EVALUATION

Impact Evaluation of the Western Highlands Integrated Program in Guatemala

Midline Report

USAID
FROM THE AMERICAN PEOPLE

MEASURE Evaluation

EVALUATION

Impact Evaluation of the Western Highlands Integrated Program in Guatemala

Midline Report

Gustavo Angeles, PhD (team lead); Aimee Benson, MA; Paul Brodish, PhD; Kristen Brugh,PhD; Roberto Molina, MSc; Martin Romero, PhD; Tory M. Taylor, MPH; Emily Weaver, PhD; Jose Urquieta, PhD

June 2018

MEASURE Evaluation
University of North Carolina at Chapel Hill
123 W. Franklin Street, Suite 330
Chapel Hill, NC 27516 USA
Phone: +1 919-445-9350 | measure@unc.edu
www.measureevaluation.org

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. TRE-18-010;
ISBN: 978-1-64232-049-7

USAID
FROM THE AMERICAN PEOPLE

MEASURE
Evaluation

ABSTRACT

The main objectives of the Western Highlands Integrated Program (WHIP), which was funded by the United States Agency for International Development (USAID)/Guatemala, were to reduce poverty and chronic malnutrition among children in 30 priority municipalities. The program combined the Rural Value Chains Project (RVCP) with a health and nutrition program. The primary evaluation question focused on the effects of the WHIP on key indicators at the population level in the program's zone of influence (ZOI). The secondary evaluation questions focused on understanding the impacts of the integrated (RVCP plus health/nutrition) program and the health/nutrition program by itself; the relative effectiveness of the integrated program compared to the health intervention alone; and the presence of spillover effects from RVCP direct beneficiaries to nonassociation members' households in RVCP areas (known as RVCP indirect beneficiaries). The evaluation used a prospective, quasi-experimental study design with a matched comparison group, and implemented a difference-in-differences analysis controlling for household-level fixed effects using pooled baseline and midline data from a panel of households. Results of the midline impact evaluation indicated that although there was no statistically significant program impact on household consumption, poverty, or hunger, these indicators were moving in the expected direction consistent with the program's theory of change. Mixed results in the time trends for infant and young child feeding practices, nutritional status, and the decreasing use of reproductive and maternal health services suggest that the cessation of the health and nutrition program in 2013/2014 may have had detrimental effects on these indicators in the ZOI areas.

ACKNOWLEDGMENTS

The authors wish to acknowledge the efforts of numerous people who supported this impact evaluation analysis and report. We are grateful to the United States Agency for International Development (USAID)/ Guatemala for entrusting us with this activity and providing the technical support and financial resources that made it possible. In particular, we thank Joshua Smith, Ángel López, Josefina Martínez, and Baudilio López, who currently work at the USAID Mission, and Tom DiVincenzo, Judith Timyan, David Delgado, and Ashley Frost who were at the Mission during the first years of the evaluation. We also thank Lisa Maniscalco, Kristen Wares, and Elisa Ballard, our colleagues at USAID/Washington, DC, and Anne Swindale and Lindsey Anna, of the USAID Bureau for Food Security, who provided important and sound technical and management guidance. We acknowledge the contributions of Juan Mendez, Pablo Toledo, and Pedro Pineda, of DevTech Systems, Inc., who provided data and documentation from the midline survey for use in this evaluation. Erin Luben, Bates Buckner, and Peter Lance, all of the USAID-funded MEASURE Evaluation, offered valuable recommendations on key technical and operational issues. We thank MEASURE Evaluation's knowledge management team for editorial and production services.

Cover: Students run to wash their dishes after having lunch at school in the town of El Quetzal, department of San Marcos, Guatemala. Photo: © 2013 Tory M. Taylor, MEASURE Evaluation

CONTENTS

FIGURES

TABLES

ABBREVIATIONS

BL	baseline
BMI	body mass index
CPI	Consumer Price Index
DID	difference-in-differences
EMEPAO	Encuesta de Monitoreo y Evaluación del Programa del Altiplano Occidental (Western Highlands Integrated Program Baseline Evaluation Survey)
ENCOVI	Encuesta Nacional de Condiciones de Vida (Living Standards Measurement Survey)
ENSMI	Encuesta Nacional de Salud Materno Infantil (National Maternal and Child Health Survey)
FE	fixed effects
FTF	Feed the Future
HH	household
HHS	Household Hunger Scale
HO	health only
MAD	Minimum Acceptable Diet
ML	midline
MSPAS	Ministerio de Salud Pública y Asistencia Social (Ministry of Public Health and Social Assistance)
NGO	nongovernmental organization
ORS	oral rehydration salts
pc	per capita
PEC	Programa de Extensión de Cobertura
PPP	purchasing power parity
PSM	propensity score matching
Q	Quetzales
RVCP	Rural Value Chains Project
USAID	United States Agency for International Development
USD	United States dollar
WHIP	Western Highlands Integrated Program
WRA	women of reproductive age
ZOI	zone of influence

EXECUTIVE SUMMARY

Since 2012, USAID in Guatemala has overseen the development of a set of coordinated programs designed to improve health and economic well-being in 30 priority municipalities located in the country's historically underserved Western Highlands region. This group of programs is known as the Western Highlands Integrated Program (WHIP). USAID has facilitated several monitoring and evaluation activities since the program's inception. A baseline household survey was conducted from July through November 2013, forming the groundwork for a prospective, quasi-experimental impact evaluation of two flagship WHIP initiatives: the Rural Value Chains Project (RVCP), which provides technical and resource support to members of smallholder producers' groups; and the health and nutrition program, a set of community-level activities offering health and nutrition education and working to improve access to high-quality health services for women and children.

This document presents the results of the midline impact evaluation of the program at its approximate midpoint (2015) on key indicators: household expenditures, poverty, food security, dietary and feeding practices, nutrition status, fertility, family planning, and use of maternal health services.

Intervention Components

The WHIP's RVCP and the health and nutrition interventions were the focus of this impact evaluation. The RVCP had two main components: technical assistance and training provided to agricultural and handicraft producers' associations, and nutritional education and information communication. There were two RVCP study populations of interest: the first was RVCP direct beneficiaries, which comprised approximately four percent of households in the 30 priority municipalities. These 30 municipalities were also termed the WHIP's zone of influence (ZOI). The second was RVCP indirect beneficiaries, consisting of households that did not have an RVCP association member but were in communities where producers' association members resided (approximately 28 percent of the total population in the ZOI).

The health and nutrition program was designed to expand the quality and availability of health and nutrition services in the public sector. Unfortunately, the health and nutrition program was not implemented as planned. The Extension of Coverage Program (Programa de Extensión de Cobertura or PEC) started in 1997 as a response to the health provision goals of the Guatemala Peace Accords signed in 1996. The Ministry of Public Health and Social Assistance (MSPAS or Ministerio de Salud Pública y Asistencia Social) contracted nongovernmental organizations (NGOs) as providers of health services to extend health coverage to primarily indigenous and impoverished rural communities without access to a health provider. In November 2013, Guatemala's Congress passed an amendment to the budget law that would halt all MSPAS funding to PEC NGOs, and health service delivery for rural populations was significantly reduced. This situation worsened in the absence of a suitable strategy to replace these NGOs. Lack of coverage exacerbated the effects of a health system hobbled by pervasive stockouts of essential medicines, family planning supplies, and vaccines. Insufficient human resources and reliance on unpaid MSPAS field staff affected service delivery, especially for primary care. Continuous changes in the MSPAS administration led to the delay in the implementation of a new model of care designed to provide appropriate health services to an estimated one-third of the total population of Guatemala.

Evaluation Objectives and Questions

The main objective of the WHIP impact evaluation was to estimate changes in key indicators that were attributable to the WHIP interventions under study. To this end, the primary evaluation question was:

1. What are the effects (impact) of the WHIP on key indicators at the population level in the WHIP ZOI?

The primary evaluation question required additional disaggregation because each program component under study covered different population groups in the ZOI. The target population of the health and nutrition program included the entire ZOI, whereas the RVCP targeted only members of households with a producer participating in the RVCP (direct beneficiaries). The direct beneficiaries made up four percent of all households in the ZOI at baseline. As noted, RVCP activities were expected to have indirect effects on nonparticipating households (i.e., indirect beneficiaries) who resided in the same communities where RVCP direct beneficiary households were located. The indirect beneficiaries constituted 28 percent of the total population in the ZOI at baseline. Therefore, the estimated impact of the program on the ZOI represented impact on three distinct groups in the ZOI:

Group 1: RVCP member households, designated as "RVCP direct beneficiaries"

Group 2: RVCP nonmember households located in RVCP areas, designated as "RVCP indirect beneficiaries"

Group 3: Households located outside the RVCP areas, designated as "Health Only"

The first two groups were, by definition, direct or indirect participants in the RVCP; however, they were also part of the target population of the health and nutrition program. They would therefore be exposed to the *integrated program* (RVCP and health and nutrition program). The third group, Health Only, was exposed to the health and nutrition program solely. This unique design permitted the investigation of the following secondary evaluation questions:

2.A. What has been the impact of the integrated program (RVCP and health and nutrition) on key outcomes in the integrated program areas (on Group 1 and Group 2)?

2.B. What has been the impact of the health and nutrition program, acting without the RVCP, on key outcomes in the Health Only areas?

2.C. Was the integrated program (RVCP and health and nutrition) more effective than the health and nutrition program alone at improving key outcomes at the population level?

2.D. Did the RVCP have indirect effects on the nonmember households located in the RVCP areas (Group 2)?

Note that question 2.C looked at the relative effectiveness of the integrated program compared with the "Health Only" program, whereas question 2.D examined the externalities or spillover effects of the RVCP on indirect beneficiaries living in the RVCP areas (Group 2).

Methods

The baseline survey gathered data from a representative sample of 4,007 households in the 30 priority municipalities (ZOI), and 2,294 households in propensity score matched comparison areas. To enable rigorous analysis of both intermediate and longer-term effects, the evaluation was based on a prospective, quasi-experimental difference-in-differences (DID) design with a matched comparison group and controls for household-level fixed effects (FE). This design required panel data at the household level. Robustness checks were conducted using data from the panel of clusters and controls for cluster-level FE. The midline survey was conducted between August and December 2015 and follow-up data were successfully obtained for 60 percent of households in the original sample. Although attrition was high overall, analysis suggested it was not selective.

Results

Table 1 presents a summary of the main results for the panel of households, both summary statistics and program impact estimates derived from the household FE DID models.[1]

The results showed a trend toward improvements in some economic and health outcomes in the WHIP ZOI but provided limited evidence that these changes were attributable to the program. Between 2013 and 2015, improvements were noted in per capita consumption, the prevalence and depth of poverty, food security, exclusive breastfeeding, stunting, and skilled birth attendance in the ZOI. However, many of these indicators also improved in the comparison group, suggesting that the program's impact on these outcomes after two years were either negligible or too small to detect without a larger sample. Other indicators, including the proportion of children ages 6 to 23 months receiving a minimum acceptable diet, modern family planning use, and receipt of four or more prenatal checkups, showed comparable declines in both the ZOI and comparison groups. Women's dietary diversity remained basically unchanged in both areas. The proportion of women who were underweight declined slightly in the ZOI and increased slightly in the comparison group, but there was not a statistically significant impact estimate in the household FE model. The program appeared to have a significant effect on fertility, but not in the expected direction; the proportion of women who gave birth in the two years before the survey decreased in the comparison group but did not change among those in the ZOI.

Analyses for the ZOI subgroups further suggest that the addition of the RVCP to the health program produced effects on most outcomes that were either negligible after two years or undetectable in a sample of this size using a household FE model.

[1] In Table 1, prevalence values are reported as proportions. Impact estimates for prevalence indicators are interpreted as the additional change in the prevalence of the outcome achieved in the WHIP ZOI relative to the change occurring in the comparison group, controlling for differences in observed characteristics and other fixed unobserved factors. For instance, an impact estimate of -0.014 for the prevalence of poverty at USD 1.25 per person per day means that in the WHIP ZOI, the proportion of people living in USD 1.25 poverty declined in magnitude by 0.014 (or 1.4 percentage points) relative to the change in the comparison group; however, the difference was not statistically significant at standard significance levels of 5 percent or 10 percent. Outcomes expressed in other units have similar interpretations. For example, for daily per capita expenditures, an estimate of 0.189 means that in the WHIP ZOI, expenditures increased by 0.189 2010 United States dollars (USD) more than the change in comparison areas, controlling for differences in observed characteristics and other fixed unobserved differences between the groups. This difference was also not statistically significant at standard significance levels. Additional information on the units of the indicators is presented in the Table 1 footnote.

Limitations

There are three important limitations to the WHIP midline impact evaluation. First, the short time interval between the 2013 baseline and 2015 midline surveys provided a limited opportunity for the program to affect change in poverty, nutritional status, health behavior, and health service use outcomes. The impact evaluation was powered to detect significant program impact after four years of implementation at endline. The midline evaluation was designed to serve as a progress check for high-level indicators, such as poverty and stunting, which are slow to move, and a midpoint impact analysis of behaviors and service use indicators that are more readily responsive to intervention. The second and most critical limitation of the midline impact evaluation is that the health and nutrition program was not implemented as planned because of the 2013 amendment to the Guatemalan national budget law. The final key limitation of the midline impact evaluation is potential problems with midline data quality, including the fidelity of households identified as panel households between baseline and midline, and issues around the reporting of birth dates in the women's questionnaire birth history. We do not consider any of these limitations to be "fatal flaws," but they are important to disclose.

Conclusions

Taken together, results of the midline impact evaluation indicate that although there is no statistically significant program impact on household consumption, poverty, or hunger, these indicators are moving in the expected direction, consistent with the program's theory of change. Mixed results in the time trends for infant and young child feeding practices, nutritional status, and the decreasing use of reproductive and maternal health services suggest that the cessation of the health and nutrition program in 2013/2014 may have had detrimental effects on these indicators in the ZOI areas.

INTRODUCTION

This document is the midline report for the impact evaluation of the USAID/Guatemala WHIP. The impact evaluation was implemented as part of a larger set of monitoring and evaluation activities. It sought to estimate the effects of WHIP interventions on key population-level outcomes. The impact evaluation was based on a prospective, quasi-experimental DID with a matched comparison group and FE evaluation design and used longitudinal data collected from the 2013 baseline survey and the 2015 midline survey. This report presents findings on the program's impact at its midpoint on key indicators: household consumption, poverty, food security, nutrition status, dietary practices, fertility, family planning, and use of maternal and child health services.

Country Background

Guatemala's population is one of the most disadvantaged in Latin America. High levels of poverty, malnutrition, and maternal and child morbidity and mortality persist. USAID works with the Government of Guatemala to implement a country strategy focused on sustainable development in the health, education, governance, environmental, and economic sectors. There are several United States-funded initiatives being implemented in Guatemala, including Feed the Future (FTF) and the Global Health Initiative, both of which were included in the WHIP. The WHIP consisted of 18 different programs designed to reduce poverty and chronic malnutrition in 30 priority municipalities in the Western Highlands region by incorporating interventions in the agriculture, health, governance, education, and climate change sectors. Among other activities, WHIP offered a package of interventions to help improve incomes among smallholder agricultural and handicraft producers, and support for strengthening maternal and child health services. WHIP officially began in mid-2012 and ran for five years, through 2017. As of 2013, WHIP covered an estimated population of 1.5 million people.

Early in the program design phase, USAID/Guatemala decided to plan a set of evaluation activities aimed at providing information on program performance and impact. In 2012, the Mission asked MEASURE Evaluation to design a joint performance and impact evaluation. Given the information needs of the Mission and the features of the program, the proposed design was a quasi-experimental DID design using a matched comparison group and data from three representative longitudinal household surveys to be conducted at baseline (2013), midline (2015), and endline (2017). Impact evaluation analyses were planned to be conducted at both midline, using baseline and 2015 data, and at the end of the program, using data from all three surveys.

In 2013, a research team, led by MEASURE Evaluation in collaboration with Guatemalan colleagues, implemented the 2013 WHIP Baseline Evaluation Survey (Encuesta de Monitoreo y Evaluación del Programa del Altiplano Occidental or EMEPAO). Fieldwork was conducted from July to November 2013. The results constituted the baseline for the impact evaluation and were reported separately (Angeles, Hidalgo, Molina-Cruz, Taylor, Urquieta-Salomón, Calderón, …Romero, 2014). In 2015, DevTech, an international consulting firm, implemented the 2015 WHIP Midline Evaluation Survey, with fieldwork conducted from August to December 2015. In 2015, USAID/Guatemala requested MEASURE Evaluation to conduct an impact evaluation analysis using baseline and midline information. This report presents results of the midline impact evaluation of the WHIP on key performance indicators from these two surveys, highlighting gains attributable to the program during the first two years of implementation.

THE WESTERN HIGHLANDS INTEGRATED PROGRAM

Objectives

The main objectives of the WHIP were to reduce chronic malnutrition among children and reduce household poverty in the 30 municipalities prioritized by the program. The WHIP combined interventions aimed at improving household economic conditions with interventions that sought to increase access to and use of nutrition, maternal and child health, and family planning services. The program coordinated the functions of several different initiatives: 1) the RVCP; 2) a health and nutrition program aimed at improving access to health, nutrition, and family planning services; 3) Title II/PL480, which provides food assistance and education about agricultural practices, nutrition, sanitation, and health; 4) a program designed to strengthen local governance; 5) a primary education program that focused on the quality of learning and on supporting bilingual education; and 6) a program aimed at reducing vulnerability to climate change. The RVCP and the health and nutrition programs (1 and 2) were the explicit focus of this impact evaluation.

WHIP Components and Interventions

Two of the WHIP initiatives—the RVCP and the health and nutrition program—formed the basis for the sampling strategy used in the impact evaluation design and associated surveys. It is important to note that the health and nutrition program covered the entire population in the 30 priority municipalities, whereas the RVCP focused only on selected producers' associations located in the 30 municipalities. The RVCP was expected to generate effects mainly among the members of the producer's associations participating in the RVCP and other members of their households, but it was also hypothesized to have indirect effects on households located in nearby neighborhoods. These externalities, or spillover effects of the RVCP, were expected to occur through increased local expenditures and by employment opportunities generated from increased productive activities among RVCP households.

The Rural Value Chains Project

The RVCP was based on the value chain model promoted by FTF. It had two main components. The first component consisted of technical assistance and training provided to agricultural producers' associations working in horticulture and coffee value chains and to handicraft producers on how to increase production, improve product quality, expand market competitiveness, and gain access to national and international markets. This first component was expected to improve household income, and dietary quantity and quality were expected to subsequently improve. The second component sought to improve families' nutritional status through education and information communication in RVCP members' households. It was expected that the combination of education and income generation would lead to positive changes in nutrition-related behavior, improve household food availability, and increase children's and women's dietary diversity and quality, resulting in improved nutritional status among these populations. The RVCP was implemented by the Asociación Nacional del Café and the Asociación Guatemalteca de Exportadores. In early 2013, the program included 118 associations with members from approximately four percent of households in the 30 priority municipalities. Households with an RVCP producer were located in census tracts encompassing 28 percent of the total population in the ZOI. Thus, areas there had direct beneficiaries along with a much larger population of indirect beneficiaries.

The Health and Nutrition Program

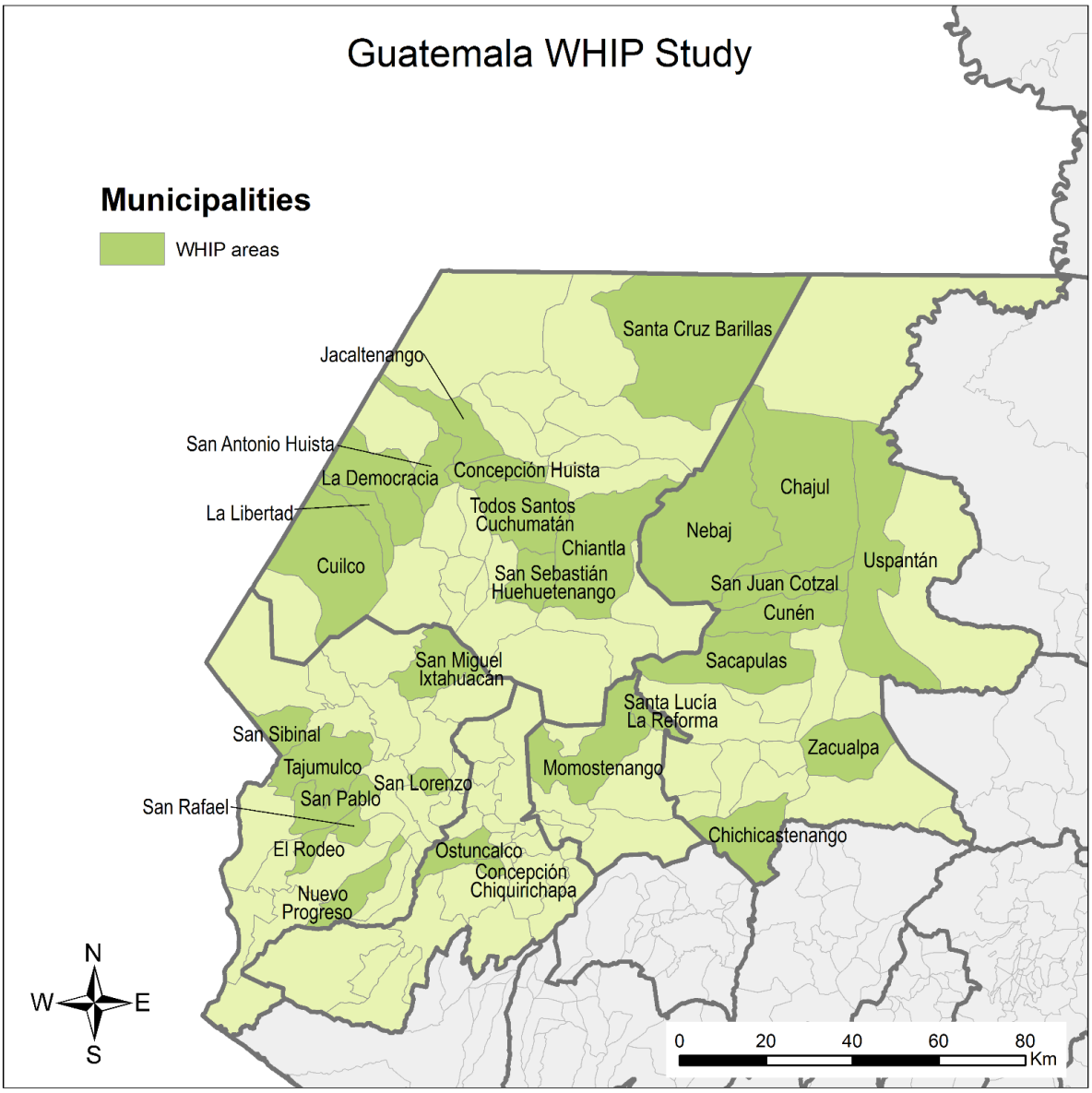
The health and nutrition program was developed as an integrated set of interventions to improve health and nutrition outcomes among women and children in the Western Highlands region by promoting improvements in health and nutrition practices at the household level, expanding the availability of health and nutrition services, and improving the quality of the health care system and services. According to its original design, the program was to support behavior change to improve home health practices and increase the use of maternal and child health, reproductive health, and family planning services. The program was also to include initiatives to improve the cultural relevance, transparency, and accountability of public health services.

Unfortunately, 2013 and 2014 cuts to the Guatemalan public health system reduced the availability of services the health and nutrition program sought to provide. The PEC began in 1997 as a response to the health provision goals of the Guatemalan Peace Accords signed in 1996. The MSPAS contracted NGOs as providers of health services to extend coverage to primarily indigenous and impoverished communities without access to a health provider. In November 2013, Guatemala's Congress passed an amendment to the budget law that would halt all MSPAS funding to PEC NGOs by 2017. In 2014, the MSPAS precipitously terminated the contracts of the PEC NGOs and health service delivery for rural populations was significantly reduced. The situation worsened in the absence of a suitable strategy to replace these NGOs. Lack of coverage exacerbated the effects of a health system hobbled by pervasive stockouts of essential medicines, family planning supplies, and vaccines. Insufficient human resources and unpaid MSPAS field staff worsened the delivery of care, especially at the primary level of care. Continuous changes in MSPAS administration have led to delayed implementation of a new model of care to provide appropriate health services to an estimated one-third of the total Guatemalan population.

WHIP Zone of Influence: Target Population and Geographic Areas

The WHIP target area was composed of 30 priority municipalities in five departments located in the Western Highlands region: Quiché, San Marcos, Huehuetenango, Totonicapán, and Quetzaltenango. In accordance with FTF terminology, the 30 priority municipalities as a whole were also known as the WHIP ZOI. Figure 1 presents the location of the WHIP municipalities. Appendix A lists the WHIP priority municipalities and their mid-2013 estimated population totals.

Figure 1. WHIP priority municipalities



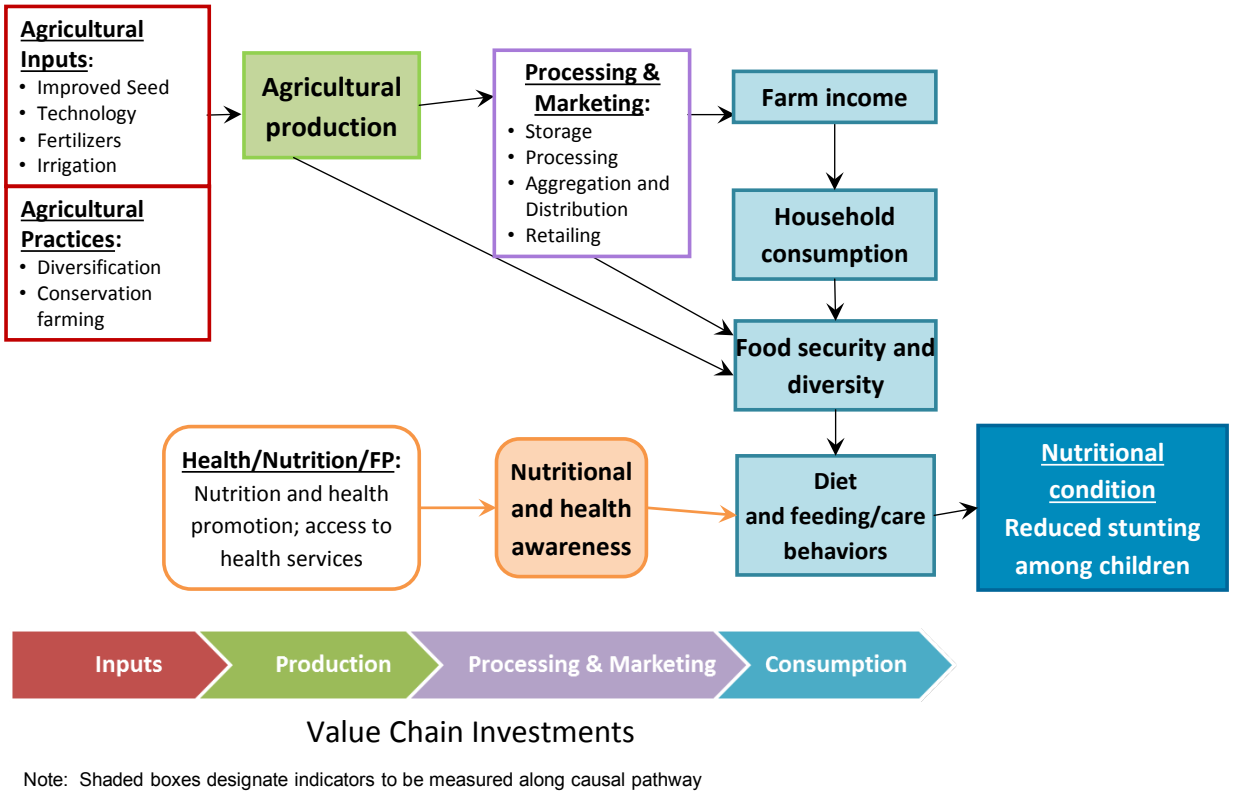
THEORY OF CHANGE

The WHIP's conceptual framework is based on the FTF/FEEDBACK approach to improving nutrition, both the immediate and underlying causes of malnutrition. Nutrition-specific interventions, including nutrition education, supplementary feeding, and micronutrient fortification, can effectively address the direct determinants of malnutrition. However, if left unaddressed, common underlying causes of malnutrition, such as the scarcity of assets (e.g., food and income) may impede improvements in the nutritional status of beneficiaries.

In Guatemala, the WHIP was designed to implement health and nutrition interventions across all 30 municipalities. Nutrition activities focused on improving diet and feeding/care behaviors that were ultimately expected to improve the nutritional status of children under five (Figure 2). The RVCP, which was implemented in a small subset of households in the WHIP municipalities, was expected to have an additional effect on nutrition, on both the household members participating in the RVCP and on neighboring households, as previously noted. The causal pathway posits that the RVCP would increase producers' farm/handicraft income, thus boosting household consumption, increasing food security and diversity, and ultimately improving the nutritional status of household members. The RVCP activities specifically focused on the use of improved seeds and fertilizers and more effective cropping techniques to increase agricultural productivity. Additional intervention activities focused on improving processing and marketing, which would contribute to increasing household income and consumption. Taken together, increased food production and income among RVCP direct beneficiary households were expected to lead to greater household food security, enhanced dietary diversity, and improved nutritional status.

As discussed above, there was another group of households that could indirectly benefit from the RVCP: households that did not receive RVCP interventions themselves, but were located in the same communities as the RVCP direct beneficiaries. The RVCP component of the WHIP could generate positive externalities and indirect impact in RVCP areas through both increased community access to nutritious foods produced by RVCP households and through new employment opportunities on and off RVCP farms. RVCP direct beneficiary households that increased their production of labor-intensive crops could create employment opportunities for day laborers, and the income earned by both the farm and the farm employees would circulate in the local nonfarm economy, stimulating further economic growth, poverty reduction, and subsequent improvements in nutrition and health in the RVCP community at large.

Figure 2. Causal chain: RVCP and health/nutrition/family planning program



Note: Shaded boxes designate indicators to be measured along causal pathway

METHODS: THE WHIP IMPACT EVALUATION

Objective and Evaluation Questions

The main objective of the impact evaluation was to estimate the changes in key indicators that were attributable to the WHIP interventions. The primary evaluation question was:

1. What were the effects (impact) of the WHIP on key indicators at the population level in the WHIP ZOI?

Key results to examine were program impact on indicators of poverty, nutrition, and reproductive health. As proposed in the overall evaluation design, program effects would be examined twice: at the program's midpoint (from 2013 to 2015, as reported here) and over the life of the program (from 2013 to 2017).

The primary evaluation question required additional disaggregation because each program component under study covered different population groups in the ZOI. The target population of the health and nutrition program included the entire ZOI, whereas the RVCP targeted only members of households with a producer participating in an RVCP producer's association (direct beneficiaries). The direct beneficiaries made up four percent of all households in the ZOI at baseline. As noted, RVCP activities were also expected to have indirect effects on nonparticipating households (i.e., indirect beneficiaries) who resided in the same communities in which RVCP direct beneficiary households were located. The indirect beneficiaries constituted approximately 28 percent of the total population in the ZOI at baseline. The estimated impact of the program on the ZOI therefore represented impact on three distinct groups in the ZOI:

Group 1: RVCP member households, designated as "RVCP direct beneficiaries"

Group 2: RVCP nonmember households located in RVCP areas, designated as "RVCP indirect beneficiaries"

Group 3: Households located outside the RVCP areas, designated as "Health Only"

The first two groups were, by definition, direct or indirect participants in the RVCP; however, they were also part of the target population of the health and nutrition program. They were therefore exposed to the *integrated program* (RVCP and health and nutrition program). The third group, Health Only, was exposed to the health and nutrition program solely. This unique design permitted the investigation of the following secondary evaluation questions:

2.A. What has been the impact of the integrated program (RVCP and health and nutrition) on key outcomes in the integrated program areas (on Group 1 and Group 2)?

2.B. What has been the impact of the health and nutrition program, acting without the RVCP, on key outcomes in the Health Only areas?

2.C. Was the integrated program (RVCP and health and nutrition) more effective than the health and nutrition program alone at improving key outcomes at the population level?

2.D. Did the RVCP have indirect effects on the nonmember households located in the RVCP areas (Group 2)?

Note that question 2.C looked at the relative effectiveness of the integrated program compared with a "Health Only" program, whereas question 2.D examined the externalities or spillover effects of the RVCP on indirect beneficiaries living in RVCP areas (Group 2).

Study Design

The impact evaluation sought to answer evaluation questions 1, 2.a, 2.b, 2.c, and 2.d. The assessment of program impact required the evaluators to estimate what would have happened if the WHIP had not been implemented, which necessitated a comparison group. Given that the WHIP priority areas were purposively selected before the beginning of the evaluation, it was not feasible to consider a randomized control design, the preferred option for assessing impact. The evaluation strategy was therefore based on the second-best feasible alternative: a prospective, quasi-experimental, DID design with a matched comparison group and FE controls. This design estimates program impact by comparing changes in the program group between the 2013 baseline and 2015 midline (or the endline in 2017) to changes in the comparison group over the same period, controlling for household- and community-level differences between the two groups.

The validity of the impact estimates obtained by this design depends on what is called "the parallel trends assumption," which assumes that the comparison group provides a good approximation of the change that would have occurred in the WHIP areas if the program had never been implemented. To find an appropriate comparison group, propensity score matching (PSM) was used to identify a group of census tracts (*sectores censales*) located in Western Highlands municipalities that were *not part* of the WHIP ZOI, but were statistically similar in demographic and socioeconomic characteristics to the census tracts selected in the program areas for the baseline survey. Matching was performed at the census tract level using information contained in the 2002 Population Census and the 2002 Agricultural Census. Potential comparison census tracts were identified from an expanded list of municipalities eligible for the FTF project, excluding the 30 WHIP priority municipalities. The PSM was performed based on census tract-level demographic, agricultural, and economic development characteristics. Additional information on the matching procedures and the identification of the comparison groups is provided in the sampling discussion found in the section below titled Selection of Treatment and Comparison Groups, and in Annex 1 of the baseline report (Angeles, et al., 2014). Two matched comparison groups were identified: one for the RVCP areas (for the clusters where Groups 1 and 2 lived) and another for the Health Only group (for the clusters where Group 3 lived).

It is important to note that due to the nonrandomized nature of the impact evaluation, additional variables were included in the impact estimation models to control for observed differences between the program and the comparison groups that persisted even after the matching process. Moreover, to control for unobserved differences between the groups, analyses included household-level FE controls. To be able to employ household-level FE, surveys should be conducted longitudinally in households. As a secondary specification and for robustness check purposes, we also used and compared results from models with cluster-level FE.

Estimation Strategy

Primary Evaluation Question

To answer evaluation question 1—what were the main effects of the WHIP program?—we used the following equation:[2]

$$Y_{ijt} = \beta_0 + \beta_1 P_j + \beta_2 T_t + \beta_3 P_j T_t + \beta_4 X_{ijt} + \lambda_{ij} + \varepsilon_{ijt} \quad (1)$$

[2] Equation (1) presents the classic DID specification, including both the program and time dummy variables individually, but note that P_j will be perfectly collinear with the household-level dummies λ_{ij}. Therefore, an equivalent specification would be Equation (1′), where the coefficients of equations (1) and (1′) are equivalent.

$$Y_{ijt} = \beta_0 + \beta_2 T_t + \beta_3 P_j T_t + \beta_4 X_{ijt} + \lambda_{ij} + \varepsilon_{ijt} \quad (1')$$

where,

Y_{ijt}: Outcome of interest, at the individual or household level (i) in community j at time t

P_j: Dummy for the WHIP program in the community equal to 1 if community j is in the WHIP ZOI and equal to 0 if not

T_t: Dummy for time of the observation equal to 1 if the observation is from the 2015 midline survey and equal to 0 if it is from the 2013 baseline

X_{ijt}: Observed time-varying individual, household, and community characteristics

λ_{ij}: Unobserved characteristics of the households that do not change during the period of observation (these were the household dummy variables used to control for household-level FE)

ε_{ijt}: Random shocks to individual or household i in community j at time t

The $\beta's$ represented the parameters or coefficients to be estimated. Equation (1) was estimated with FE estimation methods applied on pooled data from the panel of households included in both the baseline and midline surveys. Summary statistics and impact estimates were calculated using sample weights (Appendix D), and standard errors were corrected for clustering at the cluster (*sector censal*) level. In this model, the coefficient of interest was β_3, the coefficient of the program*time interaction term, which was the DID program impact estimate. Its value was interpreted as the additional change in the outcome achieved in the WHIP ZOI area relative to the change occurring in the comparison group, controlling for differences in the observed characteristics, X_{ijt}, and for fixed unobserved differences between the households and communities.

To conduct robustness checks, we used a variant of Equation 1 to control for cluster-level FE only:

$$Y_{ijt} = \gamma_0 + \gamma_1 P_j + \gamma_2 T_t + \gamma_3 P_j T_t + \gamma_4 X_{ijt} + \lambda_j + \varepsilon_{ijt} \quad (2)$$

where λ_j represented the community-level characteristics that did not change over the observation period (cluster-level FE). The $\gamma's$ represented the coefficients to be estimated. Equation (2) was estimated with the FE estimation methods applied to pooled data from the panel of clusters included in both the baseline and midline surveys. The coefficient of interest was γ_3, which was the DID program impact estimate, and it was interpreted as the additional change in the outcome achieved in the WHIP ZOI area relative to the change occurring in the comparison group, controlling for differences in the observed characteristics, X_{ijt}, and for fixed unobserved differences between the communities.

In the results sections, we report three sets of impact estimates:

Simple DID: program impact estimated from the simplified version of Equation (1) without controls for individual, household, or community characteristics or FE, estimated among the household panel with standard errors corrected for clustering at the cluster level. It presents the difference between the changes in the means observed in the ZOI and those in comparison areas.

Household FE: program impact estimated using Equation (1), including relevant individual, household, and community controls and household-level FE, estimated among the household panel with standard errors corrected for clustering at the cluster level.

Cluster FE: program impact estimated using Equation (2), including relevant individual, household, and community controls and cluster-level FE, estimated among the cluster panel with standard errors corrected for clustering at the cluster level.

We also investigated the presence of heterogeneous program impact by baseline household expenditure levels for all key indicators and for child sex and age group for under five feeding and nutrition indicators. The household FE model (Equation 1) was run separately by subgroup, and the results are presented in Appendices F and G. We investigated impact heterogeneity by baseline consumption level by analyzing program impact among households below (lower 50 percent) and above (upper 50 percent) the baseline sample-weighted median household expenditure level (USD 3.34 2010 USD per capita daily expenditure).

Secondary Evaluation Questions

The following model was used to answer evaluation questions 2.a, 2.b, 2.c, and 2.d:[3]

$$Y_{ijt} = \alpha_0 + \alpha_1 T_t + \alpha_2 RVCD_{ij} + \alpha_3 RVCI_{ij} + \alpha_4 HealthO_{ij} + \alpha_5 RVCD_{ij} T_t + \alpha_6 RVCI_{ij} T_t + \alpha_7 HealthO_{ij} T_t + \alpha_8 X_{ijt} + \lambda_{ij} + \varepsilon_{ijt} \quad (3)$$

where,

$RVCD_{ij}$: Dummy indicating the household is in Group 1 (RVCP direct beneficiary); equal to 1 if household *i* in community *j* is in Group 1, equal to 0 otherwise

$RVCI_{ij}$: Dummy indicating the household is in Group 2 (RVCP indirect beneficiary); equal to 1 if household *i* in community *j* is in Group 2, equal to 0 otherwise

$HealthO_{ij}$: Dummy indicating the household is in Group 3 (Health Only); equal to 1 if household *i* in community *j* is in Group 3, equal to 0 otherwise

The *a*'s represented the coefficients to be estimated. Equation (3) was estimated with household-level FE estimation methods applied on pooled data from the panel of households included in both the baseline and midline surveys. As in the previous models, standard errors were corrected for clustering at the cluster level. The estimated parameters of this model allowed us to answer the secondary evaluation questions. We reported the following DID impact estimates and impact comparisons in the secondary program impact results tables:

RVCP Direct:	Evaluation question 2.a: impact of the integrated program on Group 1 (RVCP direct beneficiaries)	α_5
RVCP Indirect:	Evaluation question 2.a: impact of the integrated program on Group 2 (RVCP indirect beneficiaries)	α_6
Health Only:	Evaluation question 2.b: impact of the health and nutrition program only on Group 3 (Health Only)	α_7

[3] Similar to Equation (1), Equation (3) presents the full DID specification for the secondary impact evaluation equations, including the program and time dummy variables individually. The program dummy variables will be perfectly collinear with the household-level dummies λ_{ij}, and an equivalent specification that excludes coefficients for the program dummy variables is given in equation (3') and the estimates parameters would be equivalent to those estimated from equation (3):

$$Y_{ijt} = \alpha_0 + \alpha_1 T_t + \alpha_5 RVCD_{ij} T_t + \alpha_6 RVCI_{ij} T_t + \alpha_7 HealthO_{ij} T_t + \alpha_8 X_{ijt} + \lambda_{ij} + \varepsilon_{ijt} \quad (3')$$

RVCP Direct – Health Only:	Evaluation question 2.c: relative impact of the RVCP direct program compared with the Health Only program	$\alpha_5 - \alpha_7$
RVCP Indirect – Health Only:	Evaluation question 2.d: spillover effect of the RVCP on Group 2 (RVCP indirect beneficiaries compared with Health Only); if there was no spillover effect, then the RVCP indirect impact would not be statistically different from the Health Only impact	$\alpha_6 - \alpha_7$

Last, to conduct robustness checks, we used a variant of Equation (3) controlling for cluster-level FE only. The model therefore becomes:

$$Y_{ijt} = \alpha_0 + \alpha_1 T_t + \alpha_2 RVCD_{ij} + \alpha_3 RVCI_{ij} + \alpha_4 HealthO_{ij} + \alpha_5 RVCD_{ij} T_t + \alpha_6 RVCI_{ij} T_t + \alpha_7 HealthO_{ij} T_t + \alpha_8 X_{ijt} + \lambda_j + \varepsilon_{ijt} \quad (4)$$

Equation (4) was estimated on pooled data from the panel of clusters. The interpretation of the parameters remained the same as above, but in the context of controlling for cluster-level FE only to account for fixed unobserved differences between the intervention and comparison areas.

Data

As previously noted, the nonrandomized nature of the WHIP impact evaluation design required us to control for differences between the groups on unobserved household and community characteristics. To this end, the estimation models used FE strategies. The implementation of a FE specification requires longitudinal data, that is, several observations over time for each unit of analysis. The baseline, midline, and endline household surveys therefore planned to be longitudinal at the household level.

Baseline Survey

Objectives

The 2013 baseline evaluation survey was designed to achieve several objectives: 1) provide representative estimates of key indicators for the ZOI; 2) provide a demographic, socio-economic, and health/nutrition profile of the WHIP population at the start of the program; and 3) provide a comparison group for the WHIP population and test the extent of the balance between the WHIP intervention and comparison groups before program implementation. The baseline survey was the basis for the longitudinal midline and endline surveys planned for implementation in 2015 and 2017. The baseline survey was based on a representative sample of households from the three groups in the ZOI (RVCP direct, RVCP indirect, and Health Only) and inclusive of a sample of households from clusters in neighboring municipalities matched to the ZOI clusters selected for the baseline.

Selection of Treatment and Comparison Groups

Randomized control trials are the gold standard for evaluating the impact of development interventions. The randomized control trial framework was not an option for the Guatemala WHIP impact evaluation because ZOI locations were purposively targeted based on community need and the ability to successful implement and use WHIP intervention components. Before the start of the evaluation, the USAID/Guatemala Mission identified 30 priority municipalities for inclusion in the ZOI based on poverty levels,

high rates of child malnutrition, and an economic and agricultural infrastructure conducive to successful program implementation and uptake (e.g., functioning food and farm input markets, roads, financing, and other necessary infrastructure). The 30 priority WHIP municipalities were selected from a larger list of approximately 70 municipalities that were eligible for the FTF project (Hernández, January 2011; Maruyama, January 2011). Because the randomized control trial strategy was not a viable option, the WHIP impact evaluation design used the next best option, a prospective longitudinal DID design with controls for FE and a matched comparison group. The validity of the impact estimates depended on identifying an appropriate comparison group. Achieving a high degree of similarity between treatment and comparison areas both in terms of outcomes and observed determinants lends more confidence to the assumption that treatment and comparison groups are also similar in terms of unobservable characteristics.

Because a primary objective of the baseline survey was to obtain representative samples from the three ZOI domains (RVCP direct beneficiaries, RVCP indirect beneficiaries, and Health Only), and given that RVCP direct beneficiaries comprised only four percent of the baseline ZOI sample, the sampling of intervention areas, the selection of comparison groups, and the sampling of comparison areas were an iterative process. First, a list of census tracts for the 30 priority WHIP municipalities and a list of census tracts for the remaining 40 eligible municipalities were developed. Based on information obtained from the RVCP implementing partners, we identified census tracts where at least one member of an RVCP association resided at baseline. These tracts were defined as the combination of Groups 1 and 2, and in each of these tracts, households with an RVCP member were identified as Group 1 (RVCP direct beneficiaries), whereas the remaining households constituted Group 2 (RVCP indirect beneficiaries). The Health Only tracts (Group 3) did not include any RVCP-associated households.

Once the RVCP and Health Only treatment census tracts were identified, PSM methods were used to select comparison census tracts from the 40 nonpriority eligible municipalities using data from the 2002 Population Census and the 2002 Agricultural Census (the most recent data available at baseline). Matching procedures were undertaken to find two comparison groups. Group 4, the comparison group for the RVCP direct and indirect beneficiaries, was identified using matching procedures based on census tract aggregate characteristics, including population size; improved housing structures; improved water, sanitation, and cooking facilities; indigenous and Spanish-speaking population percentages; percentage of agricultural producers in the population; farm characteristics and agricultural production profiles; and other department-specific variables. The second comparison group (Group 5) was selected for the Health Only intervention group through PSM using similar characteristics, but with a greater focus on characteristics relative to public administration and less focus on agricultural profiles.

Based on the PSM results, 95 intervention census tracts for Groups 1 and 2 were matched with 78 comparison census tracts (comparison Group 4), and 37 Health Only intervention census tracts for study Group 3 were matched with 30 comparison census tracts (comparison Group 5). Appendix B presents a map of the ZOI and comparison area municipalities included in the impact evaluation. If one or more treatment or comparison census tract was selected through PSM to be included in the evaluation study, the entire municipality is shaded in the map. These matched intervention and comparison census tracts formed the preliminary evaluation survey sampling frame. Additional information on the matching procedures and survey methodology, including detailed sampling procedures, is available in the 2013 baseline report, Annex 2 (Angeles, et al., 2014).

Data Collection

Fieldwork for the baseline survey was conducted by MEASURE Evaluation from July to November 2013. The survey used the following data collection instruments: questionnaires for household, consumption, women (with modules for children's health), anthropometry and hemoglobinometry (women and children), Women's Empowerment in Agriculture Index, community, and health facility. Baseline interviews were completed in 4,007 households in the WHIP ZOI and 2,294 households in the comparison group areas. Additional information about the baseline survey design and results are provided in the 2013 WHIP baseline evaluation survey report (Angeles, et al., 2014).

Balance Between Intervention and Comparison Groups

We undertook a series of balance checks after the baseline data were collected from intervention and matched comparison households. The purpose of the balance checks was to assess the statistical equivalence of intervention and comparison groups based on observable characteristics to determine whether the matching process at the municipality and cluster (*sector censal*) levels succeeded in generating similar treatment and comparison groups, and whether the parallel trends assumption seemed plausible in our intervention and matched comparison groups. Linear regression models with correction for clustering were estimated for 75 variables, including outcome indicators, household characteristics, and other important contextual factors. Outcome indicators were consumption and poverty, food security and diet diversity, nutritional status of young children and women of reproductive age, feeding practices, and use of reproductive health services. Contextual variables were individual and household characteristics, housing physical characteristics, detailed nutritional status and farming information, fertility and family planning, and general child health status and service use variables. Statistical tests were performed to compare the RVCP groups with comparison Group 4 and to compare the Health Only group with comparison Group 5.

Results from the baseline balancing analyses established that the RVCP domains (Groups 1 and 2) were similar, on average, to comparison Group 4 on 56 of 75 indicators (75 percent of tested variables), and that the Health Only group and comparison Group 5 were statistically similar on 65 of 75 indicators (87 percent). Although this degree of similarity was not high, it was also not unexpected given the nonrandomized selection of the comparison group. Treatment and comparison groups were balanced at the 95 percent significance level for key WHIP/FTF indicators, except for the percentage of households with moderate/severe hunger, exclusive breastfeeding among children under six months of age, the percentage of births attended by a physician or nurse, the percentage of births with four or more prenatal care checkups, and the total fertility rate. Detailed sample means, differences, and balance test p-values are given in Table 14.1 and Table 14.2 in the baseline survey report (Angeles, et al., 2014).

The results indicate the continuation of differences between program and comparison groups, and reinforce the need to control for these baseline differences in the impact estimation models. The DID estimation approach allowed for the inclusion of additional control variables to account for characteristics not adequately controlled for by the matching process used to select the comparison group, and FE specifications were used to control for unobserved, time-invariant differences across study groups.

Midline Survey

The 2015 midline survey was conducted by DevTech, with fieldwork implemented from August through December 2015. It was designed to provide a second set of key indicator outcome measures for baseline households and representative estimates for the situation in the ZOI as of 2015. The midline survey used

the same data collection instruments as the baseline survey. The midline survey intended to revisit baseline households plus additional households selected to account for modifications in the RVCP intervention areas, although the realized sample had only a subset of the intended household panel sample (those baseline households located and responding at midline). The additional households sampled only at midline were not included in the midline impact evaluation report because the DID strategy requires panel data for the household FE and the cluster FE specifications.

Sample Attrition

Attrition occurs when households from the baseline sample are missing in the follow-up sample. Attrition may be caused by multiple events, such as migration, death, dissolution of households, or any other event that makes it difficult to locate a household during follow-up data collection. Attrition is important for estimating program impact because it decreases the sample size, leading to less precise impact estimates, but also could introduce bias in the evaluation sample. If attrition is selective, that is, if those leaving the sample are different from those who remain, it could lead to inaccurate program impact estimates and/or affect the representativeness of the sample.

We examined both overall and differential attrition from the baseline to the midline surveys. Overall attrition refers to the total share of observations missing at follow up from the original baseline sample. Overall attrition can change the characteristics of the remaining sample, rendering it nonrepresentative of the population from which it was obtained. In that sense, overall attrition could affect our ability to generalize the evaluation results to the population of interest: the WHIP ZOI and its subgroups. Differential attrition occurs when the treatment and control samples differ in the types of households that leave the sample. Differential attrition can create biased samples and impact estimates by increasing the differences between the program and comparison groups obtained at baseline.

Table 2 presents information on the number of households in the baseline sample, the panel (those also interviewed in the midline survey), and those lost to follow up at midline. Note that households sampled only at midterm are not included in the impact analysis or attrition assessment. Overall attrition was high, at 40.1 percent, with very small variation in the ZOI (40.6 percent) and the comparison areas (39.3 percent). Attrition was similarly high in each of the program subgroups.

Table 2. Households in the baseline survey, panel, and attrition

Groups	2013 Baseline	In the panel	Lost to follow up	Retained in the panel (%)	Attrition rate (%)
1	1,264	782	482	61.9	38.1
2	1,746	996	750	57.0	43.0
3	997	602	395	60.4	39.6
Total ZOI	4,007	2,380	1,627	59.4	40.6
Comparison 1+2	1,438	916	522	63.7	36.3
Comparison 3	856	476	380	55.6	44.4
Total comparison	2,294	1,392	902	60.7	39.3
Total sample	**6,301**	**3,772**	**2,529**	**59.9**	**40.1**

We examined overall attrition in the ZOI and the total sample by comparing the average baseline values for households and individuals in the panel with those for households and individuals lost to follow up (attritors) on 205 indicators. (Appendix C, Section C.1 provides results tables for overall attrition.) We found that in the ZOI, only 15 of 205 indicators (7.3 percent) were statistically different at the five percent significance level. The results at the 10 percent significance level were 27 of 205, or 13.2 percent. We found similar results for the total sample. These results indicated that despite levels of loss to follow up, overall selective attrition in terms of bias introduced in the panel might not be a serious problem in the analysis sample. It is important to note, however, that although we did not find evidence of significant overall attrition differences in terms of household demographic profiles, we found significant differences in several important WHIP/FTF indicators between attritors and panel households. We found evidence of systematically higher incidence of stunting among attritors (seven percentage point difference) and lower incidence of minimum acceptable diet among children ages 6 to 23 months (eight percentage point difference), receipt of four or more prenatal checkups (10 percentage point difference), and lower modern contraceptive prevalence (6.5 percentage points) among attritors relative to panel households. These findings suggest that households that left the panel fared worse on average than those that remained in the panel for important indicators of food security, child nutritional status, and health service use. The potential therefore exists that the midline impact evaluation sample is no longer representative of the baseline population on these characteristics (e.g., we may be estimating impact using a subsection of the population with better-than-average outcomes).

We investigated differential attrition by testing for balance between the ZOI and the total comparison group using baseline data on the panel of households, that is, those included in both the baseline and midline surveys. (Appendix C, Section C.2 provides results tables for differential attrition.) We found no statistically significant difference in group means between the ZOI panel and the comparison group panel on 178 of 205 indicators (86.8 percent), which is a percentage similar to the balance results obtained for the original baseline sample. These findings indicate that the balance between the intervention and matched comparison groups has not deteriorated because of the high attrition rate; however, these results also support the need to use impact estimation methods that control for observed and unobserved differences between the groups. At the same time, we found evidence of significant differential attrition in two key indicators: the comparison group had systematically lower prevalence of minimum acceptable diet for children ages 6 to 23 months, and lower prevalence of receiving four or more prenatal checkups relative to the ZOI. These findings were consistent with the lack of balance discovered in these two indicators at baseline. We did not find evidence of differential attrition based on young child stunting or the number of children under five in the household between the ZOI and comparison panel households.

PROGRAM IMPACT ON CONSUMPTION AND POVERTY

The WHIP aimed to improve economic and health conditions among the households in the ZOI, with expected results involving measurably increased consumption expenditures and reductions in poverty. The baseline and midline surveys therefore had an extensive module designed to measure household consumption and expenditures. This section details the impact of the WHIP on these key program indicators at midline. In this section and the following two sections, results correspond to the DID with the household-level FE model using data from the panel of households. This is our preferred model. As described above in the sections on the Study Design and Estimation Strategy, this model specification allowed us to control for observed differences between the program and comparison groups and for unobserved fixed differences between the groups, which we considered necessary given the nonexperimental nature of the impact evaluation.

Outcome Measures

Per Capita Expenditures

The main measure used in the evaluation was the estimate of total daily household consumption per capita. Consumption estimates comprised cash expenses for goods and services consumed in the household and estimates of the cash value of services received and of assets, such as homes, plus the value of durable goods available to members of the household. The monetary value of goods produced in the household and any cash transfers from social assistance programs or remittances were also included.

Consumption per capita was obtained by dividing the estimate of total household consumption by the number of individuals in the household. The baseline and midline survey instruments adopted the national norms for measuring consumption, specifically, the expenditures and consumption module was modeled after the 2011 Living Standards Measurement Survey (Encuesta Nacional de Condiciones de Vida 2011 or ENCOVI 2011*).* Furthermore, algorithms used to process EMEPAO consumption data were based on those used in the 2006 and 2011 ENCOVI surveys. These approaches allowed for direct comparison of WHIP evaluation survey results with those from official national surveys. Consumption results were obtained in Quetzales and converted to 2010 USD. Appendix E provides a detailed explanation of the conversion factors applied.

Prevalence of Poverty at USD 1.25

We estimated the impact of WHIP on poverty using the prevalence of poverty with a cut-off line of USD 1.25 daily per capita (in international USD, 2005 PPP), equivalent to 8.66 Quetzales (2013) daily per capita and to 9.01 Quetzales (2015) daily per capita. This poverty line is commonly used by international agencies for inter-country comparisons. The USD 1.25 poverty line (2005 PPP) was also adopted by the FTF program to monitor performance in countries where the initiative is being implemented.

Depth of Poverty

The third measure used in this analysis was the depth of poverty indicator. This indicator was estimated using the poverty gap index, which measures the extent to which individuals' consumption, on average, falls below the poverty line as a percentage of the poverty line. This indicator captures changes in poverty that the prevalence of poverty (or headcount index) does not detect. For example, if consumption levels among the poor increase without necessarily crossing the poverty line, this change is not reflected in the prevalence of poverty, but it is reflected in the depth of poverty.

Main Program Impact

Results for the first evaluation question, the overall impact of the WHIP program on household expenditures and poverty in the aggregated WHIP ZOI population, are given in Table 3. The first set of columns presents summary statistics (means and standard errors) of the indicators for the panel of households. The last three columns present DID impact estimates using three model specifications, as previously discussed in the descriptions of the simple DID model, the household FE model (our preferred model), and the cluster FE model. The cluster FE model was only included as a robustness check.

At baseline, levels of consumption in the ZOI and comparison groups were low, at an average of 2010 USD 3.66 and USD 4.07, respectively. At midline, daily per capita consumption had risen to USD 3.89 in the ZOI but decreased to USD 3.99 in the comparison group, on average. Figure 3 presents the distribution of daily per capita expenditures, by survey round and intervention/comparison group, and shows that the expenditure distribution shifted to the right, indicating higher consumption levels, for both the ZOI and comparison areas between the baseline and midline survey rounds. The unadjusted model (simple DID) showed a small positive and marginally significant (at the 10 percent level) net change in per capita expenditures in the ZOI relative to the comparison group of USD 0.309. However, after including control variables and household-level FE, the impact estimate was smaller (USD 0.189) and no longer statistically significant. Results suggest that, although program impact was not significant at midline, per capita consumption was improving more in the ZOI relative to the comparison group. It is possible that the impact of WHIP was still too small to be statistically detected by our models given the available sample size.

Although we found decreased average expenditures in the comparison group, we observed a declining prevalence of poverty in both the ZOI and comparison groups, with a larger total decline in ZOI areas (from 6.6 percent to 4.3 percent) than in comparison areas (5.6 percent to 4.9 percent). Decreased poverty, with a corresponding decrease in mean expenditures in the comparison group, can best be explained by changes in the distribution of household consumption, particularly in the lower and upper tails of the expenditure distribution. As seen in Figure 3, the expenditure distribution above USD 8.00 per person per day in the comparison areas was larger at baseline than at midline, but the distribution of households below USD 1.25 was also larger at baseline than at midline; therefore, although the mean consumption level in the comparison areas was higher at baseline, the median consumption value (where the expenditure distribution peaks) was higher at midline, and the relative density of households below the poverty line was lower at midline. We also observed declines in the depth of poverty in both the ZOI and comparison areas, with a greater decline in the ZOI (net change of negative USD 0.138). Although the estimates of program impact were negative for both prevalence and depth of poverty, suggesting that ZOI areas experienced greater reductions in poverty prevalence and depth relative to the comparison areas, none of the impact estimates were significant for either poverty outcome. It is possible that the effects of the WHIP might have been occurring in the expected direction, but they were still relatively small to be detected after only two years of program implementation, particularly given the panel sample size constraints.

Findings of no significant program impact on household consumption expenditures, prevalence of poverty, or depth of poverty were consistent for households consuming both above and below the baseline median consumption level. (See Appendix F: Heterogeneous Impact Analysis by Baseline Consumption Level.)

Table 3. Main program impact on consumption and poverty

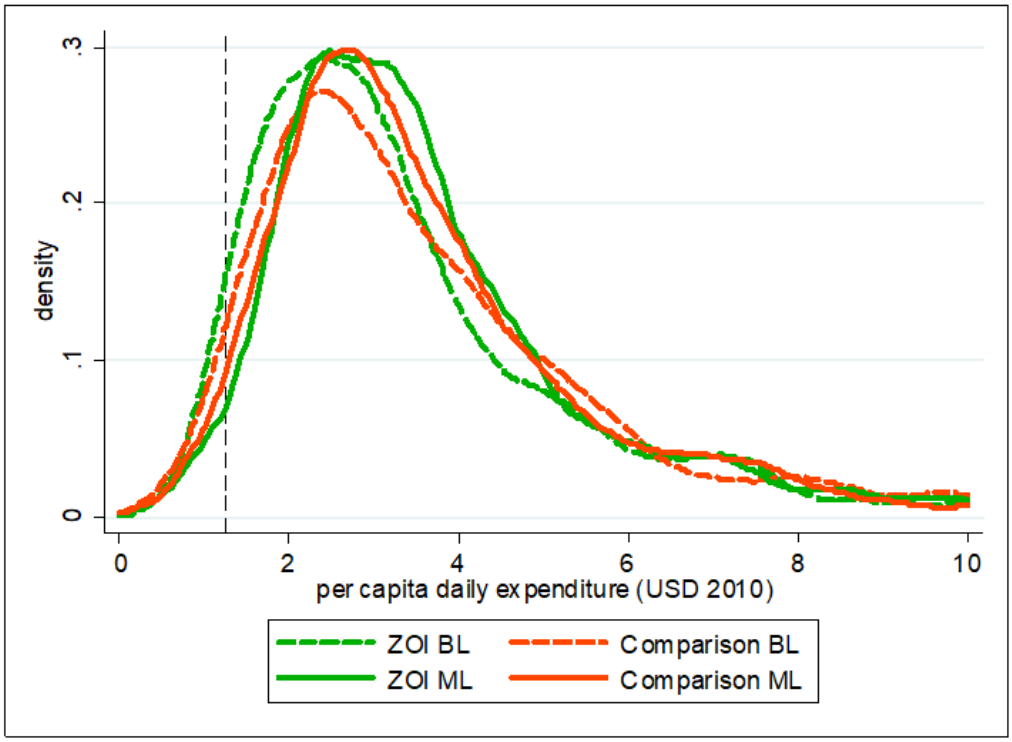
		Household Panel Summary Statistics				DID Impact Estimates		
		ZOI		Comparison		Household Panel		Cluster Panel [a]
Outcome Indicator	**N**	**Baseline**	**Midline**	**Baseline**	**Midline**	**Simple DID**	**Household FE**	**Cluster FE**
Daily per capita expenditures								
2010 USD	7,402	3.66	3.89	4.07	3.99	0.309+	0.189	0.214
		(0.19)	(0.18)	(0.21)	(0.16)	(0.167)	(0.236)	(0.152)
Prevalence of poverty (proportion)								
USD 1.25 (2005 PPP)	7,402	0.066	0.043	0.056	0.049	-0.016	-0.014	-0.014
		(0.014)	(0.010)	(0.013)	(0.011)	(0.021)	(0.026)	(0.017)
Depth of poverty								
Relative to USD 1.25	7,402	1.23	0.97	1.14	1.02	-0.138	-0.148	-0.242
		(0.30)	(0.29)	(0.32)	(0.32)	(0.509)	(0.617)	(0.440)

Notes: Summary statistics are means and standard errors for the panel of households. The simple DID model was estimated among the panel of households and did not include FE or control variables. The household FE model was estimated among the panel of households and included FE at the household level and control variables. The cluster FE model was estimated among households in the panel of clusters and included FE at the cluster level and control variables. Standard errors are in parentheses. Control variables comprised contemporaneous vector of household control variables, specifically, head of household characteristics (i.e., gender, age, literacy, education, principal occupation is farming, native language is Spanish, religion is Catholic, indigenous ethnic group), municipality, rural, had damage from 2012 earthquake, has improved drinking water, has improved sanitation, has soap and water to wash hands, has separate kitchen, uses modern cook-fuel, participation in social programs (i.e., fertilizer program, Mi Bono Seguro, Mi Bolsa Segura), household size, household crowding index, migration of household member for work, single woman-headed household, household dependency ratio, household member receives any other program benefits, and percentage of household members under five years old.

Significance levels: + p<0.10; * p<0.05; ** p<0.01; *** p<0.001

[a] The cluster FE model was estimated using the households in the panel of clusters.

Figure 3. Distribution of consumption expenditures by intervention group and study wave



Notes: The vertical reference line is set at the USD 1.25 poverty line (2010 USD). The consumption expenditure distribution is graphed for the panel of households and is truncated at USD 10.00 per capita daily for display purposes. The figure includes 97.12 percent of panel households. There was no significant difference in the proportion of ZOI vs. control households in the USD 0 to USD 10 range compared with the USD 10 and above range.

Secondary Impact by Program Component

Table 4 presents results related to evaluation questions 2.a through 2.d for the consumption and poverty indicators using the household FE DID model (Equation 3). Appendix H, Table H.1 presents subgroup means and standard errors for ZOI components. The first three columns with DID impact estimates show program impact on each of the three ZOI subgroups in the columns marked (1), (2), and (3). The next column presents the results of analyses examining the effectiveness of the integrated program relative to the health program acting alone (RVCP direct compared with Health Only), and the last column shows the externality effect of the RVCP on the RVCP indirect beneficiary group relative to the Health Only group.

Consistent with findings from the main impact analysis, we did not detect any significant program impact on the WHIP subcomponents (columns 1 through 3), and we did not find significant effects on the estimates of relative effectiveness and externalities (last two columns). For example, the finding of -0.241 for RVCP direct vs. Health Only program impact on daily per capita expenditures can be interpreted as RVCP direct beneficiaries spent -USD 0.241 less per person per day, on average, relative to the Health Only group, but this difference was not statistically significant. Similarly, the finding of -0.080 for RVCP indirect vs. Health Only indicates that the RVCP indirect households spent an average of USD 0.080 less than households in the Health Only areas, but this difference was not significant, leading us to conclude that there were not significant externalities of the RVCP direct intervention on non-RVCP households in RVCP areas (Group 2).

It is interesting to note that the average daily per capita expenditures increased for each of the WHIP intervention groups between baseline and midline, but the prevalence of poverty and depth of poverty slightly increased in both the RVCP direct and indirect areas (Appendix H, Table H.1). This is a similar situation to that previously described in the overall comparison group. Appendix I presents consumption expenditure graphs for the ZOI subgroups. It can be seen from the right-shifts of the expenditure distributions for each subgroup that the median consumption values were also increasing from baseline to midline in each subgroup, but further exploration of the bottom tail of the distribution (Appendix I, Figure I.2) shows that the density of RVCP direct and RVCP indirect households under USD 0.50 were higher at midline than at baseline, which corresponds to the slightly increased prevalence and depth of poverty observed among these groups at midline relative to baseline.

Table 4. Secondary program impact on consumption and poverty

		Household Panel DID Impact Estimates				
Outcome Indicator	N	RVCP Direct (1)	RVCP Indirect (2)	Health Only (HO) (3)	Direct vs. HO (1)-(3)	Indirect vs. HO (2)-(3)
Daily per capita expenditures						
2010 USD	7,402	-0.031	0.131	0.210	-0.241	-0.080
		(0.271)	(0.334)	(0.253)	(0.224)	(0.316)
Prevalence of poverty (proportion)						
USD 1.25 (2005 PPP)	7,402	0.007	0.009	-0.021	0.027	0.030
		(0.036)	(0.031)	(0.027)	(0.031)	(0.026)
Depth of poverty						
Relative to USD 1.25	7,402	0.332	0.264	-0.263	0.595	0.527
		(0.966)	(0.754)	(0.658)	(0.858)	(0.669)

Notes: The household FE model was estimated using the panel of households and included controls for FE at the household level and observed control variables. Standard errors corrected for clustering at the cluster (*sector censal*) level are in parentheses. Control variables include a contemporaneous vector of household control variables, as described in the Table 3 footnote.

Significance levels: + p<0.10; * p<0.05; ** p<0.01; *** p<0.001

PROGRAM IMPACT ON FOOD SECURITY AND NUTRITION

Outcome Measures

Prevalence of Food Insecurity – Household Hunger Scale

Food security was measured using the Household Hunger Scale (HHS) in accordance with FTF Guidelines (Feed the Future, 2016). The scale includes a short series of questions related to the adequacy of food supply and consumption, and the physical consequences of hunger among members of the household during the four weeks preceding the survey.[4]

The scale value begins at zero and is based on three source questions about situations pertaining to the lack of food: In the past four weeks "…was there ever a time when there was no food in your house?"; "…have you or any member of your household gone to bed hungry at night because there was not enough food?" and; "… have you or any member of your household spent one whole day and night without eating anything because there was not enough food?" For each affirmative response, participants are asked if the situation occurred "a few times (1 or 2 times)," or "sometimes (4 to 10 times)" and if so, one point is added to the HHS score. If the interviewee reports that the situation has occurred "many times (more than 10 times)," two points are added. The scale has a maximum value of six. Values totaling two or three indicate moderate hunger, whereas values from four to six inclusive indicate severe hunger.

Infant and Young Child Feeding

Exclusive breastfeeding and minimum acceptable diet indicators were derived from data in the women's questionnaire (birth history section). Only the most recent live birth (natural born children) for the women's questionnaire respondents was included in these indicators to maintain consistency with the methodology employed by other national Guatemalan demographic and health surveys (e.g., National Maternal and Child Health Survey—[Encuesta Nacional de Salud Materno y Infantil or ENSMI]).

Exclusive Breastfeeding among Children Ages 0 to 5 Months

Children ages 0 to 5 months at the time of the survey whose mothers reported that they were given only breastmilk on the previous day were considered exclusively breastfed. By definition, this included infants who received milk expressed (or from a wet nurse). Infants in this category may have also received oral rehydration salts (ORS), vitamins, minerals, and/or other medicines, but did not receive any additional nonbreastmilk food or liquid, including water.

Minimum Acceptable Diet in Children Aged 6 to 23 Months

The Minimum Acceptable Diet (MAD) indicator reflects adherence to basic standards for feeding frequency and dietary diversity in children ages 6 to 23 months old.[5] Dietary diversity is used as a proxy for the adequacy of micro-nutrient intake. The standard for breastfed children in this age range is to consume foods from at least four of the following seven food groups: 1) grains, roots, and tubers; 2) legumes and nuts; 3) dairy products (milk, yogurt, and cheese); 4) meats (beef, fish, poultry, and innards); 5) eggs; 6) fruits and vegetables high in Vitamin A; 7) other fruits and vegetables. Children who are not breastfed must consume milk/dairy and foods from at least four of the remaining six food groups. Feeding frequency is a proxy for

[4] http://www.fantaproject.org/sites/default/files/resources/HHS-Indicator-Guide-Aug2011.pdf.
[5] http://whqlibdoc.who.int/publications/2008/9789241596664_eng.pdf?ua=1 and (Feed the Future, 2016).

energy adequacy, and standards also vary according to the child's breastfeeding status and age. Children who are breastfed should receive solid, semi-solid, or soft foods at least twice a day from 6 to 8 months of age, and three times a day after 9 months. All children from 6 to 23 months of age who are not receiving breastmilk must be fed at least four times a day, including two milk feedings, to achieve adequacy.

Nutritional Status

The nutritional status of boys and girls under five years old and of women ages 15 to 49 was evaluated using anthropometric indices. Weight and height (recumbent length in the youngest children) were measured for all household members under age 60 months. Using this protocol, anthropometric results comprised children whose birth mothers were deceased, reside outside the household, and/or were not age-eligible to participate in the women's questionnaire.

Young Child Chronic Malnutrition – Stunting

Stunting, or low height-for-age, reflects chronic undernutrition. Stunting prevalence reflects the percentage of children with a height-for-age Z-score less than two standard deviations (<-2SD) below the reference mean, as defined by the World Health OrganizationWHO.[6] Stunting is an indication of linear growth retardation, which is often a consequence of prolonged exposure to poor health and inadequate diet.

Young Child Acute Malnutrition – Wasting

Wasting, or low weight-for-height, results from acute malnutrition and appears to be very rare among children under age five in the WHIP ZOI (<1 percent). Although we did examine the impact of the program on moderate and severe wasting (<-2SD below the reference mean weight-for-height), interpretation should take the very low prevalence into account.

Young Child Global Malnutrition – Underweight

Underweight, or low weight-for-age, reflects both chronic and acute undernutrition and may contribute substantially to the disease burden in medium- and low-income countries. We evaluated the impact of the program on the prevalence of children who were moderately or severely underweight (<-2SD below the reference mean weight-for-age).

Women of Reproductive Age – Dietary Diversity

Women's dietary diversity is a measure of micronutrient adequacy in the diets of women of reproductive age. The indicator is generated by summing the number of food groups consumed during the previous day and night. To calculate this indicator, information about foods consumed in the past 24 hours was collected and grouped in nine food groups as defined by the FTF Indicator Handbook (Feed the Future, 2016). The groups are: 1) grains, roots, and tubers; 2) legumes and nuts; 3) dairy products; 4) organ meat; 5) eggs; 6) flesh foods and other miscellaneous small animal protein; 7) vitamin A dark green leafy vegetables; 8) other vitamin A-rich vegetables and fruits; and 9) other fruits and vegetables. The mean number of food groups consumed was calculated by averaging the number of food groups consumed across all women ages 15 to 49 in the sample.

[6] World Health Organization Child Growth Standards. Available at http://www.who.int/childgrowth/en/.

Women of Reproductive Age – Underweight

The nutritional status of nonpregnant reproductive age women (15 to 49 years old) was measured using the body mass index (BMI). We report whether women are considered underweight, which is defined as having a BMI less than 18.5. BMI data come from the women's anthropometric module. BMI is calculated as weight in kilograms divided by height in meters-squared.

Main Program Impact

Table 5 presents summary statistics and the main effects of the WHIP on household food security, infant and young child feeding practices, and nutritional status outcomes for women and children.

Impact on Food Security

At baseline, moderate or severe hunger affected 14.4 percent of households in the ZOI compared with 16.7 percent of households in the comparison group. The prevalence of household hunger declined to 9.8 percent in the ZOI, a decrease of 4.6 percentage points. There was a similar decline of 4.9 percentage points in the comparison group to 11.8 percent. We did not find evidence of program impact on the prevalence of household hunger using either household- or cluster-level FE models. The lack of program impact on household hunger held for households consuming below and above the baseline median expenditures.

Impact on Infant and Young Child Feeding

Summary statistics and impact estimates for exclusive breastfeeding are presented using children from the panel of clusters, not the panel of households, given the low sample sizes for children ages 0 to 5 months in the household panel. (To be included in the household panel for this indicator, a household had to have an infant under six months of age at both baseline and midline surveys for the household FE specification to work.) Exclusive breastfeeding during the first six months of a child's life is common but not universal among children in the ZOI and comparison groups (76.6 percent and 73.4 percent at baseline, respectively). Prevalence increased at midline for both groups. We did not find evidence of significant program impact on the percentage of children ages 0 to 5 months who were exclusively breastfed.

At baseline, 41.5 percent of children ages 6 to 23 months in the ZOI and 34.5 percent in the comparison group (panel of households) were receiving a minimum acceptable diet. By midline, the prevalence of children in this age range with a minimum acceptable diet declined substantially to 24 percent for both study groups. Although the magnitude of the prevalence decrease was greater in the ZOI, the difference was not significantly different from that observed in the comparison group, as indicated by the simple DID estimate. After controlling for household FE and other observed characteristics, we still did not detect evidence of a significant program impact on MAD. We also did not detect heterogeneous program impact on MAD by baseline consumption level or child sex (Appendix F, Table F.1 and Appendix G, Table G.1).

Impact on Children's and Women's Nutritional Status

At baseline, 65.6 percent of children under five in the ZOI and 62.6 percent of children in the comparison group were affected by moderate or severe chronic malnutrition. The prevalence of stunting dropped to similar levels at midline, to 59.7 percent and 59.4 percent, respectively. As indicated by the simple DID estimate, the larger decline in the ZOI was not statistically different from that in the comparison group. We

did not detect any program impact on stunting in either the household panel or the cluster panel, and we did not find evidence of heterogeneous program impact on the prevalence of stunting by baseline household consumption, child sex, or child age (Appendix F, Table F.1, and Appendix G, Tables G.1 and G.2).

Acute malnutrition (wasting, low weight-for-height) was extremely rare, affecting less than 1 percent of children under five in the ZOI and comparison groups at baseline. Between baseline and midline, there were small and insignificant changes in wasting for both groups. We found no evidence of program impact on acute malnutrition. We also did not find evidence of program impact by baseline household consumption, child sex, or child age (Appendix F, Table F.1, and Appendix G, Tables G.1 and G.2).

Global malnutrition (underweight or low weight-for-age) was slightly higher in the comparison group (18.0 percent) than in the ZOI (15.3 percent) at baseline. By midline, the prevalence of underweight declined in the comparison group to 15.9 percent but increased in the ZOI to 18.9 percent. Program impact estimates from the household FE model indicated a significant 9.1 percentage point increase in the prevalence of underweight among WHIP households relative to comparison households ($p<0.05$). This significant impact appeared to be driven by program impact among households in the bottom 50 percent of the baseline consumption distribution (Appendix F, Table F.1). Children from low-expenditure WHIP households were 11.4 percentage points more likely to be underweight than children from low baseline expenditure comparison households ($p<0.05$), but there was no significant impact detected in the household panel for children in households consuming above the baseline median. We detected only a significant program impact on prevalence of underweight using household FE for male children (12.5 percentage points, $p<0.05$), and found no significant effects for female children (Appendix G, Table G.1).

On average, women in the ZOI and comparison groups had consumed foods from at least four of the nine key food groups on the day and night before the baseline survey (an indicator known as the Women's Dietary Diversity Score). Although there was no established norm for adequacy on this indicator, higher values indicate a higher probability of sufficient micronutrient intake for women and their breastfed children. Midline results showed that average dietary diversity using this indicator remained largely unchanged. Although we did not find evidence of a significant program impact on women's average dietary diversity using the household FE model for the overall sample, we did find that women in WHIP households with low baseline consumption consumed foods from 0.436 fewer food groups at midline relative to women from low baseline consumption in comparison households ($p<0.05$). Of note, since the WHIP evaluation began, a new indicator of Minimum Dietary Diversity for Women has been recommended for use. It involves minor changes to the Women's Dietary Diversity Score food groups, uses 10 groups instead of nine, and can be reported as the percentage of women consuming foods from at least 5 of the 10 groups, which is indicative of minimum dietary diversity (Leroy, Ruel, Frongillo, Harris, and Ballard, 2015).

Prevalence of underweight in women was very low in this population. At baseline, only 2.0 percent and 1.9 percent of women had a low BMI (were underweight) in the ZOI and comparison groups, respectively. At midline, the prevalence of underweight among women of reproductive age was slightly lower in the ZOI (1.2 percent) and slightly higher in the comparison group (2.3 percent). Results from the simple DID model showed that there was a marginally significant difference between the decreased prevalence of underweight in the ZOI compared with the change in the comparison group over the same period. Although we did not detect a significant program impact on underweight among women in the household FE model, we found a very similar coefficient estimate of -0.013 using the cluster FE model. This impact estimate was marginally significant at the 10 percent level, which may be due to the increased sample size in the cluster panel and the fact that the cluster FE model estimates fewer parameters (fewer FE) and is thus more efficient than the household FE model.

Secondary Impact by Program Component

Table 6 shows results for the secondary evaluation questions on food security and nutrition outcomes. The impact estimates were obtained from the household FE DID model, which was modified to enable the examination of impact in each of the three ZOI subgroups and to test hypotheses about relative effectiveness and externalities. These results are complemented by findings on trends presented in Appendix H.

Although the prevalence of food insecurity was low in the ZOI at baseline (14.4 percent overall, and 13.2 percent and 14.9 percent for the RVCP indirect and Health Only groups, respectively), it was less than half that level among RVCP direct beneficiaries (5.8 percent). Both the RVCP indirect and Health Only groups showed declines over time in food insecurity of between three to five percentage points but declines of similar magnitude were also observed in the comparison group. Accordingly, even after controlling for differences between the areas, the household FE model did not detect primary program impact on this indicator.

The program's impact on the prevalence of children receiving a minimum acceptable diet was not significant in any subgroup, but trend results presented in Appendix H indicate important declines on this indicator in all three subgroups; however, a similar decline was observed in the comparison group. Similarly, no program impact was detected for the prevalence of exclusive breastfeeding for the subgroups, but there were positive trends in the RVCP indirect and Health Only groups (Appendix H), indicating convergent trajectories at a higher level. The lack of significance in detecting impact is very likely due to increases that were also observed in the comparison group and small subgroup sample sizes.

As was the case for overall program impact, we did not detect significant impact, relative effectiveness, or externalities for under-five stunting and for wasting. It also appears that the significant overall program impact of 0.091 ($p<0.05$) on the prevalence of under-five underweight was driven largely by the 9.5 percentage point ($p<0.05$) impact in the Health Only group, which may be a reflection of the detrimental effect of the Government of Guatemala's decision to halt NGO health and nutrition service provision on this rural and vulnerable population.

We did not find evidence of subgroup impact, relative effectiveness, or externalities for women's dietary diversity. Last, we did see a significant negative impact of -0.026 ($p<0.05$) in underweight women among RVCP indirect beneficiaries. This effect should be interpreted with caution given the low prevalence of underweight in this subgroup.

Table 5. Main program impact on food security and nutrition

		Household Panel Summary Statistics				DID Impact Estimates		
		ZOI		Comparison		Household Panel		Cluster Panel
Outcome Indicator	N	Baseline	Midline	Baseline	Midline	Simple DID	Household FE	Cluster FE
Household food security								
Prevalence of households with moderate or severe hunger	7,544	0.144	0.098	0.167	0.118	0.003	0.012	-0.003
		(0.017)	(0.012)	(0.017)	(0.014)	(0.024)	(0.036)	(0.024)
Child health								
Exclusive breastfeeding in infants Ages 0-5 months old	700	0.766	0.826	0.734	0.760	0.034	-- [a]	0.019
		(0.049)	(0.041)	(0.061)	(0.061)	(0.089)	--	(0.133)
Prevalence of children ages 6-23 months receiving a minimum acceptable diet	1,363	0.415	0.248	0.345	0.239	-0.061	-0.101	-0.115
		(0.033)	(0.043)	(0.028)	(0.044)	(0.072)	(0.404)	(0.082)
Stunting	5,755	0.656	0.597	0.626	0.594	-0.027	0.038	-0.031
		(0.031)	(0.029)	(0.022)	(0.025)	(0.030)	(0.052)	(0.030)
Wasting	5,747	0.007	0.014	0.007	0.005	0.008	0.002	0.010
		(0.003)	(0.009)	(0.003)	(0.002)	(0.009)	(0.010)	(0.008)
Underweight	5,755	0.153	0.189	0.180	0.159	0.058+	0.091*	0.056+
		(0.019)	(0.022)	(0.016)	(0.016)	(0.029)	(0.040)	(0.029)
Women's health								
Dietary diversity	9,750	4.46	4.43	4.34	4.41	-0.109	-0.185	-0.167
		(0.091)	(0.091)	(0.069)	(0.085)	(0.128)	(0.170)	(0.111)
Underweight	9,383	0.020	0.012	0.019	0.023	-0.011+	-0.012	-0.013+
		(0.004)	(0.003)	(0.004)	(0.005)	(0.007)	(0.008)	(0.007)

Notes: Summary statistics are means and standard errors for the panel of households. The simple DID model was estimated among the panel of households and did not include FE or control variables. The household FE model was estimated among the panel of households and included FE at the household level and control variables. The cluster FE model was estimated among households in the panel of clusters and included FE at the cluster level and control variables. Standard errors adjusted for clustering at the cluster level are in parentheses. Control variables include a contemporaneous vector of household controls for household food security; for child health, included were household controls and mother's age; for child anthropometric outcomes, included were household controls and child age in months, child sex, caregiver sex, caregiver age and caregiver education; for women's health, included were household controls and woman's age and education. [†]This model was estimated on the number of households in the panel of clusters.

Significance levels: + p<0.10; * p<0.05; ** p<0.01; *** p<0.001

[a] The household FE model for exclusive breastfeeding in infants ages 0 to 5 months old was not appropriate for this sample of analysis. We present program estimates from the model using the panel of clusters.

Table 6. Secondary program impact on food security and nutrition

		Household Panel DID Impact Estimates				
Outcome Indicator	N	RVCP Direct (1)	RVCP Indirect (2)	Health Only (HO) (3)	Direct vs. HO (1)-(3)	Indirect vs. HO (2)-(3)
Household food security						
Prevalence of households with moderate or severe hunger	7,544	0.048	0.018	0.009	0.039	0.009
		(0.031)	(0.033)	(0.042)	(0.037)	(0.044)
Child health						
Prevalence of children ages 6-23 months receiving a minimum acceptable diet	1,363	-0.088	0.069	-0.177	0.089	0.246
		(0.508)	(0.558)	(0.470)	(0.551)	(0.653)
Exclusive breastfeeding in infants ages 0-5 months[a]	700	-0.018	0.129	0.041	-0.059	0.088
		(0.245)	(0.136)	(0.127)	(0.237)	(0.130)
Stunting	5,755	-0.019	0.002	0.047	-0.066	-0.045
		(0.071)	(0.055)	(0.060)	(0.078)	(0.062)
Wasting	5,747	-0.007	-0.017	0.006	-0.013	-0.023
		(0.012)	(0.012)	(0.012)	(0.015)	(0.014)
Underweight	5,755	0.068	0.069	0.095*	-0.027	-0.026
		(0.054)	(0.054)	(0.046)	(0.061)	(0.061)
Women's health						
Dietary diversity	9,750	-0.213	-0.142	-0.195	-0.018	0.053
		(0.167)	(0.193)	(0.190)	(0.174)	(0.202)
Underweight	9,383	-0.000	-0.026*	-0.012	0.012+	-0.013
		(0.008)	(0.011)	(0.008)	(0.007)	(0.012)

Notes: The household FE model was estimated using the panel of households and included controls for FE at the household level and observed control variables. Standard errors corrected for clustering at the cluster (*sector censal*) level are in parentheses. Control variables include a contemporaneous vector of household control variables (see Table 5 Notes).

[a] Exclusive breastfeeding in infants ages 0 to 5 months derived from the cluster panel sample using the cluster FE model due to the small sample sizes in the household panel for this indicator.

Significance levels: + p<0.10; * p<0.05; ** p<0.01; *** p<0.001

PROGRAM IMPACT ON REPRODUCTIVE HEALTH

Outcome Measures

The reproductive health indicators were non-FTF indicators and were calculated per the ENSMI methodology. The subpopulation of interest was women of reproductive age. The indicators were restricted to the woman's most recent live birth in the previous two years to compare equal baseline and midline time frames that did not overlap.

Family Planning – Use of Modern Contraceptive Methods

Family planning is an important proximate determinant of fertility. The baseline and midline EMEPAO surveys gathered information to help identify the prevalence of use of specific family planning methods, the most common supply sources, and unmet family planning needs. The prevalence of the use of modern contraceptive methods is considered a proxy for access to family planning services and is widely used to evaluate the performance of reproductive health programs. We report the indicator expressed as the proportion of married women or women in a consensual union ages 15 to 49 who, at the time of the survey, were using a modern contraceptive method. Modern contraceptive methods are female sterilization, male sterilization, IUD, injectables, implants, pills, condoms, and lactational amenorrhea.

Fertility – Births in the Previous Two Years

The association between fertility and poverty in developing countries is well established. Guatemala has one of the highest fertility rates in Latin America. The program's impact on fertility at midline was evaluated using the percentage of women of reproductive age with a birth in the two years before the interview.

Use of Maternal Health Services

Prenatal Care

The World Health Organization recommends that women have at least four prenatal checkups during their pregnancies. This practice contributes to the timely detection of possible obstetric problems and increases the probability of women receiving recommended prenatal services, such as tetanus toxoid vaccination, micronutrient supplementation, and nutrition and family planning counseling. We used a binary indicator of whether the woman received four or more prenatal checkups during her most recent pregnancy resulting in a live birth in the past two years.

Skilled Birth Attendance and Place of Delivery

Expanding access to skilled care at birth is one of the most important maternal health interventions because it has the potential to dramatically reduce maternal and neonatal mortality. We examined the impact of the WHIP on the proportion of deliveries taking place at health facilities and attended by skilled providers as part of the midline impact analysis.

Postnatal Care

Last, postnatal care is important to mothers and newborns' health because it allows for the detection and treatment of post-delivery complications and represents a crucial window of opportunity to provide newborn healthcare services and counseling to mothers. We reported program impact on whether women received a postnatal checkup within two days of delivery for their most recent live birth in the past two years.

Main Program Impact

Table 7 presents the results of the primary impact evaluation questions for reproductive health indicators. These results are complemented by findings on subgroups that are presented in Appendix F.

Family Planning and Fertility

At baseline, nearly half of women ages 15 to 49 in the ZOI (46 percent) and comparison group (45.1 percent) reported current use of a modern contraceptive method. Modern contraceptive method use declined in both groups to 37 percent at midline. We did not find evidence of program impact on modern contraceptive use prevalence in either the household FE or cluster FE models.

The percentage of women in the ZOI who reported giving birth in the two years immediately before the survey remained nearly unchanged from baseline to midline, at 22 percent but declined by four percentage points in the comparison group, from 26.1 percent to 22.1 percent. Results from the household FE model showed a positive significant impact of the WHIP on fertility, with women residing in ZOI households 6.1 percentage points more likely than women in comparison households to have had a birth in the previous two years ($p<0.05$). This positive impact was due to the decline in the comparison group, in contrast to unchanged fertility among women in program areas. The positive impact appeared to be driven by a more pronounced decline in fertility among women in low baseline consumption comparison households, from 33.0 percent at baseline to 23.2 percent at midline, and an impact of 9.7 percentage points ($p<0.01$) (Appendix F, Table F.1). We did not detect a significant impact in the household FE model for women in households consuming above the baseline median level.

Use of Maternal Health Services

Coverage of four or more prenatal checkups during the most recent pregnancy in the past two years was high in the ZOI at baseline, at 80.2 percent compared with 68.1 percent in comparison areas. Coverage rates decreased in both study groups between baseline and midline, by six percentage points and four percentage points in the comparison group and ZOI, respectively. We did not detect program impact on the prevalence of receiving four or more prenatal care checkups in the overall sample or in the baseline expenditure subgroups.

We also investigated whether the program had an impact on delivery practices. The prevalence of skilled birth attendance improved in the ZOI, from 41.7 percent to 51.6 percent, and in the comparison group, from 42.1 percent to 65.2 percent. Although the change over time in ZOI areas was significantly less than that in the comparison areas, we did not find significant program impact using the household FE model, and we did not find evidence of heterogeneous impact by baseline consumption level. We found a significant negative impact of 13.2 percentage points ($p<0.05$) using the cluster FE model, which can be explained by the greater improvement in skilled birth attendance in the comparison areas compared with in the ZOI. Given the similar

impact estimates from the household FE and cluster FE models, it is likely that the significant impact was detected using the cluster FE model because the cluster FE model is more efficient and used a larger sample of women than did the household FE model.

The percentage of births taking place in a health facility declined from 41.0 percent to 40.4 percent in the ZOI, but increased substantially, from 41.8 percent to 47.8 percent in the comparison areas. Despite the relatively large change in the comparison areas between baseline and midline, we did not detect evidence of significant program impact on the prevalence of health facility deliveries, nor did we find evidence of heterogeneous impact by baseline consumption level.

Last, the prevalence of timely postnatal care for mothers declined between baseline and midline in both program and comparison groups, from 83.1 percent to 78.9 percent in the ZOI, and from 79.2 percent to 71.0 percent in the comparison areas. Although the magnitude of the decrease was lower in the ZOI, the change over time was not significantly different between study areas. We did not detect program impact on the prevalence of postpartum care to mothers.

Secondary Impact by Program Component

Table 8 presents results related to evaluation questions 2.a through 2.d for reproductive health outcomes using the household FE DID model, which was modified to enable the examination of impact in each of the three ZOI subgroups and to test hypotheses about relative effectiveness and externalities. These results are complemented by findings on trends presented in Appendix H.

For current use of modern family planning methods, there were similar declines between baseline and midline in the three ZOI subgroups and in the comparison group. No program impact was found in any subgroup.

For the fertility indicator, there were declines in the two RVCP subgroups, but almost no change in the Health Only group, whereas there was a decline of four percentage points in the comparison group. The impact results for the different program beneficiary groups showed marginally significant coefficients on fertility. The results indicated that there was a marginally significant positive impact on fertility in the Health Only beneficiary group (6.5 percentage points, $p<0.10$), but we did not find evidence of differential effectiveness in the RVCP direct or indirect groups, or evidence of spillover effects in the RVCP indirect group.

The general trends indicated that all areas of the ZOI saw increases between baseline and midline in skilled birth attendance and facility delivery, with similar changes also occurring in the comparison group. The secondary impact estimation model did not detect any program effect, differential impact, or externalities on these indicators in ZOI subgroups.

Last, in the ZOI, trends in the data suggest that for the RVCP indirect beneficiaries and the Health Only groups, receipt of timely postpartum care decreased and it also declined in the comparison group. No significant impact was found for this indicator.

Table 7. Main program impact on reproductive health

		Household Panel Summary Statistics				DID Impact Estimates		
		ZOI		Comparison		Household Panel		Cluster Panel
Outcome Indicator	N	Baseline	Midline	Baseline	Midline	Simple DID	House-hold FE	Cluster FE
Family Planning								
Modern family planning prevalence	5,303	0.460	0.376	0.451	0.372	-0.005	-0.039	-0.009
		(0.031)	(0.027)	(0.023)	(0.028)	(0.038)	(0.065)	(0.031)
Fertility								
WRA with birth in past two years	9,863	0.226	0.224	0.261	0.221	0.038	0.061*	0.017
		(0.014)	(0.016)	(0.017)	(0.016)	(0.023)	(0.030)	(0.019)
Pregnancy outcomes								
At least four prenatal care visits	2,257	0.802	0.757	0.681	0.621	0.015	0.014	0.024
		(0.025)	(0.033)	(0.029)	(0.051)	(0.059)	(0.183)	(0.063)
Skilled birth attendance	2,256	0.417	0.516	0.421	0.652	-0.131*	-0.124	-0.132*
		(0.039)	(0.053)	(0.041)	(0.051)	(0.065)	(0.179)	(0.060)
Health facility delivery	2252	0.415	0.404	0.418	0.478	-0.071	-0.033	-0.070
		(0.039)	(0.044)	(0.041)	(0.048)	(0.056)	(0.145)	(0.054)
Postpartum care to mothers within two days of birth	1,488	0.831	0.789	0.792	0.710	0.040	-0.107	-0.019
		(0.031)	(0.037)	(0.029)	(0.050)	(0.070)	(0.469)	(0.085)

Notes: Summary statistics are means and standard errors calculated for the panel of households. The simple DID model was estimated among the panel of households and did not control for FE or control variables. The household FE model was estimated among the panel of households and controls for FE at the household level and control variables. The cluster FE model was estimated among households in the panel of clusters and controls for FE at the cluster level and control variables. Standard errors corrected by clustering at the cluster (*sector censal*) level are in parentheses. Control variables comprised a contemporaneous vector of household controls and the woman's age. For family planning and fertility, the woman's education was also included. †This model was estimated on the number of households in the panel of clusters.

Significance levels: + p<0.10; * p<0.05; ** p<0.01; *** p<0.001

Table 8. Secondary program impact on reproductive health

Outcome Indicator	N	Household Panel DID Impact Estimates: RVCP Direct (1)	RVCP Indirect (2)	Health Only (3)	Direct vs. HO (1)-(3)	Indirect vs. HO (2)-(3)
Family Planning						
Modern family planning prevalence	5,303	-0.052	-0.024	-0.043	-0.009	0.019
		(0.079)	(0.065)	(0.075)	(0.089)	(0.080)
Fertility						
WRA with birth in past two years	9,863	0.043	0.047	0.065+	-0.022	-0.018
		(0.031)	(0.031)	(0.035)	(0.041)	(0.039)
Pregnancy outcomes						
At least four prenatal care visits	2,257	0.001	-0.090	0.038	-0.037	-0.128
		(0.238)	(0.245)	(0.192)	(0.220)	(0.223)
Skilled birth attendance	2,256	-0.138	-0.057	-0.139	0.001	0.082
		(0.220)	(0.236)	(0.192)	(0.197)	(0.227)
Health facility delivery	2252	-0.105	-0.037	-0.031	-0.074	-0.006
		(0.205)	(0.171)	(0.154)	(0.174)	(0.151)
Postpartum care to mothers within two days of birth	1,488	-0.083	-0.225	-0.066	-0.017	-0.159
		(0.650)	(0.614)	(0.477)	(0.482)	(0.480)

Notes: The household FE model was estimated using the panel of households and it controls for FE at the household level and control variables. Standard errors corrected by clustering at the cluster level are in parentheses. Control variables comprised contemporaneous vector of household- and individual-level variables.

Significance levels: + p<0.10; * p<0.05; ** p<0.01; *** p<0.001

WHIP IMPACT EVALUATION LIMITATIONS

There are three important limitations to the WHIP midline impact evaluation that warrant discussion: 1) the short time interval between the baseline and midline surveys; 2) the limited implementation of the health and nutrition intervention component due to the 2013 Guatemalan Congressional budget law; and 3) midline data quality issues.

The first limitation of the midline impact evaluation is the short length of time for which the WHIP was implemented between baseline and midline, approximately two years. As described in the 2013 baseline report (Angeles, et al., 2014), the midline impact evaluation was intended to serve monitoring purposes to check whether indicators were moving in the expected directions, although it was hypothesized that some significant effects could be detected in consumption, health, and nutrition behaviors and service use indicators in such a short time period. As such, the impact evaluation was not powered at baseline to detect significant changes after only two years, but rather, was powered to detect changes over the life of the program from baseline to 2017 endline. In recognition of the expectation that key higher-level outcomes, such as household poverty and woman and child nutritional status take longer to change, the lack of statistically significant program impact at midline should not be interpreted to mean that the WHIP was not achieving desired results. Instead, trends over time in the ZOI and comparison groups provide evidence about whether these indicators were changing as expected by the program's theory of change. For example, we did find that, despite a lack of significant program impact, the household consumption and poverty indicators were evolving as expected, with increased consumption in ZOI households and decreased prevalence of poverty and depth of poverty in ZOI households.

The second and most critical limitation of the midline impact evaluation was that the health and nutrition program was not implemented as planned. In fact, health services and supply chains were shut down due to the 2013 national budget law passed by the Guatemalan Congress. This potentially affected both behavioral and health service use indicators. Behavioral indicators, such as child feeding practices and women's diet diversity, were hypothesized to respond to both the income effect of the RVCP and the nutrition education/counselling services in the health and nutrition program. The service use indicators, such as pregnancy, delivery, and postpartum care, were hypothesized to respond primarily to increased service provision and improved service quality components of the health and nutrition program.

The impact evaluation was also affected by midline data quality issues, which are described in Appendix J. Although the data quality assessment did not uncover any "fatal flaws," it is important to disclose these concerns and discuss possible implications for impact results. One important concern relates to the quality of the household panel. We found a high attrition rate, approximately 40 percent among panel households but concluded that there were no major problems with selective attrition. After identifying the household panel based on household identification variables, we conducted basic data quality checks to determine whether certain household characteristics (e.g., sex, indigenous status, primary language, primary occupation, and religion of the household head, household size) had changed in unexpected ways over time, which could be an indication of household identification and verification problems. To be clear, it is possible that these characteristics could legitimately change over two years, especially if there was a different household questionnaire respondent across survey rounds in a household. However, as explained in Appendix J, the magnitude of some of the changes was higher than expected. The implications are therefore difficult to assess because conclusions about program impact using the household panel are valid if the changes over time are true and we are following the same households over time. Conversely, if the large number of changes observed indicates that different households participated at midline, then the effects on program impact

estimates would be similar to measurement error in that we expect the effects to be attenuated (e.g., we may underestimate the magnitude of program impact) and it may be more difficult to detect significant impact as a discordance between baseline and midline households would increase the variance of our estimates. Assessing both household FE in the household panel and cluster FE in the panel of clusters is one way to assess the robustness of our impact estimates to potential problems with whether households in the household panel were the same over time.

A final important data quality issue was that there was some difficulty obtaining birth dates for children in the women's questionnaire birth history with certainty. The date of birth and date of interview variables were used to create the child's age in months, which is critical for restricting appropriate subsamples for child feeding indicators, such as exclusive breastfeeding (0 to 5 months) and minimum acceptable diet (6 to 23 months), and births occurring in the previous two years (for fertility, prenatal care, delivery care and location, and postnatal care indicators). As explained in Appendix J, we do not think that this issue affects a significant proportion of observations, especially because our analysis for these indicators was restricted to the woman's most recent birth. However, it is nevertheless important to discuss because it could lead to biased monitoring results and time trends. For example, the exclusive breastfeeding indicator was restricted to a small time frame, and if a child's age in months is off by even one to two months, bias could be introduced in the indicator. If we were underestimating the children's age in months and including children over six months of age in the exclusive breastfeeding indicator, we are likely to find a lower than expected prevalence of exclusive breastfeeding because children over six months of age are being given complementary feeding. Although this has implications for monitoring trends over time, we do not expect the age data quality issue to bias impact estimates because we do not have reason to expect that the data quality issues that likely affect the age-in-months variable are systematically different between ZOI and comparison groups.

CONCLUSIONS

The main objectives of the 2015 midline impact evaluation are to estimate changes in key indicators that are attributable to the WHIP RVCP and health and nutrition interventions. The primary evaluation question focuses on the direct impact of the WHIP on key indicators at the population level in the ZOI. Secondary evaluation questions examine the impact of the integrated RVCP and health and nutrition program, the impact of the health and nutrition program without the RVCP, the relative effectiveness of the integrated program compared with the health and nutrition intervention only, and whether the RVCP has indirect effects (spillovers or externalities) on nonmember households located in RVCP areas. The impact evaluation is based on a quasi-experimental DID design with a matched comparison group and controls for household fixed effects. As such, our preferred impact estimation strategy is to use household FE with pooled data from the panel of households, and we also implement cluster FE in the panel of clusters as a robustness check.

From the household FE results, we did not find significant primary or secondary effects on household expenditures, prevalence of poverty, or depth of poverty at midline, but we did observe that impact is occurring in the expected direction (increased expenditures and decreased prevalence and depth of poverty). Similarly, although we did not detect a significant impact of the WHIP on the prevalence of households with moderate or severe hunger, we observe decreased prevalence of hunger in both the ZOI and comparison areas at midline relative to baseline.

We did not detect significant primary or secondary program impact on infant and young child feeding practices or dietary diversity among women of reproductive age at midline. Likewise, we did not find evidence of program impact on under-five stunting or wasting prevalence, but did detect a positive impact of 9.1 percentage points ($p<0.05$) on the prevalence of underweight, which appears to be driven by an impact of 9.5 percentage points ($p<0.05$) in the Health Only group. The association between the WHIP and increased incidence of global malnutrition may be partially attributable to the lack of health and nutrition services in the ZOI, highlighting the importance of outreach, education, and service provision in improving the nutritional status of young children in the rural Western Highlands. Although we did not see a significant overall program impact on the prevalence of underweight women, we found that the prevalence of underweight women was 2.6 percentage points lower among RVCP indirect beneficiaries relative to women in comparison households ($p<0.05$). However, the prevalence of underweight among women was very low in this population.

Last, we found no evidence of program impact on the use of reproductive health services, including modern family planning methods and the use of maternal health services. We did see a WHIP impact estimate of a 6.1 percentage point increase in the percentage of women with a live birth in the two years before the survey ($p<0.05$), but this result was mainly driven by a decline in the fertility indicator in the comparison group.

Taken together, the results of the midline impact evaluation indicate that although there were no statistically significant program effects on household consumption, poverty, or hunger, these indicators were moving in the expected direction consistent with the program's theory of change. Mixed results in time trends for infant and young child feeding practices, nutritional status, and the decreasing use of reproductive and maternal health services suggest that the cessation of the health and nutrition program in 2013/2014 had detrimental effects on these indicators in ZOI areas.

REFERENCES

Angeles, G., Hidalgo, E., Molina-Cruz, R., Taylor, T., Urquieta-Salomón, J., Calderón, C…Romero, M. (2014). *Monitoring and evaluation survey for the Western Highlands Integrated Program: Baseline 2013* (TR-14_100). Chapel Hill, NC: MEASURE Evaluation, Carolina Population Center. Retrieved from https://www.measureevaluation.org/resources/publications/tr-14-100.

Feed the Future (2016). *Feed the Future indicator handbook definition sheets*. Washington, DC: USAID. Retrieved from https://feedthefuture.gov/sites/default/files/resource/files/Feed_the_Future_Indicator_Handbook_25_July_2016.pdf.

Leroy, J.L., Ruel, M., Frongillo, E.A., Harris, J., and Ballard, T.J. (2015). Measuring the food access dimension of food security: a critical review and mapping of indicators. *Food and Nutrition Bulletin*, 36(2), 167-195. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/0379572115587274.

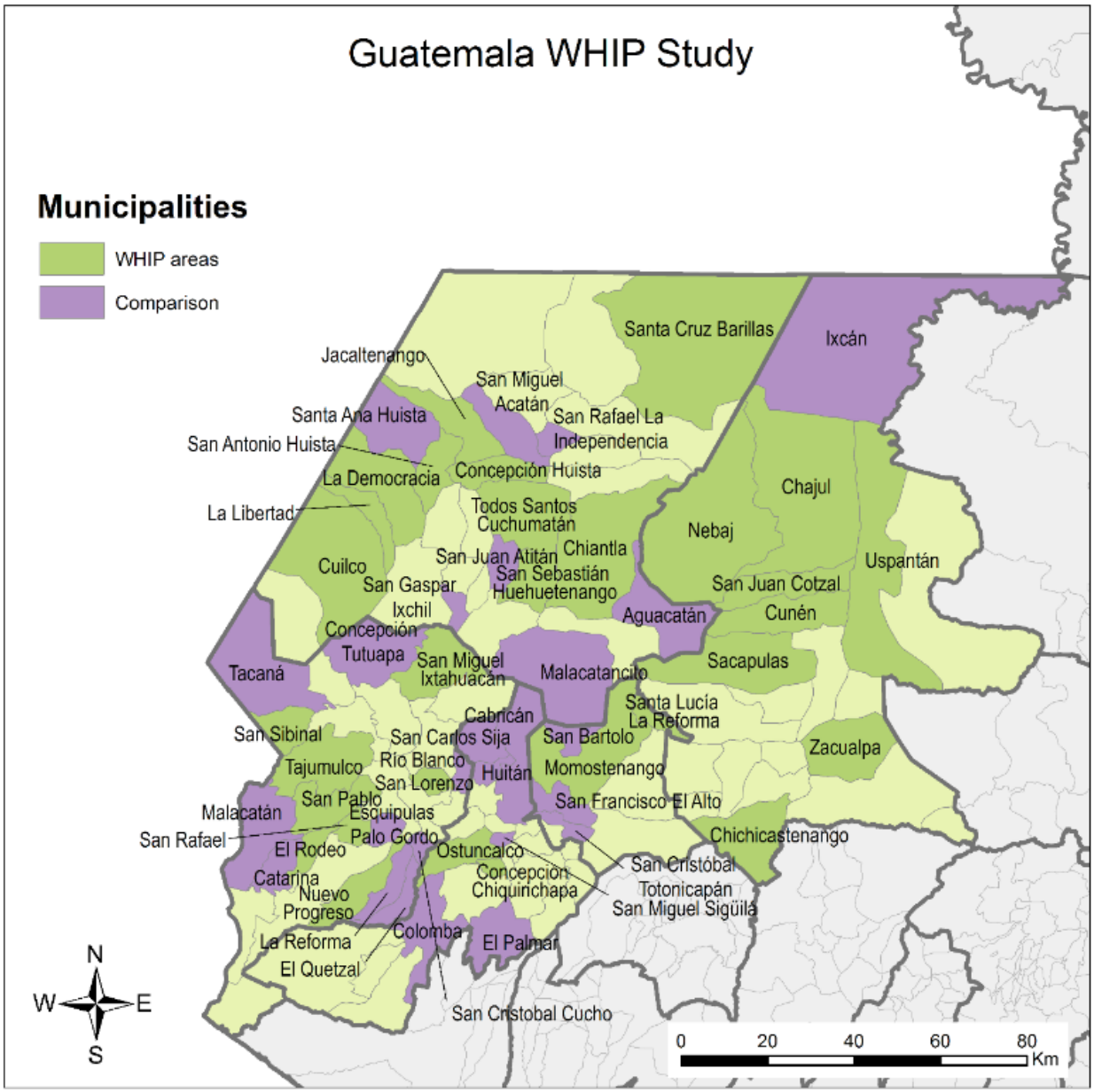
APPENDIX A. LIST OF PRIORITY WHIP MUNICIPALITIES

Table A.1. Priority WHIP municipalities

Municipality	Department	2013 population (estimated)
Cunén	Quiché	37,473
Nebaj	Quiché	88,542
Sacapulas	Quiché	48,428
Uspantán	Quiché	69,462
Chajul	Quiché	55,438
San Juan Cotzal	Quiché	28,692
Zacualpa	Quiché	49,258
Chichicastenango	Quiché	152,833
San José el Rodeo	San Marcos	17,295
San Rafael Pie de la Cuesta	San Marcos	15,978
San Lorenzo	San Marcos	12,198
San Pablo	San Marcos	54,659
Tajumulco	San Marcos	58,409
Nuevo Progreso	San Marcos	37,954
Sibinal	San Marcos	16,585
San Miguel Ixtahuacán	San Marcos	37,303
Jacaltenango	Huehuetenango	45,458
Chiantla	Huehuetenango	95,986
San Sebastián Huehuetenango	Huehuetenango	29,930
Todos Santos	Huehuetenango	36,009
Santa Cruz Barrillas	Huehuetenango	147,314
Cuilco	Huehuetenango	60,306
Concepción Huista	Huehuetenango	19,154
San Antonio Huista	Huehuetenango	18,641
La Libertad	Huehuetenango	39,048
La Democracia	Huehuetenango	45,201
Momostenango	Totonicapan	132,854
Santa Lucía La Reforma	Totonicapán	23,231
San Juan Ostuncalco	Quetzaltenango	53,687
Concepción Chiquirichapa	Quetzaltenango	18,437
Total		**1,545,765**

APPENDIX B. MAP OF ZOI AND COMPARISON AREAS INCLUDED IN THE MIDLINE IMPACT EVALUATION

Figure B.1. Map of municipalities included in the WHIP midline impact evaluation



APPENDIX C. ATTRITION TABLES

C.1. Overall Attrition

Table C.1.1. Roster characteristics

	Individuals lost to follow-up		Panel		Mean	Diff	
Variables	Mean	N1	Mean	N2	Diff	SE	p-value
Age	22.289	9,572	22.668	14,209	0.380	0.556	0.495
Female	0.532	9,574	0.519	14,210	-0.013	0.008	0.114
Married or in consensual union	0.626	5,273	0.633	8,188	0.007	0.014	0.631
Years of schooling, age 18 and older	2.311	4,546	2.462	6,996	0.151	0.098	0.128
No education, age 18 and older	0.372	4,557	0.357	7,013	-0.015	0.019	0.418
Completed or not completed primary education, age 18 and older	0.488	4,560	0.482	7,014	-0.007	0.020	0.734
Completed or not completed secondary education, age 18 and older	0.124	4,560	0.144	7,014	0.020	0.019	0.288
Cannot read and write	0.324	5,279	0.305	8,192	-0.019	0.017	0.284
Member ages 5-18 currently studying	0.752	3,225	0.752	4,882	-0.001	0.020	0.970

Table C.1.2. Household characteristics

	Attriters		Panel		Mean	Diff	p-
Variables	Mean	N1	Mean	N2	Diff	SE	value
Total number of individuals in household	5.748	1,627	5.820	2,380	0.072	0.197	0.715
Number of adults ages 18-64 years	2.526	1,627	2.560	2,380	0.034	0.071	0.630
Child under 5 in household	0.571	1,627	0.564	2,380	-0.007	0.031	0.810
Child age 5-17 in household	0.776	1,627	0.777	2,380	0.001	0.026	0.954
Member age 65 or older in household	0.174	1,627	0.180	2,380	0.006	0.022	0.772
Household has WRA	0.921	1,627	0.916	2,380	-0.005	0.016	0.755
Single woman	0.216	1,627	0.180	2,380	-0.036	0.021	0.087
Household in rural area	0.812	1,627	0.752	2,380	-0.060	0.042	0.156
Municipality	9.774	1,627	10.569	2,380	0.795	0.777	0.307
Altitude in meters	1,737.145	1,627	1,662.660	2,380	-74.485	108.475	0.493
Used translator during interview	0.176	1,627	0.191	2,380	0.016	0.036	0.668
Household interview duration in hours	33.306	1,624	32.688	2,378	-0.618	1.229	0.616

Table C.1.3. Household head characteristics

Variables	Attriters Mean	Attriters N1	Panel Mean	Panel N2	Mean Diff	Diff SE	p-value
Age	43.490	1,626	43.907	2,380	0.416	0.873	0.634
Female	0.219	1,627	0.196	2,380	-0.023	0.021	0.279
Married or in consensual union	0.861	1,626	0.877	2,380	0.016	0.015	0.288
Years of schooling, age 18 and older	2.138	1,621	2.300	2,369	0.163	0.127	0.202
No education, age 18 and older	0.381	1,627	0.348	2,378	-0.034	0.027	0.221
Completed or not completed primary education, age 18 and older	0.538	1,627	0.550	2,378	0.011	0.031	0.719
Completed or not completed secondary education, age 18 and older	0.061	1,627	0.079	2,378	0.019	0.015	0.230
Cannot read and write	0.347	1,626	0.319	2,380	-0.029	0.026	0.280

Table C.1.4. Housing characteristics/items

Variables	Attriters Mean	Attriters N1	Panel Mean	Panel N2	Mean Diff	Diff SE	p-value
Dirt or sand floors	0.535	1,627	0.450	2,380	-0.085	0.033	0.011
Cement floors	0.337	1,627	0.402	2,380	0.065	0.028	0.020
Corr. zinc or metal roof	0.763	1,627	0.760	2,380	-0.003	0.039	0.940
Concrete, reinforced concrete, ceramic roof	0.120	1,627	0.119	2,380	-0.001	0.023	0.982
Cinder block walls	0.346	1,627	0.433	2,380	0.087	0.035	0.013
Adobe or covered adobe walls	0.395	1,627	0.327	2,380	-0.068	0.048	0.159
Household has only one bedroom	0.524	1,622	0.479	2,378	-0.045	0.032	0.159
Crowding index (household size/ number of bedrooms)	3.855	1,622	3.870	2,378	0.015	0.153	0.924
Household has electricity	0.808	1,604	0.839	2,365	0.031	0.026	0.242
Connected to public water system	0.741	1,627	0.767	2,380	0.026	0.039	0.510
Improved water source or appropriately treated unimproved source	0.965	1,627	0.958	2,380	-0.007	0.008	0.360
Observed location with water and soap for handwashing	0.258	1,627	0.265	2,380	0.007	0.020	0.719
Has latrine/outhouse/cesspits	0.636	1,627	0.567	2,380	-0.068	0.042	0.106
Connected to sewage system	0.190	1,627	0.257	2,380	0.066	0.034	0.055
Improved sanitation	0.797	1,627	0.829	2,380	0.033	0.022	0.142
Room used exclusively for cooking	0.621	1,627	0.666	2,380	0.045	0.035	0.201
Firewood main source of cooking fuel	0.932	1,627	0.935	2,380	0.003	0.015	0.815
Television	0.570	1,227	0.591	1,970	0.021	0.033	0.535
Radio	0.468	1,227	0.433	1,970	-0.035	0.035	0.323
DVD player	0.158	1,227	0.175	1,970	0.016	0.019	0.394
CD player	0.131	1,227	0.198	1,970	0.067	0.024	0.005
Computer	0.066	1,227	0.087	1,970	0.021	0.016	0.197
Printer	0.041	1,227	0.050	1,970	0.009	0.010	0.370
Video player	0.038	1,227	0.048	1,970	0.010	0.012	0.432
Still camera	0.034	1,227	0.033	1,970	-0.001	0.008	0.887
Video camera	0.011	1,227	0.019	1,970	0.009	0.006	0.133
Pickup truck	0.074	1,604	0.071	2,365	-0.003	0.015	0.826
Moto	0.041	1,604	0.065	2,365	0.025	0.013	0.059
Bike	0.109	1,604	0.115	2,365	0.006	0.019	0.732
Home is owned and fully paid for	0.847	1,604	0.896	2,365	0.048	0.033	0.140
Member left for employment in past 12 months	0.291	1,627	0.288	2,380	-0.002	0.040	0.952
Household received cash remittance from family in foreign country	0.103	1,627	0.130	2,380	0.027	0.020	0.186

Table C.1.5. Program indicators

Variables	Attriters Mean	Attriters N1	Panel Mean	Panel N2	Mean Diff	Diff SE	p-value
Received gov't benefits through fertilizer program in past 12 months	0.249	1,627	0.272	2,380	0.023	0.039	0.560
Received gov't benefits through Bono Seguro program in past 12 months	0.240	1,627	0.267	2,380	0.027	0.034	0.428
Have heard about USAID or can identify its logo	0.171	1,614	0.211	2,350	0.040	0.020	0.046
Reports USAID provides services	0.314	391	0.272	636	-0.041	0.062	0.503

Table C.1.6. Consumption by expense type

Variables	Attriters Mean	Attriters N1	Panel Mean	Panel N2	Mean Diff	Diff SE	p-value
Food and drink	9.938	1,604	9.859	2,365	-0.080	0.346	0.819
Meals outside the home	0.370	1,604	0.449	2,365	0.080	0.089	0.372
Housing	2.558	1,604	3.033	2,365	0.475	0.230	0.040
Household services	2.750	1,604	2.847	2,365	0.097	0.156	0.536
Education	1.258	1,604	1.116	2,365	-0.142	0.217	0.512
Health	3.183	1,604	3.144	2,365	-0.039	0.334	0.907
Household equipment, furnishings	0.520	1,604	0.657	2,365	0.137	0.098	0.165
Donations	0.072	1,604	0.067	2,365	-0.005	0.022	0.806
Household items, house cleaning and maintenance, and household appliances	1.448	1,604	1.499	2,365	0.051	0.115	0.659
Taxes, funeral, pensions	0.723	1,604	0.682	2,365	-0.041	0.138	0.767
Recreation, entertainment, and tourism	0.175	1,604	0.237	2,365	0.062	0.064	0.329
Personal care	0.781	1,604	0.820	2,365	0.039	0.053	0.463
Clothing and shoes	0.532	1,604	0.579	2,365	0.047	0.041	0.255
Other	1.596	1,604	1.836	2,365	0.240	0.162	0.139

Table C.1.7. Consumption, item share

	Attriters		Panel		Mean	Diff	p-
Variables	Mean	N1	Mean	N2	Diff	SE	value
Food and drink	43.681	1,604	41.633	2,365	-2.048	1.151	0.077
Meals outside the home	1.048	1,604	1.086	2,365	0.037	0.207	0.856
Housing	10.166	1,604	10.966	2,365	0.800	0.425	0.061
Household services	11.646	1,604	11.825	2,365	0.179	0.367	0.626
Education	4.018	1,604	4.039	2,365	0.021	0.262	0.937
Health	8.706	1,604	9.053	2,365	0.348	0.606	0.567
Household equipment, furnishings	1.293	1,604	1.501	2,365	0.208	0.188	0.270
Donations	0.318	1,604	0.257	2,365	-0.061	0.080	0.448
Household items, house cleaning and maintenance, and household appliances	5.233	1,604	5.252	2,365	0.019	0.234	0.935
Taxes, funeral, pensions	1.936	1,604	1.851	2,365	-0.086	0.201	0.670
Recreation, entertainment, and tourism	0.532	1,604	0.473	2,365	-0.059	0.099	0.552
Personal care	1.919	1,604	2.001	2,365	0.082	0.097	0.402
Clothing and shoes	3.302	1,604	3.371	2,365	0.068	0.177	0.700
Other	6.201	1,604	6.694	2,365	0.493	0.425	0.248

Table C.1.8. Prevalence of poverty, individual level

	Attriters		Panel		Mean	Diff	p-
Variables	Mean	N1	Mean	N2	Diff	SE	value
Extreme poverty line (13.18 Quetzales)	0.289	1,604	0.260	2,365	-0.029	0.035	0.418
Total poverty line (27.17 Quetzales)	0.778	1,604	0.752	2,365	-0.026	0.023	0.262
USD 1.25 poverty line	0.054	1,604	0.061	2,365	0.007	0.017	0.659
USD 2.00 poverty line	0.318	1,604	0.285	2,365	-0.033	0.033	0.313

Table C.1.9. Per capita consumption

	Attriters		Panel		Mean	Diff	p-
Variables	Mean	N1	Mean	N2	Diff	SE	value
Daily per capita consumption (Quetzales)	22.374	1,604	23.197	2,365	0.822	1.049	0.434
Daily per capita consumption (2010 USD)	3.608	1,604	3.740	2,365	0.133	0.169	0.434

Table C.1.10. Nutrition indicators, anthropometrics, among children under 5 years of age

	Attriters		Panel		Mean	Diff	p-
Variables	Mean	N1	Mean	N2	Diff	SE	value
Moderate chronic malnutrition (stunted)	0.723	1,329	0.653	1,881	-0.071	0.033	0.034
Severe chronic malnutrition	0.390	1,329	0.270	1,881	-0.120	0.041	0.004
Moderate acute malnutrition (wasted)	0.004	1,321	0.009	1,872	0.005	0.005	0.277
Severe global malnutrition	0.000	1,321	0.000	1,872	0.000	0.000	
Moderate global malnutrition (underweight)	0.203	1,325	0.158	1,878	-0.045	0.028	0.112
Severe acute malnutrition	0.037	1,325	0.018	1,878	-0.019	0.009	0.039
Overweight (high BMI for age)	0.065	1,329	0.050	1,881	-0.014	0.014	0.295

Table C.1.11. Nutrition indicators, breastfeeding ages 0 to 5 months

	Attriters		Panel		Mean	Diff	p-
Variables	Mean	N1	Mean	N2	Diff	SE	value
No breastfeeding	0.006	129	0.001	189	-0.005	0.006	0.394
Exclusive breastfeeding	0.616	129	0.697	189	0.080	0.098	0.413
Breastfeeding and other liquids	0.318	129	0.272	189	-0.046	0.094	0.627
Breastfeeding and complementary feeding	0.059	129	0.030	189	-0.029	0.033	0.383

Table C.1.12. Nutrition indicators, breastfeeding ages 6 to 23 months

	Attriters		Panel		Mean	Diff	p-
Variables	Mean	N1	Mean	N2	Diff	SE	value
No breastfeeding	0.162	407	0.161	560	-0.001	0.037	0.981
Exclusive breastfeeding	0.025	407	0.014	560	-0.011	0.014	0.452
Breastfeeding and other liquids	0.009	407	0.009	560	-0.000	0.009	0.968
Breastfeeding and complementary feeding	0.804	407	0.816	560	0.012	0.042	0.779

Table C.1.13. Nutrition indicators, supplements

	Attriters		Panel		Mean	Diff	p-
Variables	Mean	N1	Mean	N2	Diff	SE	value
Took iron in the previous 7 days	0.209	1,358	0.217	1,904	0.008	0.032	0.794
Took vitamin A during the preceding 6 months	0.588	1,358	0.651	1,904	0.063	0.034	0.061
Took deworming medication during the preceding 6 months	0.412	1,358	0.420	1,904	0.008	0.027	0.762

Table C.1.14. Nutrition indicators, minimum acceptable diet

	Attriters		Panel		Mean	Diff	p-
Variables	Mean	N1	Mean	N2	Diff	SE	value
Acceptable feeding frequency	69.779	407	72.658	560	2.879	4.710	0.542
Acceptable dietary diversity	48.267	407	49.155	560	0.887	4.648	0.849
Minimum acceptable diet	33.984	407	42.873	560	8.889	3.942	0.025

Table C.1.15. Nutrition indicators, lactation initiation

	Attriters		Panel		Mean	Diff	p-
Variables	Mean	N1	Mean	N2	Diff	SE	value
Time of first breastfeeding							
Within first hour after birth	0.566	984	0.562	1,408	-0.004	0.036	0.901
After the first hour after birth	0.316	984	0.296	1,408	-0.020	0.033	0.541
After the first day	0.107	984	0.130	1,408	0.023	0.019	0.218
Never breastfed	0.010	984	0.012	1,408	0.002	0.005	0.712

Table C.1.16. Nutrition indicators, woman level

	Attriters		Panel		Mean	Diff	p-
Variables	Mean	N1	Mean	N2	Diff	SE	value
Height less than 145 cm	0.428	1,995	0.424	3,157	-0.004	0.028	0.875
Height 145 cm or more	0.572	1,995	0.576	3,157	0.004	0.028	0.875
Low BMI (<18.5) in 15-49 year-old women	2.961	1,995	1.986	3,157	-0.975	0.713	0.173
Overweight BMI	40.778	1,995	44.041	3,157	3.263	2.338	0.164
Acceptable dietary diversity	0.701	2,176	0.741	3,390	0.040	0.029	0.168
Dietary diversity in women: Average number of food groups ingested	4.358	2,176	4.472	3,390	0.113	0.092	0.219

Table C.1.17. Nutrition indicators, household-level women's nutrition

	Attriters		Panel		Mean	Diff	p-
Variables	Mean	N1	Mean	N2	Diff	SE	value
Prevalence of households with moderate or severe hunger	0.128	1,627	0.143	2,379	0.015	0.015	0.325
Produces food crops for household use	0.118	1,627	0.123	2,380	0.005	0.032	0.878
Does not produce but has available lot or land	0.435	1,627	0.447	2,380	0.012	0.027	0.647
Does not produce and has no available lot or land	0.445	1,627	0.430	2,380	-0.015	0.035	0.680
Believes malnutrition affects his/her household	0.519	1,627	0.528	2,380	0.009	0.034	0.784
Feels malnutrition is a serious problem in his/her community	0.576	1,627	0.614	2,380	0.038	0.027	0.162

Table C.1.18. Maternal health, woman-level

Variables	Attriters Mean	Attriters N1	Panel Mean	Panel N2	Mean Diff	Diff SE	p-value
Has child, among women ages 18-24 years	0.581	625	0.505	977	-0.076	0.044	0.085
Age at first birth among woman ages 18-24 was before age 18	0.416	374	0.400	513	-0.016	0.062	0.797
Named two or more key signs of prenatal risk	0.610	2,176	0.612	3,390	0.002	0.031	0.944
Named two or more key signs of delivery risks	0.648	2,176	0.646	3,390	-0.002	0.023	0.931
Named two or more key signs of postpartum risks	0.622	2,176	0.628	3,390	0.006	0.022	0.796
Named two or more key signs of newborn risks	0.477	2,176	0.494	3,390	0.017	0.035	0.628

Table C.1.19. Maternal health, pregnancy-level

Variables	Attriters Mean	Attriters N1	Panel Mean	Panel N2	Mean Diff	Diff SE	p-value
Received any prenatal care	0.915	1,004	0.940	1,428	0.025	0.020	0.198
Received four or more prenatal checkups	0.696	1,005	0.798	1,429	0.102	0.031	0.001
Skilled provider attended birth	0.909	1,408	0.939	1,960	0.030	0.028	0.285
Physician, ambulatory physician, or nurse attended birth	0.344	1,408	0.366	1,960	0.022	0.039	0.576
Place of delivery was facility	0.339	1,406	0.365	1,960	0.026	0.036	0.477
Received postpartum care (postnatal woman)	0.733	1,004	0.721	1,427	-0.013	0.032	0.695
Had plan to ensure transportation	53.151	324	56.377	529	3.226	5.326	0.545
Had plan to save money	78.683	324	72.724	529	-5.960	4.597	0.196
Had identified blood donors	7.713	324	6.834	529	-0.880	2.507	0.726
Had selected a place for the delivery	60.239	324	52.221	529	-8.018	6.544	0.222

Table C.1.20. Fertility and family planning

Variables	Attriters Mean	Attriters N1	Panel Mean	Panel N2	Mean Diff	Diff SE	p-value
Average number of live births	2.606	2,176	2.505	3,390	-0.102	0.117	0.386
Average number of living children	2.207	2,330	2.228	3,508	0.021	0.109	0.850
Uses any contraceptive method	0.443	1,404	0.525	2,098	0.081	0.029	0.006
Uses any modern contraceptive method	0.349	1,404	0.414	2,098	0.065	0.030	0.034
Unmet need	0.184	1,404	0.167	2,098	-0.017	0.022	0.432
Total demand for family planning	0.628	1,404	0.692	2,098	0.064	0.029	0.030

Table C.1.21. Children's health

Variables	Attriters Mean	Attriters N1	Panel Mean	Panel N2	Mean Diff	Diff SE	p-value
Received postnatal care	0.827	1,004	0.823	1,425	-0.003	0.029	0.906
Pentavalent 1 (12-23 months)	0.978	293	0.985	384	0.008	0.018	0.663
Pentavalent 2 (12-23 months)	0.968	293	0.979	384	0.011	0.020	0.574
Pentavalent 3 (12-23 months)	0.947	293	0.958	384	0.012	0.030	0.699
MMR/measles (12-23 months)	0.895	293	0.871	384	-0.024	0.050	0.637
Pentavalent booster 1 (12-23 months)	0.414	293	0.397	384	-0.016	0.059	0.782
Pentavalent 1 (24-59 months)	0.982	802	0.992	1,129	0.009	0.005	0.087
Pentavalent 2 (24-59 months)	0.974	802	0.985	1,129	0.012	0.010	0.242
Pentavalent 3 (24-59 months)	0.948	802	0.964	1,129	0.017	0.016	0.304
All three pentavalent vaccines (12-59)	0.914	1,127	0.935	1,561	0.021	0.016	0.200
MMR/measles (24-59 months)	0.957	802	0.973	1,129	0.015	0.012	0.187
Pentavalent booster 1 (24-59 months)	0.830	802	0.864	1,129	0.034	0.030	0.251
Pentavalent booster 2 (24-59 months)	0.198	802	0.241	1,129	0.043	0.022	0.054
All vaccines (24-59 months)	0.198	802	0.241	1,129	0.043	0.022	0.054
Had diarrhea in the past two weeks	0.216	1,358	0.220	1,904	0.004	0.024	0.881
Was taken to a health facility	0.351	287	0.461	411	0.110	0.057	0.054
Was given ORS or zinc	0.374	287	0.426	411	0.052	0.056	0.354
Had acute respiratory infection in the past two weeks	0.161	1,358	0.145	1,904	-0.017	0.023	0.460
Was taken to a health facility	0.553	219	0.664	286	0.112	0.072	0.121
Was administered antibiotics or other medication	0.869	219	0.843	286	-0.026	0.043	0.554

Table C.1.22. Participation in agricultural activities among household members 12 and older

Variables	Attriters Mean	Attriters N1	Panel Mean	Panel N2	Mean Diff	Diff SE	p-value
Household members 12 years or over who work in agriculture	26.172	6,180	27.580	9,517	1.408	1.369	0.305

Table C.1.23. Participation in agricultural activities, household

Variables	Attriters Mean	Attriters N1	Panel Mean	Panel N2	Mean Diff	Diff SE	p-value
Member age 12 or older is farmer	0.439	1,627	0.509	2,380	0.069	0.030	0.023
Member age 12 or older is farm laborer	0.251	1,627	0.269	2,380	0.018	0.031	0.575
Household produces food for household consumption	0.104	1,599	0.116	2,358	0.012	0.032	0.713
Household member participates in an association	0.036	1,627	0.038	2,380	0.001	0.006	0.835
Agricultural association	0.646	479	0.636	784	-0.010	0.064	0.879
Commercial association	0.162	479	0.074	784	-0.088	0.060	0.143
Savings/loan association	0.160	479	0.209	784	0.049	0.077	0.527
Household has access to land for agriculture	0.854	1,604	0.841	2,365	-0.013	0.024	0.603
Household harvested agricultural product	0.766	1,604	0.754	2,365	-0.012	0.030	0.687
White, yellow, or black corn	0.856	1,356	0.815	1,986	-0.041	0.051	0.421
Coffee	0.238	1,356	0.250	1,986	0.011	0.037	0.765
Black beans	0.391	1,356	0.395	1,986	0.004	0.047	0.938
Raised animals	0.696	1,604	0.703	2,365	0.008	0.034	0.827
Produced animal product	0.449	1,604	0.458	2,365	0.009	0.029	0.758
Irrigation pump or sprayer	0.352	1,448	0.364	2,112	0.012	0.033	0.711
Barrel	0.134	1,448	0.099	2,112	-0.035	0.019	0.063
Wagon or Cart	0.081	1,448	0.096	2,112	0.015	0.017	0.360
Household has agricultural/livestock facility	0.479	1,627	0.473	2,380	-0.006	0.038	0.880
Hen house	0.325	1,448	0.364	2,112	0.039	0.030	0.192
Pigpen	0.160	1,448	0.177	2,112	0.017	0.037	0.648
General pen	0.177	1,448	0.163	2,112	-0.014	0.028	0.621
Received technical assistance	0.051	1,448	0.043	2,112	-0.008	0.009	0.380
Cooperative assisted	0.319	355	0.397	547	0.077	0.084	0.355
Private company assisted	0.281	355	0.276	547	-0.004	0.082	0.957
Ministerio de Agricultura, Ganadería y Alimentación (MAGA) assisted	0.302	355	0.210	547	-0.091	0.096	0.341

C.2. Differential attrition

Table C.2.1. Roster characteristics

Variables	Comparison Mean	Comparison N1	Treatment Mean	Treatment N2	Mean Diff	Diff SE	p-value
Age	23.277	7,982	22.668	14,209	-0.609	0.678	0.370
Female	0.516	7,985	0.519	14,210	0.003	0.007	0.722
Married or in consensual union	0.632	4,530	0.633	8,188	0.001	0.016	0.928
Years of schooling, age 18 and older	2.546	3,943	2.462	6,996	-0.085	0.148	0.568
No education, age 18 and older	0.329	3,953	0.357	7,013	0.028	0.029	0.341
Completed or not completed primary education, age 18 and older	0.479	3,955	0.482	7,014	0.003	0.023	0.901
Completed or not completed secondary education, age 18 and older	0.170	3,955	0.144	7,014	-0.026	0.025	0.303
Cannot read and write	0.289	4,531	0.305	8,192	0.016	0.026	0.538
Member ages 5-18 currently studying	0.771	2,727	0.752	4,882	-0.019	0.024	0.410

Table C.2.2. Household characteristics

Variables	Comparison Mean	Comparison N1	Treatment Mean	Treatment N2	Mean Diff	Diff SE	p-value
Total number of individuals in household	5.727	1,392	5.820	2,380	0.093	0.183	0.612
Number of adults ages 18-64 years	2.561	1,392	2.560	2,380	-0.001	0.069	0.988
Child under 5 in household	0.553	1,392	0.564	2,380	0.011	0.030	0.728
Child age 5-17 in household	0.770	1,392	0.777	2,380	0.007	0.023	0.754
Member age 65 or older in household	0.212	1,392	0.180	2,380	-0.032	0.019	0.103
Household has WRA	0.882	1,392	0.916	2,380	0.034	0.014	0.016
Single woman	0.199	1,392	0.180	2,380	-0.020	0.019	0.301
Household in rural area	0.788	1,392	0.752	2,380	-0.037	0.080	0.648
Municipality	17.908	1,392	10.569	2,380	-7.339	1.294	0.000
Altitude in meters	1,335.310	1,392	1,662.660	2,380	327.349	165.411	0.049
Used translator during interview	0.153	1,392	0.191	2,380	0.038	0.053	0.473
Household interview duration in hours	32.940	1,391	32.688	2,378	-0.253	1.155	0.827

Table C.2.3. Household head characteristics

Variables	Comparison Mean	Comparison N1	Treatment Mean	Treatment N2	Mean Diff	Diff SE	p-value
Age	45.996	1,391	43.907	2,380	-2.089	0.865	0.017
Female	0.209	1,392	0.196	2,380	-0.013	0.019	0.508
Married or in consensual union	0.846	1,391	0.877	2,380	0.031	0.016	0.047
Years of schooling, age 18 and older	2.430	1,385	2.300	2,369	-0.130	0.153	0.396
No education, age 18 and older	0.317	1,390	0.348	2,378	0.030	0.032	0.346
Completed or not completed primary education, age 18 and older	0.563	1,390	0.550	2,378	-0.014	0.032	0.664
Completed or not completed secondary education, age 18 and older	0.100	1,390	0.079	2,378	-0.021	0.021	0.308
Cannot read and write	0.288	1,391	0.319	2,380	0.031	0.033	0.348

Table C.2.4. Housing characteristics/items

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Dirt or sand floors	0.462	1,392	0.450	2,380	-0.012	0.052	0.821
Cement floors	0.397	1,392	0.402	2,380	0.005	0.042	0.905
Corr. zinc or metal roof	0.775	1,392	0.760	2,380	-0.015	0.048	0.754
Concrete, reinforced concrete, ceramic roof	0.100	1,392	0.119	2,380	0.019	0.027	0.466
Cinder block walls	0.445	1,392	0.433	2,380	-0.012	0.055	0.824
Adobe or covered adobe walls	0.239	1,392	0.327	2,380	0.088	0.056	0.115
Household has only one bedroom	0.498	1,387	0.479	2,378	-0.019	0.037	0.610
Crowding index (household size/ number of bedrooms)	3.851	1,387	3.870	2,378	0.019	0.219	0.933
Household has electricity	0.870	1,376	0.839	2,365	-0.031	0.044	0.480
Connected to public water system	0.633	1,392	0.767	2,380	0.134	0.065	0.040
Improved water source or appropriately treated unimproved source	0.957	1,392	0.958	2,380	0.001	0.021	0.971
Observed location with water and soap for handwashing	0.354	1,392	0.265	2,380	-0.088	0.037	0.017
Has latrine/outhouse/cesspits	0.576	1,392	0.567	2,380	-0.009	0.060	0.885
Connected to sewage system	0.204	1,392	0.257	2,380	0.053	0.060	0.381
Improved sanitation	0.764	1,392	0.829	2,380	0.065	0.027	0.015
Room used exclusively for cooking	0.624	1,392	0.666	2,380	0.042	0.034	0.221
Firewood main source of cooking fuel	0.921	1,392	0.935	2,380	0.014	0.021	0.505
Television	0.620	1,213	0.591	1,970	-0.029	0.044	0.509
Radio	0.448	1,213	0.433	1,970	-0.016	0.029	0.589
DVD player	0.178	1,213	0.175	1,970	-0.004	0.023	0.878
CD player	0.163	1,213	0.198	1,970	0.035	0.028	0.214
Computer	0.101	1,213	0.087	1,970	-0.013	0.023	0.570
Printer	0.068	1,213	0.050	1,970	-0.018	0.015	0.246
Video player	0.066	1,213	0.048	1,970	-0.019	0.013	0.142
Still camera	0.035	1,213	0.033	1,970	-0.002	0.011	0.849
Video camera	0.019	1,213	0.019	1,970	0.000	0.007	0.990
Pickup truck	0.058	1,376	0.071	2,365	0.013	0.015	0.392
Moto	0.066	1,376	0.065	2,365	-0.001	0.015	0.944
Bike	0.154	1,376	0.115	2,365	-0.039	0.030	0.199
Home is owned and fully paid for	0.861	1,375	0.896	2,365	0.035	0.017	0.044
Member left for employment in past 12 months	0.255	1,392	0.288	2,380	0.033	0.030	0.262
Household received cash remittance from family in foreign country	0.134	1,392	0.130	2,380	-0.005	0.020	0.819

Table C.2.5. Program indicators

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Received gov't benefits through fertilizer program in past 12 months	0.326	1,392	0.272	2,380	-0.055	0.048	0.249
Received gov't benefits through Bono Seguro program in past 12 months	0.318	1,392	0.267	2,380	-0.051	0.038	0.185
Have heard about USAID or can identify its logo	0.184	1,389	0.211	2,350	0.027	0.027	0.313
Reports USAID provides services	0.212	271	0.272	636	0.060	0.059	0.310

Table C.2.6. Consumption by expense type

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Food and drink	10.946	1,376	9.859	2,365	-1.087	0.495	0.029
Meals outside the home	0.491	1,376	0.449	2,365	-0.041	0.082	0.613
Housing	2.962	1,376	3.033	2,365	0.071	0.302	0.815
Household services	2.954	1,376	2.847	2,365	-0.107	0.210	0.611
Education	1.333	1,376	1.116	2,365	-0.217	0.205	0.290
Health	4.084	1,376	3.144	2,365	-0.940	0.558	0.093
Household equipment, furnishings	0.607	1,376	0.657	2,365	0.050	0.146	0.731
Donations	0.086	1,376	0.067	2,365	-0.020	0.021	0.344
Household items, house cleaning and maintenance, and household appliances	1.548	1,376	1.499	2,365	-0.049	0.164	0.768
Taxes, funeral, pensions	0.697	1,376	0.682	2,365	-0.015	0.114	0.896
Recreation, entertainment, and tourism	0.273	1,376	0.237	2,365	-0.036	0.102	0.724
Personal care	0.681	1,376	0.820	2,365	0.139	0.062	0.027
Clothing and shoes	0.591	1,376	0.579	2,365	-0.012	0.065	0.853
Other	1.800	1,376	1.836	2,365	0.036	0.217	0.868

Table C.2.7. Consumption, item share

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Food and drink	42.545	1,376	41.633	2,365	-0.913	1.235	0.461
Meals outside the home	1.355	1,376	1.086	2,365	-0.270	0.159	0.090
Housing	10.313	1,376	10.966	2,365	0.652	0.452	0.150
Household services	11.125	1,376	11.825	2,365	0.700	0.453	0.124
Education	4.394	1,376	4.039	2,365	-0.355	0.447	0.428
Health	9.908	1,376	9.053	2,365	-0.854	0.665	0.201
Household equipment, furnishings	1.357	1,376	1.501	2,365	0.144	0.205	0.482
Donations	0.335	1,376	0.257	2,365	-0.078	0.079	0.325
Household items, house cleaning and maintenance, and household appliances	5.207	1,376	5.252	2,365	0.045	0.254	0.859
Taxes, funeral, pensions	1.759	1,376	1.851	2,365	0.092	0.165	0.579
Recreation, entertainment, and tourism	0.507	1,376	0.473	2,365	-0.034	0.086	0.692
Personal care	1.953	1,376	2.001	2,365	0.048	0.107	0.654
Clothing and shoes	2.605	1,376	3.371	2,365	0.766	0.234	0.001
Other	6.638	1,376	6.694	2,365	0.056	0.394	0.887

Table C.2.8. Prevalence of poverty, individual level

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Extreme poverty line (13.18 Qetzales)	0.213	1,376	0.260	2,365	0.048	0.031	0.132
Total poverty line (27.17 Quetzales)	0.715	1,376	0.752	2,365	0.037	0.036	0.308
USD 1.25 poverty line	0.049	1,376	0.061	2,365	0.012	0.017	0.484
USD 2.00 poverty line	0.252	1,376	0.285	2,365	0.033	0.032	0.306

Table C.2.9. Per capita consumption

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Daily per capita consumption (Quetzales)	25.430	1,376	23.197	2,365	-2.233	1.752	0.204
Daily per capita consumption (2010 USD)	4.100	1,376	3.740	2,365	-0.360	0.283	0.204

Table C.2.10. Nutrition indicators, anthropometrics, among children under 5 years of age

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Moderate chronic malnutrition (stunted)	0.626	1,096	0.653	1,881	0.026	0.037	0.472
Severe chronic malnutrition	0.272	1,096	0.270	1,881	-0.002	0.038	0.958
Moderate acute malnutrition (wasted)	0.008	1,087	0.009	1,872	0.002	0.004	0.651
Severe global malnutrition	0.000	1,087	0.000	1,872	0.000	0.000	
Moderate global malnutrition (underweight)	0.179	1,096	0.158	1,878	-0.021	0.023	0.364
Severe acute malnutrition	0.037	1,096	0.018	1,878	-0.019	0.008	0.025
Overweight (high BMI for age)	0.056	1,096	0.050	1,881	-0.005	0.011	0.640

Table C.2.11. Nutrition indicators, breastfeeding ages 0 to 5 months

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
No breastfeeding	0.008	127	0.001	189	-0.006	0.008	0.409
Exclusive breastfeeding	0.618	127	0.697	189	0.079	0.079	0.321
Breastfeeding and other liquids	0.341	127	0.272	189	-0.069	0.074	0.352
Breastfeeding and complementary feeding	0.034	127	0.030	189	-0.004	0.021	0.861

Table C.2.12. Nutrition indicators, breastfeeding ages 6 to 23 months

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
No breastfeeding	0.120	335	0.161	560	0.041	0.036	0.248
Exclusive breastfeeding	0.020	335	0.014	560	-0.006	0.010	0.553
Breastfeeding and other liquids	0.022	335	0.009	560	-0.013	0.012	0.269
Breastfeeding and complementary feeding	0.839	335	0.816	560	-0.023	0.038	0.556

Table C.2.13. Nutrition indicators, supplements

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Took iron in the previous 7 days	0.207	1,116	0.217	1,904	0.010	0.025	0.690
Took vitamin A during the preceding 6 months	0.609	1,116	0.651	1,904	0.043	0.030	0.161
Took deworming medication during the preceding 6 months	0.397	1,116	0.420	1,904	0.023	0.029	0.421

Table C.2.14. Nutrition indicators, minimum acceptable diet

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Acceptable feeding frequency	66.201	335	72.658	560	6.457	4.707	0.172
Acceptable dietary diversity	46.810	335	49.155	560	2.345	3.855	0.544
Minimum acceptable diet	33.878	335	42.873	560	8.995	4.063	0.028

Table C.2.15. Nutrition indicators, lactation initiation

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Time of first breastfeeding	0.512	816	0.562	1,408	0.050	0.034	0.148
Within first hour after birth	0.339	816	0.296	1,408	-0.042	0.030	0.157
After the first hour after birth	0.134	816	0.130	1,408	-0.004	0.019	0.830
After the first day	0.015	816	0.012	1,408	-0.004	0.006	0.557

Table C.2.16. Nutrition indicators, woman level

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Height less than 145 cm	0.379	1,699	0.424	3,157	0.044	0.035	0.208
Height 145 cm or more	0.621	1,699	0.576	3,157	-0.044	0.035	0.208
Low BMI (<18.5) in 15- to 49-year-old women	2.184	1,699	1.986	3,157	-0.198	0.504	0.695
Overweight BMI	43.107	1,699	44.041	3,157	0.934	2.360	0.693
Acceptable dietary diversity	0.679	1,826	0.741	3,390	0.062	0.028	0.030
Dietary diversity in women: Average number of food groups ingested	4.344	1,826	4.472	3,390	0.128	0.108	0.235

Table C.2.17. Nutrition indicators, household-level women's nutrition

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Prevalence of households with moderate or severe hunger	0.168	1,389	0.143	2,379	-0.025	0.023	0.268
Produces food crops for household use	0.120	1,392	0.123	2,380	0.003	0.036	0.943
Does not produce but has available lot or land	0.386	1,392	0.447	2,380	0.061	0.034	0.077
Does not produce and has no available lot or land	0.492	1,392	0.430	2,380	-0.062	0.041	0.136
Believes malnutrition affects his/her household	0.536	1,392	0.528	2,380	-0.007	0.037	0.840
Feels malnutrition is a serious problem in his/her community	0.649	1,392	0.614	2,380	-0.035	0.026	0.173

Table C.2.18. Maternal health, woman level

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Has child, among women ages 18-24 years	0.579	503	0.505	977	-0.074	0.042	0.079
Age at first birth among woman ages 18-24 was before 18	0.444	291	0.400	513	-0.045	0.053	0.398
Named two or more key signs of prenatal risk	0.647	1,826	0.612	3,390	-0.035	0.039	0.376
Named two or more key signs of delivery risks	0.648	1,826	0.646	3,390	-0.002	0.025	0.931
Named two or more key signs of postpartum risks	0.678	1,826	0.628	3,390	-0.049	0.028	0.076
Named two or more key signs of newborn risks	0.506	1,826	0.494	3,390	-0.013	0.039	0.750

Table C.2.19. Maternal health, pregnancy level

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Received any prenatal care	0.919	824	0.940	1,428	0.021	0.020	0.293
Received four or more prenatal checkups	0.678	824	0.798	1,429	0.120	0.029	0.000
Skilled provider attended birth	0.907	1,147	0.939	1,960	0.033	0.023	0.153
Physician, ambulatory physician, or nurse attended birth	0.410	1,147	0.366	1,960	-0.044	0.050	0.382
Place of delivery was facility	0.413	1,147	0.365	1,960	-0.048	0.052	0.356
Received postpartum care (postnatal woman)	0.728	824	0.721	1,427	-0.007	0.038	0.850
Had plan to ensure transportation	56.420	350	56.377	529	-0.043	4.877	0.993
Had plan to save money	77.763	350	72.724	529	-5.040	4.128	0.224
Had identified blood donors	7.663	350	6.834	529	-0.829	2.541	0.744
Had selected a place for the delivery	57.546	350	52.221	529	-5.325	5.401	0.325

Table C.2.20. Fertility and family planning

Variables	Comparison Mean	N1	Treatment Mean	N2	Mean Diff	Diff SE	p-value
Average number of live births	2.705	1,826	2.505	3,390	-0.200	0.117	0.089
Average number of living children	2.396	1,893	2.228	3,508	-0.168	0.108	0.123
Uses any contraceptive method	0.504	1,135	0.525	2,098	0.020	0.033	0.536
Uses any modern contraceptive method	0.407	1,135	0.414	2,098	0.007	0.037	0.844
Unmet need	0.195	1,135	0.167	2,098	-0.028	0.021	0.191
Total demand for family planning	0.699	1,135	0.692	2,098	-0.008	0.028	0.779

Table C.2.21. Children's health

Variables	Comparison Mean	Comparison N1	Treatment Mean	Treatment N2	Mean Diff	Diff SE	p-value
Received postnatal care	0.681	824	0.823	1,425	0.142	0.034	0.000
Pentavalent 1(12-23 months)	0.996	218	0.985	384	-0.010	0.010	0.326
Pentavalent 2 (12-23 months)	0.989	218	0.979	384	-0.010	0.013	0.454
Pentavalent 3 (12-23 months)	0.943	218	0.958	384	0.016	0.021	0.467
MMR/measles (12-23 months)	0.942	218	0.871	384	-0.071	0.040	0.074
Pentavalent booster 1 (12-23 months)	0.390	218	0.397	384	0.007	0.055	0.896
Pentavalent 1 (24-59 months)	0.981	640	0.992	1,129	0.011	0.008	0.183
Pentavalent 2 (24-59 months)	0.959	640	0.985	1,129	0.026	0.013	0.047
Pentavalent 3 (24-59 months)	0.934	640	0.964	1,129	0.030	0.016	0.066
All three pentavalent vaccines (12-59)	0.907	888	0.935	1,561	0.028	0.017	0.100
MMR/measles (24-59 months)	0.964	640	0.973	1,129	0.009	0.013	0.467
Pentavalent booster 1 (24-59 months)	0.782	640	0.864	1,129	0.082	0.036	0.024
Pentavalent booster 2 (24-59 months)	0.219	640	0.241	1,129	0.022	0.026	0.393
All vaccines (24-59 months)	0.216	640	0.241	1,129	0.025	0.026	0.329
Had diarrhea in the past two weeks	0.285	1,116	0.220	1,904	-0.065	0.022	0.003
Was taken to a health facility	0.408	300	0.461	411	0.053	0.054	0.330
Was given ORS or zinc	0.463	300	0.426	411	-0.037	0.060	0.540
Had acute respiratory infection in the past two weeks	0.216	1,116	0.145	1,904	-0.071	0.025	0.004
Was taken to a health facility	0.488	234	0.664	286	0.176	0.065	0.007
Was administered antibiotics or other medication	0.820	234	0.843	286	0.023	0.049	0.631

Table C.2.22. Participation in agricultural activities among household members 12 and older

Variables	Comparison Mean	Comparison N1	Treatment Mean	Treatment N2	Mean Diff	Diff SE	p-value
Household members age 12 years or over who work in agriculture	27.873	5,253	27.580	9,517	-0.293	1.873	0.876

Table C.2.23. Participation in agricultural activities, household

Variables	Comparison Mean	Comparison N1	Treatment Mean	Treatment N2	Mean Diff	Diff SE	p-value
Member age 12 or older is farmer	0.477	1,392	0.509	2,380	0.031	0.052	0.549
Member age 12 or older is farm laborer	0.269	1,392	0.269	2,380	-0.000	0.033	0.998
Household produces food for household consumption	0.112	1,372	0.116	2,358	0.004	0.036	0.912
Household member participates in an association	0.006	1,392	0.038	2,380	0.032	0.007	0.000
Agricultural association	0.336	11	0.636	784	0.299	0.169	0.080
Commercial association	0.000	11	0.074	784	0.074	0.029	0.012
Savings/loan association	0.401	11	0.209	784	-0.192	0.172	0.266
Household has access to land for agriculture	0.799	1,376	0.841	2,365	0.042	0.043	0.332
Household harvested agricultural product	0.702	1,376	0.754	2,365	0.052	0.051	0.306
White, yellow, or black corn	0.816	1,007	0.815	1,986	-0.001	0.064	0.989
Coffee	0.186	1,007	0.250	1,986	0.064	0.051	0.212
Black beans	0.385	1,007	0.395	1,986	0.010	0.046	0.832
Raised animals	0.662	1,376	0.703	2,365	0.041	0.046	0.373
Produced animal product	0.448	1,376	0.458	2,365	0.010	0.038	0.788
Irrigation pump or sprayer	0.415	1,125	0.364	2,112	-0.051	0.040	0.205
Barrel	0.165	1,125	0.099	2,112	-0.066	0.023	0.005
Wagon or cart	0.081	1,125	0.096	2,112	0.016	0.017	0.360
Household has agricultural/livestock facility	0.445	1,392	0.473	2,380	0.028	0.048	0.557
Hen house	0.354	1,125	0.364	2,112	0.010	0.036	0.778
Pigpen	0.135	1,125	0.177	2,112	0.042	0.037	0.263
General pen	0.132	1,125	0.163	2,112	0.031	0.027	0.253
Received technical assistance	0.040	1,125	0.043	2,112	0.003	0.010	0.752
Cooperative assisted	0.060	51	0.397	547	0.337	0.076	0.000
Private company assisted	0.091	51	0.276	547	0.185	0.098	0.061
MAGA assisted	0.654	51	0.210	547	-0.444	0.099	0.000

APPENDIX D. SAMPLING ISSUES: ADJUSTING FOR ATTRITION AND NONRESPONSE

This appendix describes the procedures used to adjust weights in the baseline and midline WHIP surveys to conduct the impact evaluation. We studied the need for adjustments to address household and census tract attrition, nonresponse on the consumption survey, and nonresponse on the women's survey (including nonparticipation in anthropometric measurement for both women and children, and women's nonresponse on survey questions about children included in the sample).

Weights were adjusted for attrition observed between the baseline and midline surveys. The first procedure described here accounts for household-level attrition because only households with interview data from both rounds were included in the impact analysis. The second procedure is an adjustment for census tract attrition (with census tracts constituting sample clusters) because the impact analysis included only households from tracts that were visited in both survey rounds. Two hundred seventeen (217) baseline clusters were revisited at midline, but nine were not.

Adjustment for Household Attrition

Household weights were adjusted using the following factor, calculated for the households in each poststratum h:

$$factorP_h = (\sum_{total\ en\ h} pesohogar_{h,i}) / (\sum_{panel\ en\ h} pesohogar_{h,i})$$

Where the sum in the numerator represents all households in the poststratum h, and the sum in the denominator represents only households in the panel (insample_panel_hh = 1) in the poststratum.

In keeping with the baseline design, to avoid possible bias and increases in the variance of the estimators, we defined the poststrata using the following variables:

-- dominio

-- quintile_rexp_pcdaily (quintil)

The values of these factors appear in Table D.1.

Adjusted weights for each household i in the poststratum h were calculated as follows:

$$pesohogP_{h,i} = factorP_h * pesohogar_{h,i}$$

Women's and children's case weights were calculated using the household weight, the number of eligible women in a household, the number of women interviewed, the number of women measured, the number of eligible children, the number of children for whom interview data were provided, and the number of children measured. The calculations were as follows:

factorme = count_women / count_women_interviewed

factormm = count_women / count_women_measured

factornm = count_children / count_children_measured

factorne	= max(1, count_children / nino_entrevista_mujer)	baseline (time = 0)
	= max(1, count_children / count_children_interviewed_ml)	midline (time = 1)
pesomeP	= factorme * pesohogarP	Woman interviewed
pesommP	= factormm * pesohogarP	Woman measured
pesonmP	= factornm * pesohogarP	Child measured
pesoneP	= factorne * pesomeP	Child with interview data

Table D.1. Weight adjustments for household attrition

Quintile/panel	Household weight	Number of cases	factorP
Domain 1: RVCP Direct			
10	265.6	57	
11	486.9	92	1.546
1	752.5	149	
20	320.5	84	
21	604.6	129	1.530
2	925.1	213	
30	285.4	76	
31	558.4	132	1.511
3	843.8	208	
40	487.3	129	
41	756.5	178	1.644
4	1,243.7	307	
50	629.3	136	
51	1,179.6	251	1.534
5	1,808.9	387	
Domain 2: RVCP Indirect			
10	2,137.4	115	
11	3,722.2	151	1.574
1	5,859.5	266	
20	3,182.1	160	
21	4,752.9	181	1.669
2	7,935.0	341	
30	2,325.3	112	
31	5,098.4	185	1.456
3	7,423.7	297	
40	3,995.6	183	
41	5,243.3	198	1.762
4	9,238.9	381	
50	4,947.3	180	
51	8,513.3	281	1.581
5	13,460.6	461	

Table D.1. Weight adjustments for household attrition (continued)

Quintile/panel	Household weight	Number of cases	factorP
Domain 3: Health Only			
10	9,543.4	69	
11	13,750.0	83	1.6941
1	23,293.3	152	
20	9,488.5	74	
21	16,985.4	104	1.5586
2	26,473.8	178	
30	11,037.4	77	
31	18,395.1	113	1.6000
3	29,432.5	190	
40	11,984.1	81	
41	22,035.1	137	1.5439
4	34,019.1	218	
50	14,087.0	94	
51	26,738.3	165	1.5268
5	40,825.4	259	
Domains 4 and 5: ZOI Comparison Group			
10	825.4	117	
11	1,047.0	139	1.788
1	1,872.3	256	
20	916.4	147	
21	1,723.4	244	1.532
2	2,639.8	391	
30	1,043.0	156	
31	1,584.4	221	1.658
3	2,627.4	377	
40	1,362.5	204	
41	2,554.7	361	1.533
4	3,917.3	565	
50	1,938.6	278	
51	3,030.3	427	1.640
5	4,968.9	705	

The previously calculated household weights could also be applied to consumption indicators, although consumption data were missing for some households. Accordingly, we applied weights as follows to the calculation of consumption indicators for households and individuals:

pesohog4P = pesohogarP

pesoper4P = count_members * pesohog4P

Nonresponse on the consumption survey was sufficiently rare that adjustment factors calculated for this purpose were ultimately not used in the impact analysis.

Adjustment for Census Tract Attrition

Census tract weights were adjusted using the following factor, calculated for tracts in the domain h:

$$factorU_h = (\sum_{total\ en\ h} pesosc_{h,i}) / (\sum_{panel\ en\ h} pesosc_{h,i})$$

Where

$pesosc_{h,i}$ = the weight (the inverse of the probability of selection at baseline) of the census tract i from the domain h.

The sum in the numerator represents all census tracts in the h domain, whereas the denominator represents only those tracts included in the panel (insample_panel_clust = 1) in the domain. The values of these factors are provided in Table D.2.

Next, adjusted weights for each household i from the domain h were calculated as follows for the baseline data:

$$pesohogU_{h,i} = factorU_h * pesohogar_{h,i}$$

And the women's and children's weights were calculated as follows:

pesomeU	= factorme * pesohogarU	Woman interviewed
pesommU	= factormm * pesohogarU	Woman measured
pesonmU	= factornm * pesohogarU	Child measured
pesoneU	= factorne * pesomeU	Child with interview data

Table D.2. Weight adjustments for cluster attrition

Domain	Baseline (pesosc)	Panel tract	factorU
1	308.0	305.3	1.0087
2	308.0	305.3	1.0087
3	1,051.3	940.4	1.1179
4	74.0	73.0	1.0137
5	29.0	26.0	1.1154

Using the same adjustment factors, we calculated an adjusted weight for each household i in the domain h for the midline survey. The calculation was as follows:

$$fehogarU_{h,i} = factorU_{h} * fe2002_{h,i}$$

And the women's and children's weights were calculated as follows:

femeU	= factorme * fehogarU	Woman interviewed
femmU	= factormm * fehogarU	Woman measured
fenmU	= factornm * fehogarU	Child measured
feneU	= factorne * femeU	Child with interview data

The previously calculated household weights could also be applied to the calculation of consumption indicators, although consumption data were missing for some households. Accordingly, we defined the consumption indicators as follows for households and individuals, for the baseline and midline survey data:

pesohog4U	= pesohogarU
pesoper4U	= count_members * pesohog4U
fehog4U	= fehogarU
feper4U	= count_members * fehog4U

As previously noted, nonresponse on the consumption survey was sufficiently rare that adjustment factors calculated for this purpose were ultimately not applied to the impact analysis.

APPENDIX E. CONVERSION FACTORS FOR CONSUMPTION EXPENDITURES

Table E.1. Conversion references

Item	Value
2011 Total Poverty Line[a]	Q 24.74 daily per capita
2011 Extreme Poverty Line[a]	Q 12.00 daily per capita
Base Year 2011	
Consumer Price Index (CPI), Guatemala May 2011[b]	103.68
CPI, Guatemala September 2013[b]	113.85
CPI, Guatemala September 2015[c]	119.99
Base Year 2005	
CPI, Guatemala May 2005[d]	100.0
CPI, Guatemala September 2013[d]	152.522
CPI, Guatemala September 2015[e]	158.72
CPI, US 2005[f]	100.0
CPI, US 2010[f]	111.6563
PPP Conversion Factor, Quetzales per US 2005 International Dollar[g]	4.54033

Sources:

[a] Government of Guatemala, Encuesta Nacional de Condiciones de Vida.

[b] May 2011 was the midpoint of ENCOVI field work; September 2013 was the midpoint of field work for the Poverty Assessment Tool by DevTech Systems.

[c] Instituto Nacional de Estadística. Índice de Precios al Consumidor y Costo de la Canasta Básica Alimentaria y Vital. Diciembre 2015, Base Diciembre 2010. Guatemala, Enero de 2016.

[d] World Development Indicators, World Bank, Financial Sector, Exchange rates & prices series http://databank.worldbank.org/data/views/reports/tableview.aspx#, and CPI 2013 data, INE, Guatemala.

[e] Correspondence with Pablo Toledo, DevTech Systems, May 2016.

[f] United States Department of Labor, Bureau of labor Statistics. Consumer Price Index. http://www.bls.gov/cpi/

[g] World Development Indicators, World Bank, Economic Policy & Debt, Purchasing Power Parity series http://databank.worldbank.org/data/views/reports/tableview.aspx#

Relative Poverty Lines

The Government of Guatemala's national poverty lines, obtained from the 2011 Living Standards Measurement Survey, were updated to 2013 Quetzales for the baseline data and to 2015 Quetzales for the midline data using the overall Consumer Price Index (CPI), base year 2011.

Total Poverty Line (daily per capita)

Nominal May 2011 Quetzales: Q 24.74 per capita/day

Nominal September 2013 Quetzales: Q 24.74 * (113.85/103.68) = Q 27.17 per capita/day

Nominal September 2015 Quetzales: Q 24.74*(119.99/103.68) = Q 28.63 per capita/day

Extreme Poverty Line (daily per capita)

Nominal May 2011 Quetzales: Q 12.00 per capita/day

Nominal September 2013 Quetzales: Q 12.00 * (113.85/103.68) = Q 13.18 per capita/day

Nominal September 2015 Quetzales: Q 12.00 *(119.99/103.68) = Q 13.89 per capita/day

A household was considered poor at baseline if its nominal 2013 daily per capita consumption was less than Q 27.17 pc/day, and a household was considered poor at midline if its nominal 2015 daily per capita consumption was less than Q 28.63 pc/day.

International Poverty Lines

The USD 1.25 and USD 2.00 international poverty lines (2005 PPP) were converted to 2013 and 2015 Quetzales using the overall CPI, base year 2005.

USD 1.25 Poverty Line (2005 PPP)

Nominal September 2013 Quetzales: 1.25 * 4.54033 * (152.522/100.0) = Q 8.66 pc/day

Nominal September 2015 Quetzales: 1.25 * 4.54033 * (158.72/100.0) = Q 9.01 pc/day

USD 2.00 Poverty Line (2005 PPP)

Nominal September 2013 Quetzales: 2.00 * 4.54033 * (152.522/100.0) = Q 13.85 pc/day

Nominal September 2015 Quetzales: 2.00 * 4.54033 * (158.72/100.0) = Q 14.41 pc/day

A household was considered below the USD 1.25 poverty line at baseline if its nominal 2013 daily per capita consumption was below Q 8.66, and a household was considered below the USD 1.25 line at midline if its nominal 2015 daily per capita consumption was below Q 9.01.

Consumption in Real 2013 Quetzales

Midline household consumption was converted from nominal 2015 values to real September 2013 values to facilitate comparison of household expenditures over time. Using the 2011 base year CPI, the following conversion factor was applied to midline consumption data:

Real 2013 consumption = Nominal 2015 consumption * (113.85/119.99)

= Nominal 2015 consumption * 0.948829

Consumption in 2010 USD

Baseline and midline household consumption data were converted to 2010 US dollars using the 2005 base year CPI information.

Conversion Factor from Nominal 2013 Quetzales to 2010 USD

Real 2013 consumption = Nominal 2015 consumption * (113.85/119.99)

= Nominal 2015 consumption * 0.948829

Conversion Factor from Nominal 2015 Quetzales to 2010 USD

$$2010\ USD = Q_{2015} * \left(\frac{100\ Q_{2005}}{158.72\ Q_{2015}}\right) * \left(\frac{1\ USD_{2005}}{4.54033\ Q_{2005}}\right) * \left(\frac{111.6563\ USD_{2010}}{100.0\ USD_{2005}}\right)$$

$$= Q_{2015} * 0.15494022$$

APPENDIX F. HETEROGENEOUS IMPACT ANALYSIS BY BASELINE CONSUMPTION LEVEL

Table F.1. Main results, by baseline consumption level

	Baseline Consumption - Lower 50%						Baseline Consumption - Upper 50%					
		ZOI		Comparison		DID Impact		ZOI		Comparison		DID Impact
Outcome	N	BL	ML	BL	ML	(Household [HH] FE)	N	BL	ML	BL	ML	(HH FE)
Consumption and Poverty												
Daily per capita expenditures (2010 USD)	3,467	2.236	3.032	2.251	3.149	-0.046	3,905	5.842	4.990	6.206	4.806	0.255
		(0.046)	(0.084)	(0.038)	(0.133)	(0.174)		(0.212)	(0.283)	(0.301)	(0.214)	(0.409)
Prevalence of poverty (USD1.25)[a]	3,467	0.108	0.063	0.103	0.072	-0.013	3,905	.	0.017	.	0.027	.
		(0.023)	(0.016)	(0.023)	(0.017)	(0.045)		.	(0.008)	.	(0.008)	.
Depth of poverty (USD1.25)[a]	3,467	2.039	1.525	2.121	1.637	-0.182	3,905	.	0.249	.	0.413	.
		(0.483)	(0.517)	(0.565)	(0.571)	(1.096)		.	(0.145)	.	(0.148)	.
Food Security and Nutrition												
Prevalence of households with moderate/severe hunger	3,467	0.182	0.116	0.211	0.149	0.003	4,032	0.107	0.081	0.133	0.094	0.008
		(0.028)	(0.014)	(0.022)	(0.022)	(0.056)		(0.018)	(0.015)	(0.019)	(0.013)	(0.040)
Prevalence of exclusive breastfeeding age 0-5 months[b]	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
Prevalence of minimum acceptable diet age 6-23 months	829	0.406	0.222	0.283	0.277	-0.077	533	0.431	0.286	0.454	0.205	-0.243
		(0.051)	(0.063)	(0.033)	(0.065)	(0.498)		(0.049)	(0.064)	(0.050)	(0.048)	(0.610)
Prevalence of stunting age 0-59 months	3,440	0.707	0.641	0.661	0.643	0.045	2,301	0.567	0.533	0.569	0.535	0.051
		(0.032)	(0.024)	(0.026)	(0.031)	(0.064)		(0.044)	(0.045)	(0.031)	(0.032)	(0.089)
Prevalence of wasting age 0-59 months	3,433	0.007	0.009	0.010	0.008	0.010	2,300	0.008	0.020	0.002	0.002	0.008
		(0.004)	(0.006)	(0.004)	(0.004)	(0.014)		(0.005)	(0.014)	(0.002)	(0.002)	(0.025)
Prevalence of underweight age 0-59 months	3,440	0.168	0.177	0.199	0.151	0.114*	2,300	0.126	0.207	0.149	0.167	0.091
		(0.021)	(0.020)	(0.020)	(0.019)	(0.050)		(0.025)	(0.030)	(0.022)	(0.022)	(0.064)
Women's dietary diversity	5,054	4.25	4.26	3.92	4.26	-0.436*	4,681	4.74	4.63	4.73	4.55	0.104
		(0.125)	(0.098)	(0.073)	(0.107)	(0.204)		(0.067)	(0.124)	(0.076)	(0.094)	(0.194)
Prevalence of underweight women	4,798	0.021	0.011	0.019	0.023	-0.019	4,585	0.018	0.012	0.020	0.022	-0.006
		(0.005)	(0.004)	(0.006)	(0.008)	(0.013)		(0.005)	(0.005)	(0.005)	(0.006)	(0.012)

Table F.1. Main results, by baseline consumption level (continued)

	Baseline Consumption - Lower 50%						Baseline Consumption - Upper 50%					
		ZOI		Comparison		DID Impact		ZOI		Comparison		DID Impact
Outcome	N	BL	ML	BL	ML	(Household [HH] FE)	N	BL	ML	BL	ML	(HH FE)
Reproductive Health												
Prevalence of current modern contraceptive use	2,727	0.426	0.367	0.427	0.377	-0.059	2,569	0.502	0.386	0.477	0.368	-0.004
		(0.035)	(0.034)	(0.028)	(0.034)	(0.081)		(0.032)	(0.036)	(0.029)	(0.033)	(0.086)
Prevalence of women with births in past two years	5,120	0.247	0.233	0.330	0.232	0.097**	4,728	0.198	0.213	0.196	0.212	0.030
		(0.015)	(0.022)	(0.019)	(0.025)	(0.032)		(0.022)	(0.020)	(0.017)	(0.016)	(0.047)
Prevalence ≥ four prenatal care visits	1,330	0.811	0.730	0.652	0.571	0.031	927	0.788	0.792	0.729	0.668	0.132
		(0.027)	(0.045)	(0.034)	(0.068)	(0.200)		(0.051)	(0.038)	(0.038)	(0.054)	(0.351)
Prevalence of health facility deliveries	1,328	0.356	0.312	0.342	0.341	0.043	924	0.514	0.524	0.541	0.607	-0.013
		(0.046)	(0.049)	(0.041)	(0.058)	(0.185)		(0.050)	(0.071)	(0.063)	(0.051)	(0.244)
Prevalence of deliveries with skilled birth attendant	1,329	0.356	0.446	0.347	0.527	-0.043	927	0.520	0.608	0.541	0.769	-0.109
		(0.046)	(0.063)	(0.042)	(0.068)	(0.207)		(0.049)	(0.066)	(0.063)	(0.046)	(0.301)
Prevalence of postpartum care among women within two days of delivery	846	0.835	0.786	0.792	0.708	0.123	642	0.824	0.793	0.792	0.713	0.073
		(0.039)	(0.049)	(0.041)	(0.072)	(0.490)		(0.044)	(0.048)	(0.040)	(0.062)	(1.467)

Notes: Summary statistics are means and standard errors for the panel of households. The household FE model was estimated among the panel of households and included FE at the household level and contemporaneous control variables; standard errors were adjusted for clustering at the cluster level and are presented in parentheses.

Significance levels: + p<0.10; * p<0.05; ** p<0.01; *** p<0.001

[a] Baseline summary statistics were not available for households consuming above the baseline median expenditure level (USD 3.34, 2010 USD).

[b] Heterogeneous impact by baseline consumption level were not investigated for exclusive breastfeeding among children ages 0 to 5 months due to the low subgroup sample sizes.

APPENDIX G. HETEROGENEOUS IMPACT ANALYSIS BY CHILD SEX AND AGE

Table G.1. Child feeding and nutritional status outcomes, by child sex

	Female Children						Male Children					
		ZOI		Comparison		DID Impact		ZOI		Comparison		DID Impact
Outcome	N	BL	ML	BL	ML	(HH FE)	N	BL	ML	BL	ML	(HH FE)
Prevalence of minimum acceptable diet ages 6–23 months	678	0.457	0.251	0.374	0.228	-0.567	685	0.376	0.246	0.319	0.253	-1.206
		(0.034)	(0.063)	(0.036)	(0.053)	(1.678)		(0.054)	(0.058)	(0.036)	(0.064)	(2.423)
Prevalence of stunting ages 0–59 months	2,822	0.650	0.599	0.611	0.584	0.013	2,919	0.662	0.594	0.641	0.603	0.021
		(0.039)	(0.035)	(0.030)	(0.033)	(0.096)		(0.033)	(0.036)	(0.026)	(0.028)	(0.085)
Prevalence of wasting ages 0–59 months	2,819	0.006	0.009	0.010	0.002	0.013	2,914	0.008	0.019	0.004	0.009	-0.001
		(0.004)	(0.005)	(0.005)	(0.002)	(0.023)		(0.004)	(0.013)	(0.003)	(0.004)	(0.020)
Prevalence of underweight ages 0–59 months	2,822	0.167	0.186	0.161	0.150	-0.008	2,919	0.141	0.193	0.198	0.167	0.125*
		(0.028)	(0.025)	(0.020)	(0.020)	(0.073)		(0.019)	(0.025)	(0.021)	(0.020)	(0.059)

Notes: Summary statistics are means and standard errors. Standard errors for impact estimates were corrected for clustering at the cluster level and are presented in parentheses. Refer to Table 5 for details on the household- and cluster-FE models.

Significance levels: + p<0.10; * p<0.05; ** p<0.01; *** p<0.001

Table G.2. Child nutritional status outcomes, by child age

	Children ages 0-23 months						Children ages 24-59 months					
		ZOI		Comparison		DID Impact		ZOI		Comparison		DID Impact
Outcome	N	BL	ML	BL	ML	(HH FE)	N	BL	ML	BL	ML	(HH FE)
Prevalence of stunting	2,315	0.600	0.483	0.536	0.484	0.065	3,426	0.694	0.674	0.695	0.658	0.021
		(0.034)	(0.038)	(0.026)	(0.030)	(0.177)		(0.039)	(0.031)	(0.029)	(0.033)	(0.076)
Prevalence of wasting	2,307	0.013	0.026	0.008	0.011	0.006	3,426	0.003	0.005	0.006	0.002	-0.009
		(0.006)	(0.020)	(0.004)	(0.005)	(0.035)		(0.003)	(0.005)	(0.003)	(0.002)	(0.020)
Prevalence of underweight	2,315	0.158	0.166	0.166	0.152	0.040	3,426	0.150	0.205	0.190	0.163	0.114
		(0.030)	(0.035)	(0.016)	(0.020)	(0.136)		(0.020)	(0.023)	(0.022)	(0.019)	(0.074)

Notes: Summary statistics are means and standard errors. Standard errors for impact estimates were corrected for clustering at the cluster level and are presented in parentheses. Refer to Table 5 for details on the household- and cluster-FE models.

Significance levels: + p<0.10; * p<0.05; ** p<0.01; *** p<0.001

APPENDIX H. SUMMARY STATISTICS FOR THE ZOI SUBGROUPS

Table H.1. Subgroup means and standard errors for expenditures and poverty indicators, using the panel of households

	RVCP Direct		RVCP Indirect		Health Only	
Outcome Indicators	Baseline	Midline	Baseline	Midline	Baseline	Midline
Daily per capita expenditures 2010 USD						
Mean	4.11	4.16	4.04	4.07	3.57	3.84
SE	0.256	0.291	0.297	0.264	0.219	0.211
Prevalence of poverty (proportion) USD 1.25						
Mean	0.043	0.051	0.043	0.050	0.071	0.041
SE	0.010	0.017	0.009	0.013	0.018	0.013
Depth of poverty relative to USD1.25 poverty line						
Value	0.665	1.128	0.756	1.218	1.349	0.913
SE	0.219	0.503	0.218	0.343	0.370	0.351
Number of observations	776	753	992	983	597	588

Table H.2. Subgroup means and standard errors for food security on nutrition indicators, using the panel of households

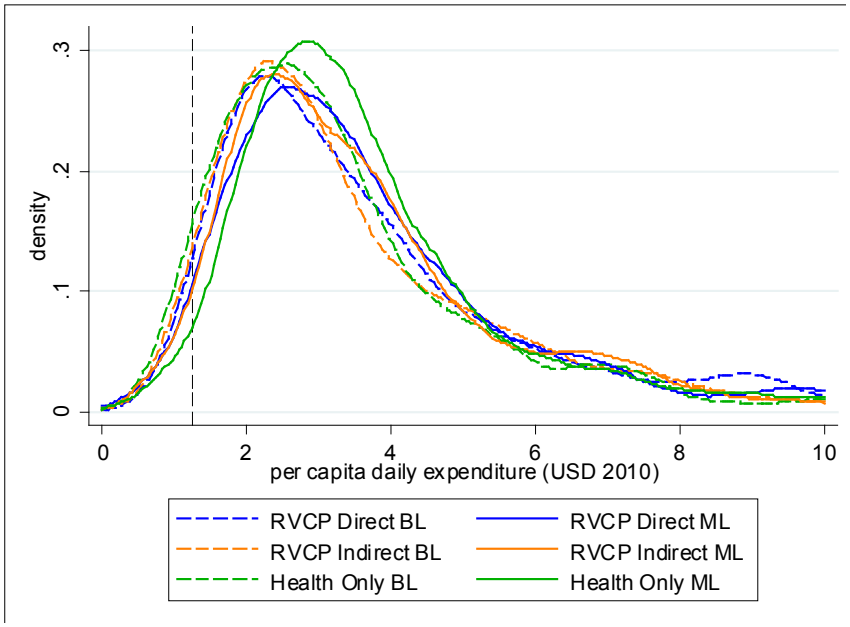
	RVCP Direct		RVCP Indirect		Health Only	
Outcome Indicators	Baseline	Midline	Baseline	Midline	Baseline	Midline
Household food security						
Prevalence of households with moderate or severe hunger						
Value	0.058	0.071	0.132	0.108	0.149	0.096
SE	0.012	0.011	0.014	0.015	0.020	0.015
N	782	779	996	999	602	602
Child health						
Prevalence of children ages 6-23 months receiving a minimum acceptable diet						
Proportion	0.422	0.359	0.375	0.213	0.423	0.253
SE	0.045	0.084	0.040	0.054	0.039	0.052
N	166	84	241	128	144	86
Prevalence of exclusive breastfeeding (on the panel of clusters)						
Proportion	0.781	0.685	0.698	0.784	0.779	0.837
SE	0.040	0.121	0.053	0.077	0.059	0.047
N	105	33	116	59	71	50
Stunting, children under age 5						
Proportion	0.617	0.592	0.620	0.586	0.665	0.604
SE	0.024	0.029	0.041	0.039	0.037	0.035
N	613	515	800	746	493	436
Underweight, children under age 5						
Proportion	0.124	0.153	0.112	0.137	0.162	0.200
SE	0.013	0.021	0.015	0.022	0.023	0.027
N	613	515	800	746	493	436
Wasting, children under age 5						
Proportion	0.005	0.003	0.013	0.004	0.006	0.016
SE	0.004	0.002	0.005	0.002	0.003	0.011
N	611	514	800	746	491	436
Women's health						
Dietary diversity						
Mean score	4.66	4.61	4.59	4.61	4.43	4.38
SE	0.072	0.093	0.084	0.126	0.110	0.107
N	1,216	971	1345	1,208	829	748
Underweight						
Proportion	0.010	0.011	0.026	0.009	0.018	0.010
SE	0.003	0.003	0.007	0.003	0.004	0.004
N	1,134	996	1,235	1,207	769	697

Table H.3. Subgroup means and standard errors for reproductive health indicators, using the panel of households

	On RVCP Direct Group		On RVCP Indirect Group		On Health Only Group	
Outcome Indicators	Baseline	Midline	Baseline	Midline	Baseline	Midline
Family planning						
Modern family planning prevalence						
Proportion	0.438	0.327	0.458	0.389	0.461	0.374
SE	0.024	0.035	0.028	0.035	0.038	0.033
N	651	498	771	649	460	412
Fertility						
WRA with birth in the past two years						
Proportion	0.191	0.178	0.232	0.205	0.226	0.229
SE	0.015	0.018	0.014	0.017	0.018	0.020
N	1,249	963	1,373	1,206	854	739
Pregnancy outcomes						
At least four prenatal care visits						
Proportion	0.76	0.634	0.739	0.679	0.818	0.775
SE	0.041	0.074	0.041	0.045	0.030	0.039
N	256	179	343	266	197	174
Skilled birth attendance						
Proportion	0.36	0.612	0.506	0.573	0.398	0.503
SE	0.046	0.067	0.053	0.069	0.046	0.062
N	256	179	343	266	197	174
Health facility delivery						
Proportion	0.357	0.44	0.506	0.399	0.395	0.404
SE	0.046	0.077	0.053	0.062	0.047	0.052
N	255	179	343	266	196	174
Postpartum care to mothers within two days of birth						
Proportion	0.782	0.898	0.919	0.754	0.809	0.792
SE	0.037	0.035	0.022	0.053	0.038	0.042
N	199	108	248	150	141	126

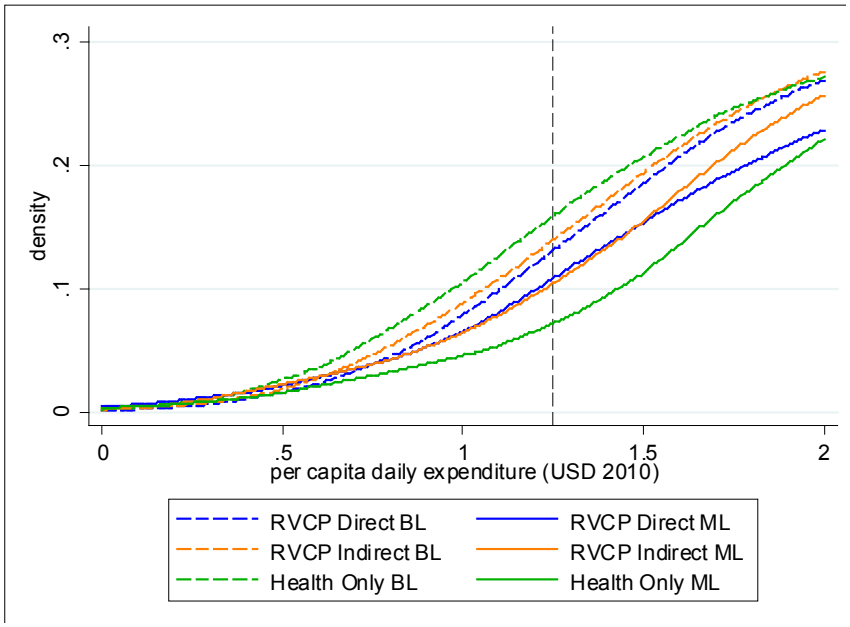
APPENDIX I. CONSUMPTION EXPENDITURE GRAPHS FOR ZOI SUBGROUPS

Figure I.1. Distribution of consumption expenditures, by intervention group and study wave



Notes: The vertical reference line is set at the USD1.25 poverty line (2010 USD). The consumption expenditure distribution is graphed for the panel of households in ZOI areas and is truncated at USD 10.00 per capita daily for display purposes; the graph includes 95.62 percent of ZOI panel households.

Figure I.2. Distribution of consumption expenditures, by intervention group and study wave



Notes: The vertical reference line is set at the USD1.25 poverty line (2010 USD). The consumption expenditure distribution is graphed for the panel of households in ZOI areas and is truncated at USD 2.00 per capita daily for display purposes.

APPENDIX J. MIDLINE DATA QUALITY ASSESSMENT

Accuracy of the Household Panel

Because of concerns related to high attrition and survey data collection at midline, we undertook an assessment to determine the level of agreement between potentially identifying variables at midline and baseline for households included in the panel analyses.

We ran checks on the differences in key demographics between baseline and midline in the panel households. We found the following differences between baseline and midline panel households:

- 27.4 percent of households reported a different predominant language spoken in the household.
- 23.6 percent of households reported a different primary household religion.
- 22.0 percent of households had different responses on whether the household head could read and write at baseline and midline.
- 19.7 percent of households had a different household type (adult male and female, adult male only, adult female only).
- 19.4 percent of households had a change of +/- 3 or more members.
- 15.6 percent of households reported different marital status of the household head.

Although these percentages were higher than expected, we concluded that it was possible that the households comprise a panel, although about 20 percent to 25 percent of them differ on key demographics.

Issues with Midline Data Collection

Indicator Definitions, Data Requirements, and Limitations of the Impact Estimates

The purpose of this document is to describe the indicators requested by the USAID/Guatemala mission for a midline impact analysis. The indicators are separated by level (household, woman, and child), the calculation of the indicator is explained, data requirements and assumptions are outlined, and limitations of an impact analysis using the indicator are discussed.

Overview of Requested Indicators

The 17 requested indicators can be separated according to whether they are at the household, woman, or child level. Several woman/child-level indicators can be further distinguished by their dependence on information from the birth history section of the woman's questionnaire.

Household Level

1. Daily per capita expenditures (2010 USD)
2. Prevalence of poverty: Percentage of people living on less than USD 1.25/day
3. Depth of poverty: Mean percentage shortfall relative to the USD 1.25 poverty line
4. Prevalence of households with moderate or severe hunger

Child Level

5. Prevalence of stunted children under five years of age
6. Prevalence of wasted children under five years of age
7. Prevalence of underweight children under five years of age

Woman Level

8. Women's Dietary Diversity: Mean number of food groups consumed by women of reproductive age
9. Prevalence of modern contraceptive method use among women of reproductive age
10. Prevalence of underweight women
11. Women's Empowerment in Agriculture Index

Indicators Dependent on Birth History

12. Prevalence of exclusive breastfeeding of children under six months of age
13. Prevalence of children ages 6-23 months receiving a minimum acceptable diet
14. Percentage of women of reproductive age who gave birth in the past two years
15. Percentage of women who attended at least four prenatal care visits
16. Percentage of births delivered in a health facility
    a. Percentage of births attended by a skilled health professional
17. Percentage of women receiving postpartum care within two days of birth

Household-Level Indicators

There were a total of 3772 households in the panel of households (domains 1 through 5). There were 6025 baseline households in the panel of clusters and 4405 midline households in the panel of clusters in domains 1 through 5.

Number of households in the panel of households (domains 1 through 5)

Domain	Panel Households	2013 BL only	2015 ML only
1	782	482	146
2	996	750	249
3	602	395	43
4	916	522	129
5	476	380	66
Total	3,772	2,529	633

Notes: Only includes households with completed interviews (hres==1 | unc_complete==1)

Number of households in the panel of clusters (domains 1 through 5)

Domain	2013 BL households not in the panel of clusters	2013 BL households in the panel of clusters	2015 ML households in the panel of clusters
1	37	1,227	925
2	40	1,706	1,248
3	90	907	645
4	19	1,419	1,045
5	90	766	542
Total	276	6,025	4,405

Notes: Only includes households with completed interviews (hres==1 | unc_complete==1). All clusters in domains 1 through 5 at midline were included in the panel of clusters.

We made two key assumptions regarding the household data.

First, we assume that the panel households in domains 1 through 5 were the same households at baseline and midline (reference midline variables PAQUETE2013 and NUMERO_VIVIENDA_2013).

Second, we assumed that the reported household size was accurate. There were some discrepancies across midline files in terms of the number of household members.

Indicator-specific concerns

1. Daily per capita expenditures (2010 USD)
2. Prevalence of poverty: Percentage of people living on less than USD 1.25/day
3. Depth of poverty: Mean percentage shortfall relative to the USD 1.25 poverty line

FTF Indicators EG.3-a (R), EG-a (R), and EG-b (RAA)

Calculation:

Calculations for the three household poverty indicators were undertaken in accordance with the updated June 2016 FTF Indicator Handbook Definition Sheets. The 2013 baseline and 2015 midline household consumption data were converted to 2010 USD using the 2005 base year CPI information. The USD 1.25 international poverty line (2005 PPP) was converted to 2013 and 2015 Quetzales using the overall CPI, base year 2005. A household was considered to be below the USD 1.25 poverty line at baseline if its nominal 2013 daily per capita consumption was below the nominal 2013 USD 1.25 poverty line (Q 8.66), and a household was considered to be below the USD 1.25 line at midline if its nominal 2015 daily per capita consumption was below the nominal 2015 USD 1.25 poverty line (Q 9.01). The depth of poverty relative to the USD 1.25 poverty line was set to zero for nonpoor households.

Data Requirements and Source:

Calculation of the poverty indicators requires consumption-expenditure data, which were collected using the CUESTIONARIO DE GASTOS Y CONSUMO. The ENCOVI definition was used to calculate the household's consumption aggregate. MEASURE Evaluation calculated the consumption aggregate at baseline. Midline data used for the impact analysis come from the consumption aggregate calculated by DevTech.

Households in ZOI (domains 1 through 3), unweighted without clustering

Indicator	Baseline Panel of Households	Baseline Panel of Clusters	Midline Panel of Households	Midline Panel of Clusters
Daily per capita expenditure (USD 2010)				
N	2,365	3,804	2,324	2,753
Mean	4.438	4.412	4.698	4.700
Minimum	0.397	0.397	0.300	0.300
Maximum	37.784	37.784	76.141	86.930
Prevalence of poverty, USD 1.25 per capita daily				
N	2,365	3,804	2,324	2,753
Mean	0.0397	0.039	0.033	0.033
Minimum	0	0	0	0
Maximum	1	1	1	1
Depth of poverty relative to USD 1.25 line				
N	2,365	3,804	2,324	2,753
Mean	0.0072	0.0072	0.0068	0.0067
Minimum	0	0	0	0
Maximum	0.716	0.716	0.785	0.785

Notes: There were five panel households and six panel cluster households with a maximum pc daily expenditure > 40. Only two values appeared to be outliers; for the panel of households, 76.141 was likely an outlier and the second highest value was 49.139, and for the panel of clusters 76.141 and 86.930 were outliers.

Assumptions:

Sample sizes and unweighted descriptive statistics for the ZOI indicate that the consumption aggregate was comparable across baseline and midline data.

Limitations for Impact Analysis:

None.

4. Prevalence of households with moderate or severe hunger

FTF Indicator HL.9-e (RAA)

Calculation:

Calculations for household hunger indicator were undertaken in accordance with the updated June 2016 FTF Indicator Handbook Definition Sheets.

Data Requirements and Source:

Calculation of the household hunger indicator used variables from the household questionnaire, Section 3, Food Insecurity.

Households in ZOI (domains 1 through 3), unweighted without clustering

Indicator	Baseline Panel of Households	Baseline Panel of Clusters	Midline Panel of Households	Midline Panel of Clusters
Prevalence of households with moderate/ severe hunger				
N	2,380	3,840	2,380	2,818
Mean	0.1121	0.1161	0.1029	0.0951
Minimum	0	0	0	0
Maximum	1	1	1	1

Assumptions:

None.

Limitations for Impact Analysis:

None.

Child-Level Indicators (Anthropometrics)

There were a total of five child-level indicators, all of which relied on the child's age in months to set the subsample of eligible children.

The two child-level diet indicators (minimum acceptable diet and exclusive breastfeeding) used variables from the women's questionnaire; we had to calculate the child's age in months from the birth history for these two indicators. The child diet indicators are discussed under "Indicators Derived from Both History Information," below.

The three anthropometric indicators were defined for children under five. In the baseline data, we were able to compare child age and birth dates across multiple survey modules, which is standard practice when assessing the quality of survey data. Although we were able to merge the midline anthropometric data file to the midline household roster at the individual-child level, some cases had inconsistent ages across data files. We were unable to merge the midline anthropometric data file to the child's birth history information from the midline women's questionnaire. Although we would have preferred to verify the child's age in months across midline data files, merging the individual-child across data files was not mandatory for the midline impact evaluation, so we only merged the anthropometric child data to household-level control variables.

Of the children reported to be under five in the anthropometric data file, 10 percent were identified to be older than five in the household roster. There were also some discrepancies between age in months from the anthropometric file and age in completed years from the household roster. We assume that the age in months variable in the midline anthropometric data file was correct and relied on this age variable in our analyses of anthropometric outcomes.

```
. tab P117

P117: NI◆AS |
    Y NI◆OS |
 MENORES DE |
  5 A◆OS DE |
       EDAD |      Freq.     Percent        Cum.
------------+-----------------------------------
         Si |      3,951       90.41       90.41
         No |        419        9.59      100.00
------------+-----------------------------------
      Total |      4,370      100.00

. tab edadmc P106, m

    Edad en |
  meses OMS |
categorizad |                        P106: EDAD EN A◆OS CUMPLIDOS
          a |         0          1          2          3          4         24          . |     Total
------------+---------------------------------------------------------------------------------+----------
  0-5 meses |       440          0          0          1          0          0          1 |       442
 6-11 meses |       427         20          1          2          0          0          4 |       454
12-23 meses |         2        794         23          3          1          1          1 |       825
24-35 meses |         1          8        827         41          2          0          9 |       888
36-47 meses |         1          1         11        803         53          0          7 |       876
48-59 meses |         1          1          2         11        892          0          4 |       911
------------+---------------------------------------------------------------------------------+----------
      Total |       872        824        864        861        948          1         26 |     4,396
```

5. Prevalence of stunted children under five years of age
6. Prevalence of wasted children under five years of age
7. Prevalence of underweight children under five years of age

FTF Indicators HL.9-a (R), HL.9-b (R), and HL.9-c (R)

Calculation:

Calculations for child anthropometric indicators were undertaken in accordance with the updated June 2016 FTF Indicator Handbook Definition Sheets.

Data Requirements and Source:

Calculation of the child anthropometric indicators used data from the anthropometric survey module.

Children ages 0-59 months in ZOI (domains 1 through 3) households, unweighted without clustering

Indicator	Baseline Panel of Households	Baseline Panel of Clusters	Midline Panel of Households	Midline Panel of Clusters
Prevalence of stunted children under five years of age				
N	1,935	3,127	1,708	2,045
Mean	0.646	0.657	0.606	0.611
Minimum	0	0	0	0
Maximum	1	1	1	1
Prevalence of wasted children under five years of age				
N	1,935	3,127	1,708	2,045
Mean	0.144	0.154	0.155	0.153
Minimum	0	0	0	0
Maximum	1	1	1	1
Prevalence of underweight children under five years of age				
N	1,931	3,123	1,707	2,044
Mean	0.008	0.006	0.007	0.006
Minimum	0	0	0	0
Maximum	1	1	1	1

Assumptions:

Because the birth dates and age in months were cleaned in the midline anthropometric file we received, we assumed that the age information in the anthropometric data file was correct, even if it conflicted with age information in the household roster.

It also seemed that some children who were measured but had nonvalid z-scores may have been dropped from the midline anthropometric data file received by MEASURE Evaluation. We were unable to confirm this, but if it was the case, then this affected the response rates and sample weight adjustments because we would not make an adjustment for a child who was measured but had a nonvalid z-score, but we would make an adjustment if that child was not in the anthropometric data set. A simple tabulation of flagged z-scores at baseline and midline illustrates this concern:

Children under five with flagged z-scores (domains 1 through 5)

	# Flagged Z-Scores	
	Baseline	**Midline**
WEI	6/5,265 (0.11%)	1/3,263 (0.03%)
LEN	23/5,265 (0.44%)	10/3,265 (0.31%)
EFL	20/5,239 (0.38%)	0/3,268
BMI	20/5,256 (0.38%)	2/3,265 (0.06%)

Given the low number of flagged cases in both the baseline and midline survey rounds, we do not expect that this created a substantial problem for the midline impact analysis.

Limitations for Impact Analysis:

The z-scores were adjusted for the sex and age of the child in months. Using incorrect age information can introduce measurement error in the anthropometric z-scores. It is standard practice to verify child age using age and birth date variables from multiple survey sections, but our ability to conduct this quality check was very limited. It seems that 10 percent of children could be counted as under five when they were in fact older than five. Because we could not merge the anthropometric age variables with the birth history data, we were unable to check the consistency of the more specific age in months. We need to assume that any measurement error due to incorrect age specification did not differ systematically between treatment and comparison households, maintaining the validity of the impact estimates.

Our inability to verify whether all measured children were included in the midline anthropometric file could result in some children having incorrect sampling weight adjustments at midline. This is likely a small issue and should not significantly influence our impact estimates.

Woman-Level Indicators

8. Women's dietary diversity: Mean number of food groups consumed by women of reproductive age

FTF Indicator HL.9.1-c (O)

Calculation:

Calculated per the June 2016 FTF Indicator Handbook Definition Series.

Data Requirements and Source:

This indicator was defined for women of reproductive age (15-49 years) and used data from the women's questionnaire.

9. Prevalence of modern contraceptive method use among women of reproductive age

Non-FTF Indicator

Calculation:

This indicator was calculated in accordance with the Guatemala ENSMI and was defined for women ages 15 to 49 years who were married or in a consensual union. Modern methods comprised female sterilization, male sterilization, IUD, injectables, implant, pill, condom, and lactational amenorrhea. Natural methods were the rhythm method and withdrawal.

10. Prevalence of underweight women

FTF Indicator HL.9-d (R)

Calculation:

Calculated per the June 2016 FTF Indicator Handbook Definition Series.

Data Requirements and Source:

This indicator was defined for nonpregnant women of reproductive age (15-49 years) and used data from the anthropometric module.

Women ages 15 to 49 years in ZOI (domains 1 through 3) households, unweighted without clustering

Indicator	Baseline Panel of Households	Baseline Panel of Clusters	Midline Panel of Households	Midline Panel of Clusters
Women's dietary diversity				
N	3,390	5,340	2,936	3,506
Mean	4.571091	4.533521	4.56233	4.560468
Minimum	1	0	1	1
Maximum	9	9	9	9
Prevalence of modern contraceptive method use				
N	2,098	3,355	1,719	2,066
Mean	0.397521	0.394337	0.337987	0.338335
Minimum	0	0	0	0
Maximum	1	1	1	1
Prevalence of underweight women				
N	3,138	4,924	2,915	3,444
Mean	0.0198	0.0207	0.0137	0.0139
Minimum	0	0	0	0
Maximum	1	1	1	1

Assumptions:

None for contraceptive method use or dietary diversity. For prevalence of underweight women, we had similar concerns as with the child anthropometric indicators: if only women with valid height and weight measurements were included in the midline data file received by MEASURE Evaluation (i.e., they excluded women who were measured but whose weight or height measurements had problems), we would not be able to correctly adjust midline sampling weights for woman-anthropometric indicators. This was likely to be a mild concern.

Limitations for Impact Analysis:

No significant concerns.

11. Women's Empowerment in Agriculture Index

FTF Indicator EG.3-b (R)

The empowerment survey module was only implemented among a subsample of households in the ZOI at baseline. Because we do not have baseline data for comparison households, we were unable to estimate program impact on the Women's Empowerment in Agriculture Index indicators.

Indicators Derived from Birth History Information

12. Prevalence of exclusive breastfeeding of children under six months of age

FTF Indicator HL.9.1-b (RAA)

13. Prevalence of children ages 6-23 months receiving a minimum acceptable diet

FTF Indicator HL.9.1-a (RAA)

Calculation:

The breastfeeding and minimum acceptable diet indicators were calculated per the June 2016 FTF Indicator Handbook Definition Sheets.

Data Requirements and Source:

Data for both indicators came from the women's questionnaire. The breastfeeding indicator was defined for children ages 0 to 5 months old, and the minimum acceptable diet was defined for children ages 6 to 23 months old. These indicators were only reported for each woman's most recent live birth to maintain consistency with the ENSMI methodology.

The remaining indicators were non-FTF indicators and were computed using the ENSMI methodology. The indicators were calculated for women ages 15 to 49 and were restricted to the woman's most recent live birth in the past two years. Data for these indicators were taken from the women's questionnaire. By restricting our analysis to the past two years, we were comparing equal time frames (two years) that did not overlap.

14. Percentage of women of reproductive age who gave birth in the past two years

Calculation:

This indicator was defined for women of reproductive age (15 to 49 years) and was restricted to births in the two years before the wave-specific interview (e.g., baseline births occurred between 2011 and 2013 and midline births occurred between 2013 and 2015).

15. Percentage of women who attended at least four antenatal care visits
16. Percentage of births delivered in a health facility

Calculation:

Health sites consisted of MSPAS facilities, private hospitals/clinics, Asociación Pro-Bienestar de la Familia de Guatemala facilities, Instituto Guatemalteco de Seguridad Social sites, and other public and private medical facilities.

16.a. Percentage of births attended by skilled health professional

Calculation:

Skilled health professionals were physicians or ambulatory physicians, nurses, and trained midwives. If multiple personnel attended the delivery, only the most qualified was considered.

17. Percentage of women receiving postpartum care within two days of birth

Calculation:

Calculation made in accordance with DHS (and ENSMI) definition.

Data Requirements:

To be consistent with the ENSMI methodology, we only reported these indicators for each woman's most recent live birth. This meant that we had to be able to identify which child was the woman's most recent birth. There were several different variables throughout the woman's questionnaire that indicate whether the child is the most recent birth; unfortunately, only two of these variables were included in the midline data files and a comparison of these variables suggests that the data were not always for the most recent birth.

```
Contains data from U:\Activities\Guatemala_WHIP\Midline
data\Midline_data_5.31.16\EMEPAO15_MUJER.dta
  obs:         7,115
 vars:         1,473
 size:   192,546,130
Sorted by:

. d MP212_01 MP501F_1 MP553_1 (variables specific to most recent birth)

MP212_01 No. de linea
MP501_1 No. de orden en historial de nacimientos
MP553_1 Comio, pan, tortilla, arroz, fideos, pastas u otra comida de granos (for
most recent live birth)

. tab MP212_01, m

 MP212. No. |
   de linea |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |      4,782       67.21       67.21
          . |      2,333       32.79      100.00
------------+-----------------------------------
      Total |      7,115      100.00

. tab MP501F_1 MP212_01, m

   MP501F. |
    No. de |
  orden en |
 historial |
```

de nacimientos	MP212. No. de linea 1	.	Total
0	2	0	2
1	2,539	0	2,539
2	108	0	108
3	86	0	86
4	74	0	74
5	79	0	79
6	55	0	55
7	69	0	69
8	39	0	39
9	26	0	26
10	24	0	24
11	10	0	10
12	8	0	8
13	6	0	6
14	5	0	5
16	2	0	2
.	1,650	2,333	3,983
Total	4,782	2,333	7,115

```
. tab MP501F_1 MP553_1, m
```

Here we should only see MP551_1 responses for children with MP501F_1==1. There are responses for MP551_1 that come from children who are not the most recent live birth if we assume that the number contained in the MP501F_1 is the actual birth order.

MP501F. No. de orden en historial de nacimientos	MP553. Comio, pan, tortilla, arroz, fideos, pastas u otra comida de granos			
	Si	No	sabe	.	Total
0	0	0	0	2	2
1	794	462	2	1,281	2,539
2	31	15	0	62	108
3	20	23	1	42	86
4	24	14	0	36	74
5	28	6	0	45	79
6	13	11	0	31	55
7	13	15	0	41	69
8	13	8	0	18	39
9	7	6	0	13	26
10	7	10	0	7	24
11	3	5	0	2	10
12	3	1	0	4	8
13	4	1	0	1	6
14	1	3	0	1	5
16	1	1	0	0	2
.	2	0	0	3,981	3,983
Total	964	581	3	5,567	7,115

Each of these indicators also required assigning an individual child an age in months. The interview date, birthdate, and age variables were included in the household roster, the birth history section of the women's questionnaire, and the anthropometric questionnaire. It is common practice to check the quality of age in months data by comparing multiple variables related to the child's age across survey modules. Unfortunately, we were not able to link individual children from the women's questionnaire birth history to the household roster or to the anthropometric data and, therefore, we were unable to implement birth date and age data quality checks, and we could not use data from elsewhere in the survey to fill in age values that were missing or incorrectly entered. To facilitate cleaning of the midline data at MEASURE Evaluation, we only cleaned birth history birth dates and age in months for children identified in the birth history as the most recent birth.

Assumptions:

We made the assumption that at midline, the enumerator correctly followed the birth history module guidelines when conducting the women's interview and that the child who was reported first was indeed the most recent birth. This meant that we deferred to MP212_01==1 as indicating that the child was the most recent birth, even if this was inconsistent with other birth order variables for the same case.

Fortunately, restricting the cleaning of child age data to the most recent live birth (as indicated by MP212_01==1) resulted in fewer than 3 percent of observations being flagged for drastic inconsistencies with birth year, birth month, or both.

Limitations for Impact Analysis:

Implications of not knowing for sure that we were using midline data from the most recent birth were not serious for the calculation of program impact on these indicators because the indicators were also restricted by age, so even if we did not get the most recent birth, we would still be using an age-eligible child (assuming the age was correct).

For exclusive breastfeeding of children ages 0 to 5 months, we could have overestimated the program impact if children who were ages 0 to 5 months appeared to be older in the midline data and were thus excluded from the indicator. We could have underestimated the program impact if children who were older than 5 months—and therefore, would likely have begun complementary infant feeding—were registered as being ages 0 to 5 months because these children would be less likely to be exclusively breastfed simply due to their older age.

There was a similar concern for children receiving a minimum acceptable diet where change in the diet we attributed to the program could have been due to measurement error in the child's age in months.

MEASURE Evaluation
University of North Carolina at Chapel Hill
123 W. Franklin Street, Suite 330
Chapel Hill, NC 27516 USA
Phone: +1 919-445-9350
measure@unc.edu
www.measureevaluation.org

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. TRE-18-010; ISBN: 978-1-64232-049-7