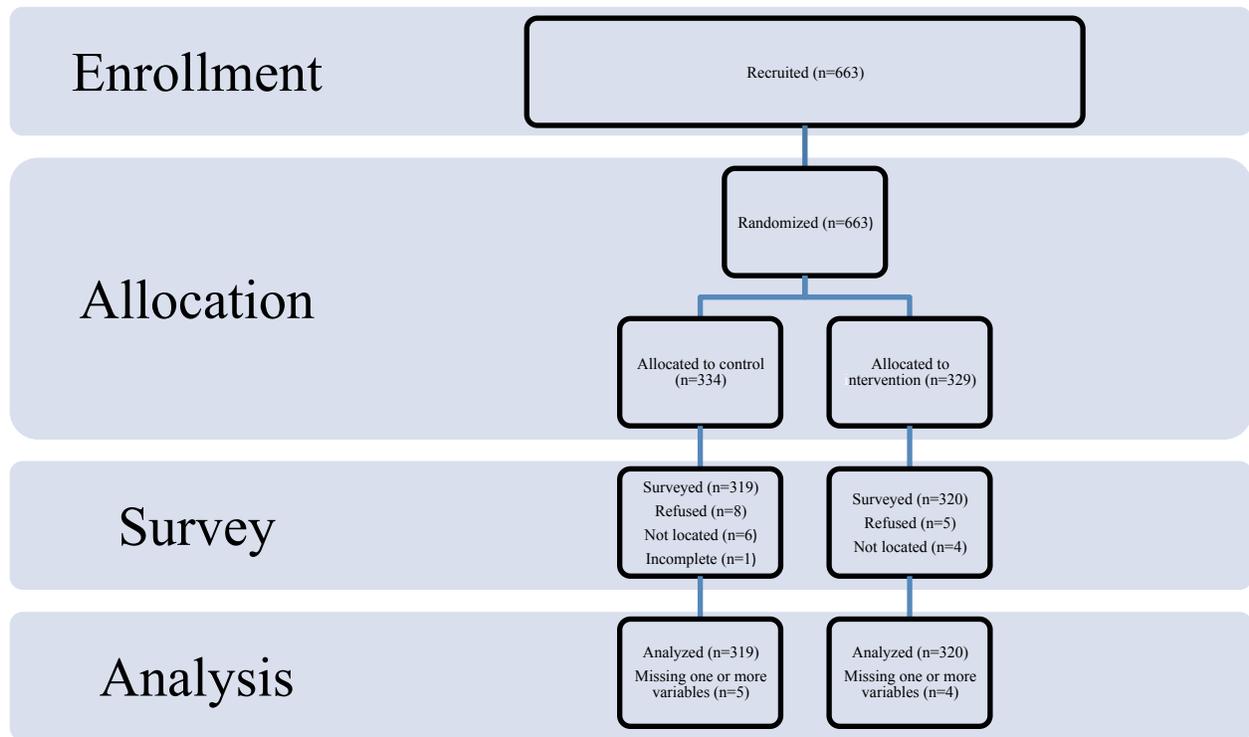
Figure 2: Rural Sample Participant Flow



The urban sample was larger than the rural sample, with 663 students assessed for eligibility (Figure 3). All 663 students were randomized, and 334 were assigned to the control while 329 were assigned to the intervention condition. Due to students refusing or not being located, 319 students were surveyed from the control and 320 were surveyed from the intervention group. There were some missing data for 5 of the students in the control group and 4 students in the intervention group.

The smaller rural sample reflects the fact that the rural ÑPKCW program beneficiaries were spread out among much smaller communities in a broad and sparsely populated geographic area as compared with the urban beneficiaries. There were fewer rural communities than urban communities that had enough eligible children to provide both an intervention and a control group.

Figure 3: Urban Sample Participant Flow



4.2 ANALYSIS

Descriptive statistics for students at the beginning of the study are presented in Tables 1 through 8 below. Baseline characteristics are presented for rural and urban students separately. Intervention and control groups were compared based on gender composition, participating school, and hours spent on schoolwork, homework, housework, agricultural work, paid work outside the home, and working in a family business. We use “housework,” encoded in the Spanish version of the time allocation module as *tareas domésticas*, to include “helping at home,” which respondents may describe as *ayuda en casa*. Thus, this category may capture such activities as cooking, doing laundry, carrying water, and farming activities conducted around the home, such as tending animals. Other types of “helping” may include working in a family business or helping parents in an outside job. The enumerators used their discretion in deciding which activities to place into each category. Many of the respondents did not consider helping to be work, as it is an expected part of childhood. However, helping constitutes a significant part of the time respondent children spend working and should be included in any summary of their work hours. Hours were compared for typical days during the work week, and for Saturdays and Sundays, separately. Sample sizes varied according to the number of valid responses collected for each variable. Therefore, sample sizes are reported in the appropriate cell for each comparison.

4.2.1 Rural Schools

Table 1 presents baseline characteristics of the rural intervention and control groups. There was a higher proportion of girls in the intervention group (64%) compared with the control group (51%). This difference was due to random sampling variability. The mean age was around 11 in both groups. Nearly all of the respondents were enrolled in school, and nearly all reported working.

The students were fairly evenly distributed across the five rural schools. The one exception was at School 5, where the control group had 11 students compared with 24 in the intervention group. This appears to be a case in which oversubscription did not occur as expected. The lottery was done using the 35 children who were found to be eligible, and the first 24 children selected were put into the intervention group in order to have enough children for a full class. In this particular school, the screening lists included 13 children who appeared to have been added to the control group without going through the screening process, which suggests that staff involved in enrollment may have tried to rebalance the block by adding children in a nonrandom way to the control group. Enumerators appropriately excluded those 13 and they were not counted in the participant flow of the impact evaluation.

The age distribution of participants is presented in Figure 4, and the distribution of average hours of work (including housework, paid work, agricultural work, and work in family businesses) is presented in Figure 5.

We found that rural students reported spending very little time during the typical work week on agricultural work, working in a family business, or doing other work for pay (Tables 2-5). They spent a bit more time on these activities during Saturdays and Sundays, though they still averaged only fractions of an hour per day. However, on Saturdays and Sundays students reported that they were spending about three hours per day on average on housework, which was more than they spent on homework (about two hours). Mondays through Fridays, students reported spending over an hour on housework per day. A calculated estimate of hours allocated to each of the activities per week is presented in Table 5. The same numbers are presented graphically in Figure 6, which compares the estimated cumulative time per week in each category for the intervention and control children, respectively.

Table 1: Baseline Characteristics of Students in Rural Schools

	Intervention			Control			Test Statistic*		
	Mean or Count	SD or Percent	n	Mean or Count	SD or Percent	n	χ^2 or t	df	p
Female	83	64%	129	58	51%	113	4.20	1	0.041
Age	11.2	2.11	128	11.2	2.18	113	0.07	233.3	0.942
Working	125	100%	128	102	99%	111	4.14	1	0.042
In school	128	100%	128	111	99%	112	2.27	1	0.132
School ID:									
5	24	69%	35	11	31%	35	4.36	4	0.363
6	28	47%	59	31	53%	59			
12	25	50%	50	25	50%	50			
13	27	53%	51	24	47%	51			
14	26	52%	50	24	48%	50			

*Test statistics are χ^2 for count data, or two-sample t-tests assuming unequal variances for means. P-values for t-tests are two-sided.

Figure 4: Age Distribution of Rural Respondents

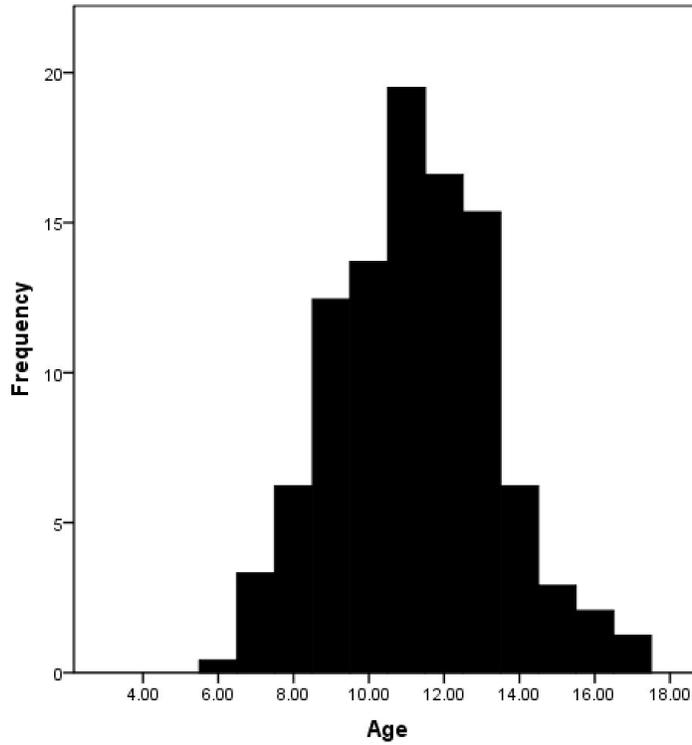


Figure 5: Distribution of Rural Respondents' Reported Weekly Hours Doing Any Work

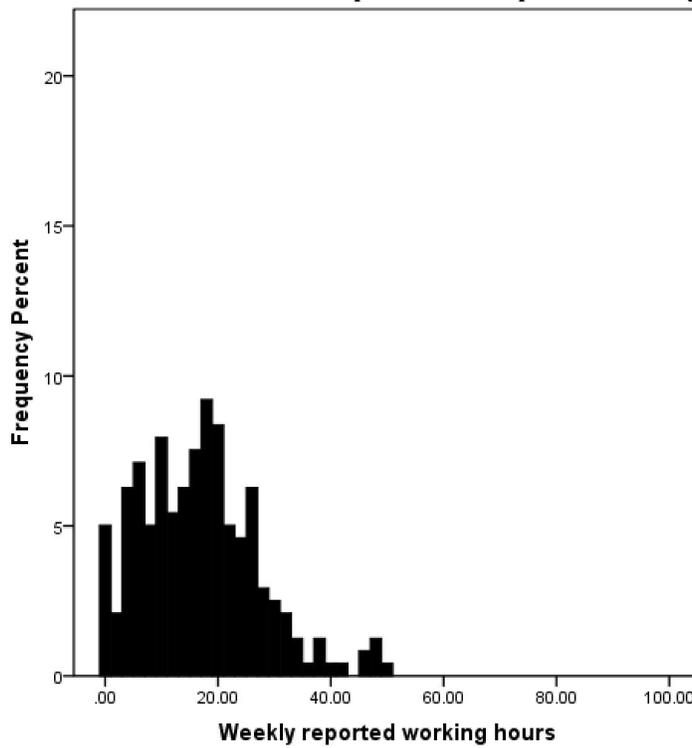


Table 2: Reported Time Allocation in Hours per Day for Rural Respondents, Mondays through Fridays

	Intervention n=128 Mean (SD)	Control n=111 Mean (SD)	t	df	p
Study in school	5.52 (1.50)	5.42 (1.40)	-0.5	235.6	0.623
Homework	2.19 (1.53)	2.01 (1.51)	-0.9	233.1	0.366
Paid work outside home	0.00 (0.00)	0.00 (0.00)	-	-	-
Agricultural work	0.16 (0.71)	0.08 (0.47)	-1.1	222.5	0.281
Family business, non-farm	0.05 (0.37)	0.03 (0.28)	-0.5	233.3	0.642
Housework	1.37 (1.33)	1.23 (1.46)	-0.7	224.9	0.465
Transportation	0.00 (0.00)	0.03 (0.16)	1.7	110.0	0.083
Play or recreation	1.24 (1.25)	1.56 (1.36)	1.9	225.7	0.064
Watching TV	0.20 (0.63)	0.30 (0.82)	1.1	205.2	0.286
Eating	3.05 (0.63)	3.10 (0.59)	0.6	235.8	0.574
Sleeping	10.09 (1.27)	10.14 (1.08)	0.3	236.9	0.786
Other	0.13 (0.50)	0.08 (0.38)	-0.8	233.6	0.445

T-tests are calculated for two samples, assuming unequal variances. Associated p-values are two-sided.

Table 3: Reported Time Allocation in Hours per Day for Rural Respondents, Saturdays

	Intervention n=128 Mean (SD)	Control n=111 Mean (SD)	t	df	p
Study in school	0.12 (0.65)	0.12 (0.53)	0.0	236.5	0.999
Homework	2.26 (1.67)	2.55 (1.85)	1.3	223.5	0.204
Paid work outside home	0.04 (0.44)	0.00 (0.00)	-1.0	127.0	0.319
Agricultural work	0.99 (1.65)	0.91 (1.76)	-0.4	227.4	0.710
Family business, non-farm	0.27 (1.03)	0.12 (0.74)	-1.3	228.8	0.197
Housework	3.70 (2.88)	3.23 (2.69)	-1.3	235.7	0.195
Transportation	0.03 (0.28)	0.07 (0.35)	1.0	209.7	0.324
Play or recreation	2.95 (2.54)	3.02 (2.39)	0.2	235.4	0.820
Watching TV	0.25 (0.88)	0.49 (1.34)	1.6	184.9	0.114
Eating	3.05 (0.66)	3.12 (0.71)	0.8	226.8	0.432
Sleeping	10.19 (1.37)	10.15 (1.29)	-0.2	235.3	0.843
Other	0.16 (0.78)	0.22 (0.80)	0.5	230.4	0.612

T-tests are calculated for two samples, assuming unequal variances. Associated p-values are two-sided.

Table 4: Reported Time Allocation in Hours per Day for Rural Respondents, Sundays

	Intervention n=128 Mean (SD)	Control n=111 Mean (SD)	t	df	p
Study in school	0.03 (0.28)	0.07 (0.40)	0.9	193.3	0.367
Homework	2.07 (1.69)	2.27 (1.75)	0.9	230.0	0.372
Paid work outside home	0.00 (0.00)	0.00 (0.00)	-	-	-
Agricultural work	0.84 (1.64)	0.55 (1.50)	-1.4	236.2	0.159
Family business, non-farm	0.19 (0.90)	0.07 (0.48)	-1.3	199.1	0.211
Housework	3.58 (2.86)	3.06 (2.91)	-1.4	230.9	0.171
Transportation	0.02 (0.15)	0.02 (0.19)	-0.2	210.0	0.810
Play or recreation	2.80 (2.57)	3.41 (2.94)	1.7	220.1	0.092
Watching TV	0.45 (0.97)	0.46 (1.07)	0.0	224.3	0.962
Eating	3.05 (0.66)	3.22 (0.71)	1.9	227.3	0.058
Sleeping	10.38 (1.43)	10.23 (1.47)	-0.8	230.1	0.426
Other	0.59 (1.50)	0.64 (1.53)	0.2	230.7	0.816

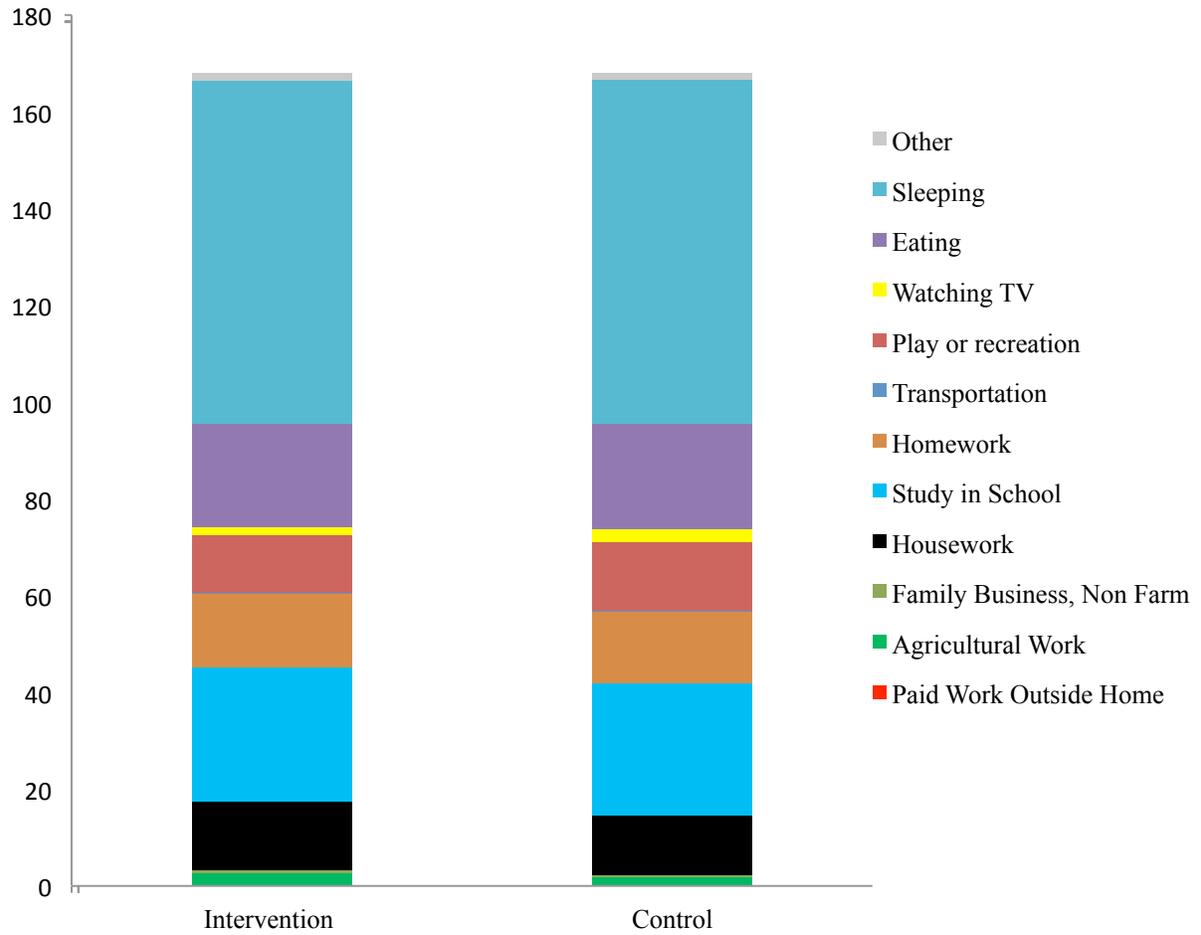
T-tests are calculated for two samples, assuming unequal variances. Associated p-values are two-sided.

Table 5: Estimated Weekly Time Allocation, Rural Respondents

	Intervention n=128 Mean (SD)	Control n=111 Mean (SD)
Study in school	27.73 (7.57)	27.31 (7.23)
Homework	15.27 (8.21)	14.86 (8.51)
Categories of work		
Paid work outside home	0.04 (0.44)	0.00 (- -)
Agricultural work	2.65 (5.17)	1.86 (3.73)
Family business, non-farm	0.69 (2.65)	0.32 (1.82)
Housework	14.12 (9.55)	12.47 (11.00)
Transportation	0.05 (0.34)	0.23 (1.07)
Play or recreation	11.95 (8.85)	14.22 (9.53)
Watching TV	1.68 (3.85)	2.43 (5.65)
Eating	21.37 (3.78)	21.83 (3.56)
Sleeping	71.03 (8.20)	71.05 (6.93)
Other	1.38 (3.69)	1.26 (2.73)

Estimated weekly hours are calculated from sum of estimated hours for Saturday and Sunday and 5 times estimated hours for a typical Monday-Friday weekday.

Figure 6. Estimated Weekly Time Allocation, Rural Respondents



4.2.2 Urban Schools

Table 6 presents baseline characteristics of the urban intervention and control groups. There was a balanced age and gender distribution in the intervention and control groups. Mean age was around 10. Nearly all children reported being enrolled in school, and most reported working (83–86%). As with the rural schools, the urban students were fairly evenly distributed across the nine urban schools. Again, there was one exception to this pattern. The exception was School 9, where the control group had 19 students compared with 29 in the intervention group. As in School 5, it appears that there was not an oversubscription situation in this community as expected. Project staff randomly selected enough children to fill a classroom and then stopped, leaving a larger number in the intervention than the control group. This results in a slightly unbalanced block but should not undermine the overall results of the impact evaluation.

The age distribution of participants is presented in Figure 7, and the distribution of average hours of work (including housework, paid work, agricultural work, and work in family businesses) is presented in Figure 8.

The urban students reported spending more time working during the week and on the weekends than their rural counterparts (Tables 7-10). In particular, they were working more in family-owned businesses than the rural students. Monday through Friday, the urban students were spending about three-quarters of an hour on average working in family-owned businesses. This average increased to about an hour on Saturdays. On Sundays, they spent over half an hour on average working for their families' businesses, so the time they spent working remained fairly consistent across the week. On top of that, they report doing about an hour of housework during each weekday and about two hours a day on Saturdays and Sundays. Perhaps as a result, the urban students spent less time on homework on the weekends compared with the rural students. Whereas the rural students reported spending about 2 hours on homework on a typical Saturday or Sunday, the urban students reported spending about 1 hour.

Estimated cumulative total time in each category for a week is presented in Table 10 and graphically in Figure 9.

Table 6: Baseline Characteristics of Urban Respondents

	Intervention			Control			Test Statistic*		
	Mean or Count	SD or Percent	n	Mean or Count	SD or Percent	n	χ^2 or t	df	p
Female	161	51%	318	154	49%	317	0.27	1	0.606
Age	10.1	2.01	318	10.1	2.12	317	0.04	632.8	0.968
Working	273	86%	316	262	83%	314	1.07	1	0.300
In school	316	100%	316	313	99%	316	3.01	1	0.083
School ID									
2	30	40%	75	45	60%	75	5.30	8	0.724
3	29	50%	58	29	50%	58			
4	29	50%	58	29	50%	58			
7	79	50%	157	78	50%	157			
8	39	50%	78	39	50%	78			
9	29	60%	48	19	40%	48			
10	26	53%	49	23	47%	49			
11	29	51%	57	28	49%	57			
15	30	51%	59	29	49%	59			

*Test statistics are χ^2 for count data, or two-sample t-tests assuming unequal variances for means. P-values for t-tests are two-sided.

Figure 7: Age Distribution of Urban Respondents

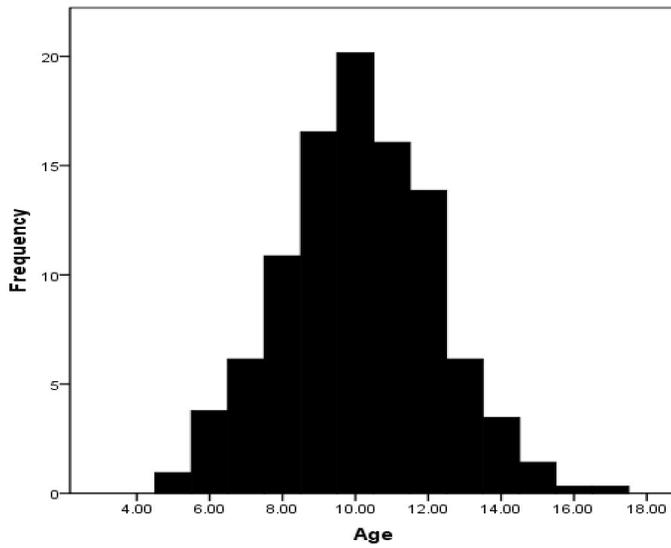


Figure 8: Distribution of Urban Respondents' Reported Weekly Hours Doing Any Work

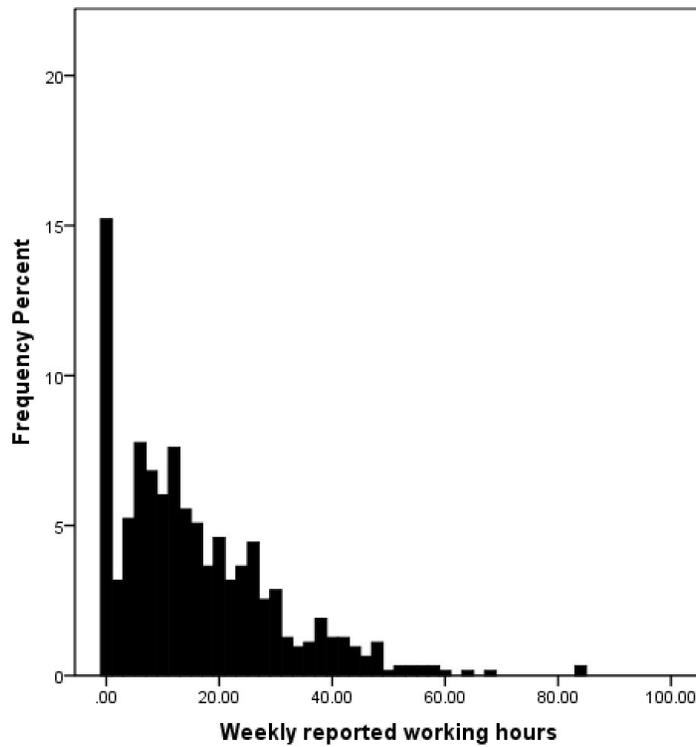


Table 7: Reported Time Allocation in Hours per Day for Urban Respondents, Mondays through Fridays

	Intervention n=316	Control n=314			
	Mean (SD)	Mean (SD)	t	df	p
Study in school	4.68 (1.00)	4.69 (1.23)	0.1	602.6	0.904
Homework	2.48 (1.54)	2.51 (1.49)	0.2	627.6	0.814
Paid work outside home	0.15 (0.91)	0.06 (0.48)	-1.4	474.4	0.159
Agricultural work	0.08 (0.48)	0.05 (0.34)	-1.0	565.0	0.341
Family business, non-farm	0.44 (1.26)	0.51 (1.37)	0.6	622.7	0.527
Housework	1.10 (1.51)	1.11 (1.68)	0.0	619.5	0.976
Transportation	0.45 (0.78)	0.41 (0.86)	-0.5	620.7	0.624
Play or recreation	1.00 (1.31)	0.86 (1.18)	-1.4	622.1	0.169
Watching TV	0.86 (1.13)	1.01 (1.19)	1.6	625.9	0.117
Eating	2.89 (1.00)	2.95 (1.09)	0.8	621.6	0.450
Sleeping	9.57 (1.52)	9.46 (1.61)	-0.9	625.4	0.359
Other	0.28 (0.92)	0.36 (1.08)	1.0	611.1	0.308

T-tests are calculated for two samples, assuming unequal variances. Associated p-values are two-sided.

Table 8: Reported Time Allocation in Hours per Day for Urban Respondents, Saturdays

	Intervention n=316	Control n=314			
	Mean (SD)	Mean (SD)	t	df	p
Study in school	0.11 (0.70)	0.08 (0.60)	-0.5	612.3	0.590
Homework	1.56 (1.77)	1.51 (1.82)	-0.4	627.1	0.724
Paid work outside home	0.34 (1.62)	0.13 (1.00)	-1.9	524.9	0.052
Agricultural work	0.20 (1.14)	0.14 (0.91)	-0.7	599.0	0.470
Family business, non-farm	1.12 (2.85)	1.04 (2.52)	-0.4	619.6	0.724
Housework	2.21 (2.48)	2.44 (2.75)	1.1	621.0	0.257
Transportation	0.10 (0.53)	0.09 (0.49)	-0.4	624.4	0.709
Play or recreation	2.66 (2.45)	2.72 (2.63)	0.3	624.2	0.774
Watching TV	1.84 (2.05)	1.97 (1.95)	0.8	626.7	0.417
Eating	2.91 (0.95)	2.90 (0.85)	-0.1	620.8	0.888
Sleeping	10.11 (1.65)	9.89 (1.53)	-1.8	624.9	0.076
Other	0.83 (2.14)	1.08 (2.38)	1.4	620.1	0.160

T-tests are calculated for two samples, assuming unequal variances. Associated p-values are two-sided.

Table 9: Reported Time Allocation in Hours per Day for Urban Respondents, Sundays

	Intervention n=316 Mean (SD)	Control n=314 Mean (SD)	t	df	p
Study in school	0.03 (0.35)	0.03 (0.43)	-0.1	602.1	0.924
Homework	1.22 (1.65)	1.20 (1.67)	-0.1	627.8	0.894
Paid work outside home	0.24 (1.57)	0.05 (0.58)	-2.0	400.3	0.045
Agricultural work	0.16 (0.89)	0.08 (0.57)	-1.3	534.0	0.188
Family business, non-farm	0.59 (2.23)	0.86 (2.60)	1.4	612.3	0.160
Housework	1.78 (2.32)	1.75 (2.46)	-0.2	625.3	0.875
Transportation	0.11 (0.55)	0.16 (0.78)	1.0	565.9	0.336
Play or recreation	3.32 (3.02)	3.22 (2.92)	-0.4	627.5	0.664
Watching TV	2.06 (2.20)	2.15 (2.22)	0.5	627.8	0.586
Eating	2.86 (1.09)	2.92 (1.08)	0.6	628.0	0.537
Sleeping	10.22 (1.78)	10.10 (1.65)	-0.9	624.9	0.357
Other	1.39 (2.66)	1.46 (2.58)	0.3	627.7	0.751

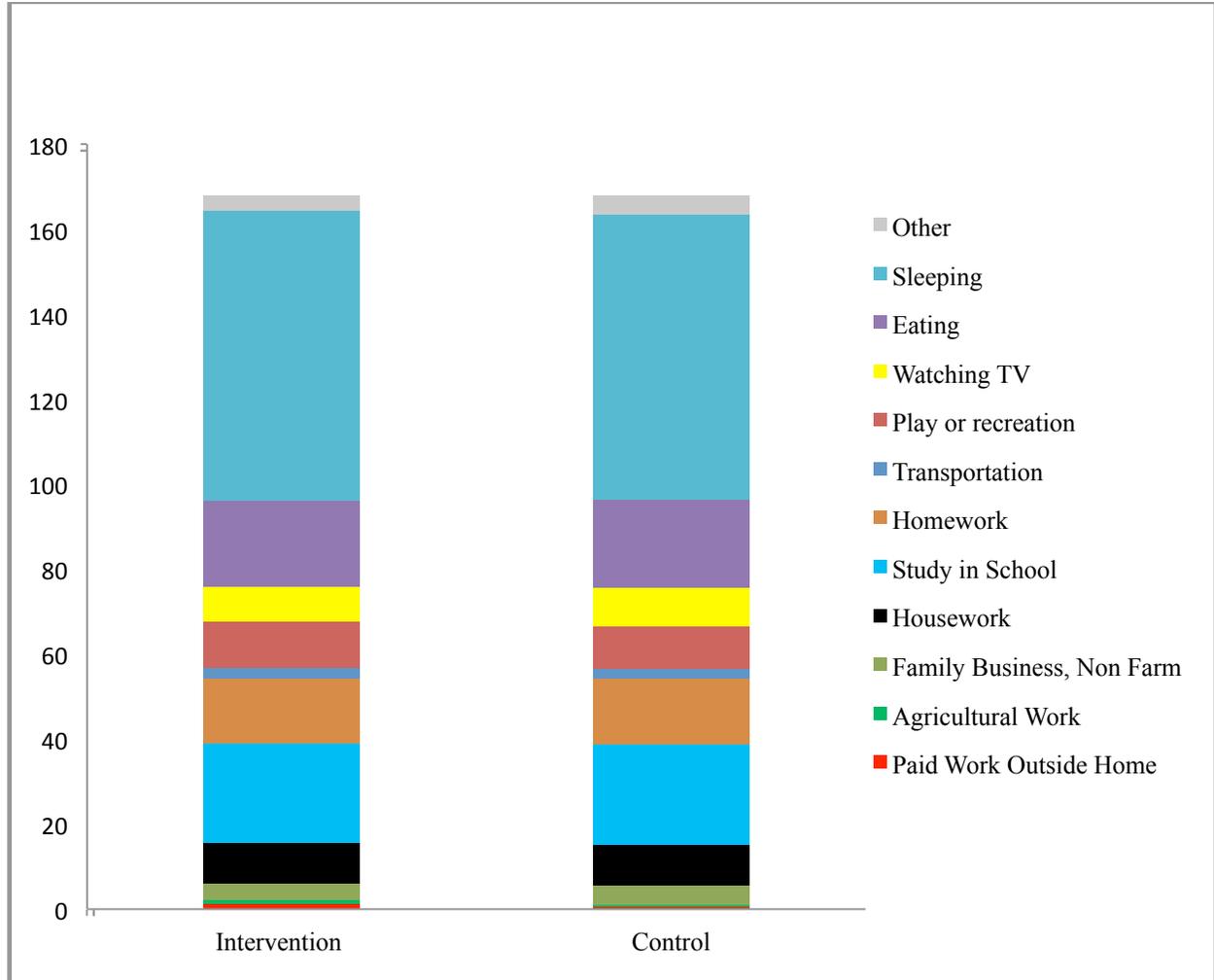
T-tests are calculated for two samples, assuming unequal variances. Associated p-values are two-sided.

Table 10: Estimated Weekly Time Allocation, Urban Respondents

	Intervention n=316 Mean (SD)	Control n=314 Mean (SD)
Study in school	23.56 (5.10)	23.58 (6.20)
Homework	15.20 (8.81)	15.28 (8.56)
Categories of work		
Paid work outside home	1.31 (6.78)	0.50 (3.19)
Agricultural work	0.77 (3.37)	0.48 (2.62)
Family business, non-farm	3.92 (9.75)	4.45 (10.09)
Housework	9.51 (9.83)	9.74 (11.64)
Transportation	2.44 (3.96)	2.32 (4.44)
Play or recreation	10.98 (9.37)	10.25 (8.68)
Watching TV	8.22 (7.47)	9.17 (7.83)
Eating	20.21 (6.13)	20.57 (6.54)
Sleeping	68.20 (10.02)	67.28 (9.90)
Other	3.61 (7.16)	4.34 (8.28)

Estimated weekly hours are calculated from sum of estimated hours for Saturday and Sunday and 5 times estimated hours for a typical Monday-Friday weekday.

Figure 9. Estimated Weekly Time Allocation, Urban Respondents



V DISCUSSION

5.1 INTERPRETATION

Although there was some evidence that some children were inadvertently added to the control group in two communities, as described above, overall the randomization process appears to have been executed as expected. This is reflected in the fact that the intervention and control groups appear to be well-balanced on the variables of interest. In the rural group, a higher percentage of girls were in the intervention group (64%) than in the control group (51%); however, the gender distribution in the urban intervention and control groups was almost equivalent. As noted above, this difference was due to random sampling variability. The difference could have been prevented by using a blocked randomization strategy, but this was not done for practical reasons. In the original sampling plan that was developed jointly with DyA and USDOL staff in early 2011, ICF suggested that the randomization process should be stratified by age and sex. This would entail conducting a separate lottery for each combination of sex and age category. In effect, the lottery would be done separately in each community for young girls, older girls, young boys, and older boys; this would guarantee a more even distribution across the intervention and control groups on both gender and age. But it emerged that this strategy was not viable in the context of the community meetings that DyA held to administer the informed consent procedure and lottery. The difference was not likely to be a material problem for the evaluation, as we included statistical controls for gender in all of the outcome models.

Students were evenly distributed across the participating schools. Attrition through the stages of study implementation to this point had been minimal.

We found that at baseline the time that students spent on various activities was equivalent between the intervention and control groups in the rural and urban samples. Students spent slightly more time in school, doing homework, and doing housework than in doing agricultural work, paid work outside the home, or working in a family business.

Students were not reporting working extremely long hours at any time during the week. In particular, during weekdays children were only reporting spending a few hours on average in various work activities. Some children may have been exposed to hazards at work, which would also represent child labor regardless of the number of hours of work. In addition, current school participation appeared to be extremely high. School attendance itself limits the number of hours that a child has available to work. These figures have been compared with monitoring data collected by DyA on its beneficiary population, and were largely in agreement. This might have made it difficult to demonstrate program effectiveness, since the program is premised in part on displacing work hours during the week with additional school activities.

5.2 GENERALIZABILITY

The particular forms and severity of child labor around the world are very context-specific, and the structure of the ÑPKCW program was tailored specifically to the Bolivian political, economic, and social situation. Therefore, results of this evaluation may not generalize directly to other programs in Latin America or to other regions throughout the world. On the other hand, there are a number of common threads running through many of the programs that OCFT funds, including interventions to strengthen

and enrich existing educational resources, awareness raising, policy advocacy, and livelihood services. As a greater number of rigorous impact evaluations are undertaken, it will become easier to assess the extent to which impacts in any given program are globally representative. As more evaluation information becomes available, the lessons learned in one context should be useful in designing and managing programs in other contexts.

Generalizability was also limited by the fact that only oversubscribed communities were included in the study. The project also provides services to smaller communities that are not oversubscribed—communities in which there are no more eligible children than can be accommodated in a single classroom. Those communities may be very different structurally and culturally from the communities included in the study. By nature, those smaller communities necessarily contribute a smaller proportion of the total population of program beneficiaries. Follow-up analyses may include examination of project records to determine the extent to which the totality of these differences is likely to restrict generalizability. Our supposition is that the impact will be minor.

5.3 OTHER LIMITATIONS

Though generally regarded as the most rigorous evaluation design, an RCT approach does present a few weaknesses. One important limitation, which has been touched in the Data Collection section above, concerns the relationship between the individual-level and group-level interventions. The impact evaluation was designed to estimate the impact of the Extended Hours program in the presence of a community-level intervention that accompanies it. It would not have been possible to estimate the impact of both the community-level and individual-level interventions together within an individual randomization design like the one used here, because children assigned to the control group were still exposed to the community-level intervention within their villages. Such a comparison would have required a place-based randomization strategy, which was not felt to be feasible within the constraints of resource availability, project structure, and community politics. DyA had to do extensive outreach with the study communities in order to gain their acceptance of the impact evaluation. This outreach work included awareness-raising about the problem of child labor and household economic strategies. It is possible that this initial community-level work was the most important element of the Extended Hours program. Factoring it out of the impact estimate may, therefore, have made the program look less effective than it actually was.

The individual randomization strategy that was chosen does have some advantages over a cluster randomization approach. Notably, individual randomization tends to be more efficient in terms of the sample size needed for a given level of statistical power. However, it carries some unique risks, including the risk of “contamination” or “spillover” effects. Because children in the intervention and control groups were living in the same villages, attending the same schools and in some cases were members of the same families, the impact of the Extended Hours program may not be well contained to the intervention group alone. In particular, the Extended Hours (intervention) group likely interacted extensively with the control group and shared many of their experiences both inside and outside of school. Anecdotal reports from the field suggested that children in the control group had been curious about the activities going on in the Extended Hours program throughout the first year of services. Because most schools in Bolivia use open-plan buildings, it may have been possible for children in the control group to observe some of the activities happening inside the classroom and perhaps even to participate unofficially in sports, arts, or other activities that happened outside the classroom. Children in the intervention group may also have talked about what they were doing in school with their friends and siblings, which may have provided another avenue for information and ideas from the intervention to cross over into the control group.

A separate issue, also raised above in the Allocation Concealment subsection, was that randomization preceded data collection, which made it impossible to blind participants and data collectors to intervention assignments. This may have affected the validity of baseline data to the extent that knowledge of intervention assignments changed participants' behavior or survey responses.

With these constraints in mind, it is important to recognize that the impact estimates may be downwardly biased relative to the true impact of the program. This is a contingency that will be explored in the final report. Findings will still be valuable in describing the participating children's experience and in framing expectations for future impact studies.

5.4 NEXT STEPS

A follow-up survey will be conducted about a year and a half after the baseline to collect information on children's educational and work outcomes following the intervention. Specifically, ICF plans to pilot a follow-up data collection instrument and train the local enumerators during the week of September 13-24, 2012, and collect follow-up data in Bolivia from September 25 through October 8, 2012. The follow-up survey will include most of the items that were included in the baseline survey and will add additional content to address children's exposure to hazards at work as well as more fine-grained information on school participation and persistence. Upon finalization of the follow-up survey, appropriate multivariate analyses will be conducted to show whether children in the Extended Hours group are more likely to participate in school and do fewer hours of work than children in the control group. Additional analyses will look at spillover effects among the study children's siblings and parents.