

Impact Evaluation of Feed the Future/USAID-ACCESO: Agriculture and Nutrition Activities in Western Honduras from 2012-2015

January 2016

Prepared for the United States Agency for International Development

It was prepared by Almanzar, Miguel, and Maximo Torero of the International Food Policy Research Institute (IFPRI)

Last revision: May 2016

Recommended Citation:

Almanzar, Miguel, and Maximo Torero. 2016. *“Impact Evaluation of Feed the Future/USAID-ACCESO: Agriculture and Nutrition Activities in Western Honduras from 2012-2015.”* International Food Policy Research Institute. Washington, DC.

Contact:

Miguel Almanzar, International Food Policy Research Institute
1201 Eye K Street NW, Washington, DC 20005
m.almanzar@cgiar.org

TABLE OF CONTENTS

List of Acronyms.....	iii
List of Tables	iv
List Of Figures.....	vii
Executive Summary.....	viii
a. Overview of Compact and Intervention(s) Evaluated.....	viii
b. Evaluation type, questions, methodology	viii
c. Findings	x
Agricultural Production	x
Labor Markets:	xi
Expenditures and Poverty	xi
Hunger, Child Health & Nutrition, Maternal Health & Nutrition	xi
Women’s Empowerment	xi
d. Conclusions	xi
1 Introduction and Project Background.....	1
2 Impact Evaluation Methodology	4
3 Geography, Treatment Assignment, and Attrition.....	8
3.1 Balancing Weights.....	10
3.2 Attrition Weights.....	15
3.3 Ex-Post Power Analysis And Simulation.....	24
4 Results	29
4.1 Agricultural Production and Market Participation.....	30
Distribution of crops	30
Use of Inputs: Fertilizer, Fungicides, Labor, and Land	35
Agriculture Production and Sales.....	50

Land Productivity or Crop Productivity	62
Coffee rust.....	65
4.2 Economic Well-Being	73
Income and Labor Supply.....	73
Expenditures	80
4.3 Nutrition and Health	85
Food Security: Hunger Scale	85
Child Nutrition and Health	87
Child Anthropometry: Stunting, Wasting, and Underweight	92
Child Anemia	93
Maternal Nutrition and Anemia.....	97
4.4 Women’s Empowerment and Time Use	101
5 Conclusions	110
Discussion of Impacts and Lessons for Future Interventions	111
Lessons for the future	112
6 References.....	114
Appendix A – Decomposition of Women Empowerment Index in Agriculture	116
Appendix B – Impact Estimates: Other Outcomes.....	120
Appendix C – Sample Methodology.....	130
Sample Design.....	130
Power Analysis	141
Appendix D - Average Treatment on the Treated Impact estimates.....	145
The Sub-sample of Households Reporting Participation in USAID-ACCESO	145
Appendix E - Statement of Differences.....	166

LIST OF ACRONYMS

5DE	Five Domains of Empowerment
BMI	Body Mass Index
GPI	Gender Parity Index
HHS	Household Hunger Scale
IFPRI	International Food Policy Research Institute
MAD	Minimum Acceptable Diet
SD	Standard Deviation
USAID	United States Agency for International Development
USG	United States Government
WEAI	Women's Empowerment in Agriculture Index
ZOI	Zone of Influence

LIST OF TABLES

Table 1 Beneficiary Attribution	9
Table 2 Composition of the Treatment Group by Status in Implementer's Beneficiary List	9
Table 3 Logit Model to Estimate the Probability of Treatment	12
Table 4 Unweighted and Weighted Matching Variables	14
Table 5 Unweighted and Weighted Comparison of Additional Variables at Baseline	15
Table 6 Sample Attrition: Percent of Households Surveyed Only Once	16
Table 7 Household Characteristics of Attriters	18
Table 8 Child and Maternal Health Characteristics of Attriters	19
Table 9 Probit for the Probability of being in both surveys: Correlates of attrition	22
Table 10 Unweighted and Weighted (for Treatment Assignment and Attrition) Characteristics at Baseline	23
Table 11 Probability Household Grows Corn, Beans or Coffee	33
Table 12 Prob. Household Grows vegetables, fruits, or tubers after ACCESO	34
Table 13 Inputs: Use of chemical and organic fertilizers, herbicides, and fungicides	37
Table 14 Inputs for Corn, Beans and Coffee: Use of fertilizer, herbicides, and fungicides	38
Table 15: Inputs for Vegetables, Fruits and Tubers: Use of fertilizer, herbicides, and fungicides	39
Table 16 Input costs for Corn, Beans and Coffee: Products and Labor	41
Table 17: Input costs for Vegetables, Fruits, and Tubers: Products and Labor	42
Table 18 Descriptive Statistics of Area Planted (ha.)	44
Table 19 Area Planted (Ha.): All, Corn, Beans, and Coffee	46
Table 20 Area Planted (ha.): Vegetables, Fruits, and Tubers	47
Table 21 Percentage area Planted (Ha) Vegetables, Fruits, and Tubers	48
Table 22 Percentage area planted (Ha) Corn, Beans, and Coffee	49
Table 23 Total Production (Kgs.)	54
Table 24 Total Production (Kgs) and Value: Corn, Beans, and Coffee	55
Table 25 Total Production (Kgs) and Value: Vegetables, Fruits, and Tubers	56
Table 26: Log Total Production (kgs) and Values: Corn, Beans, and Coffee	57
Table 27: Log Total Production (kgs) and Values: Vegetables, Fruits, Tubers	58
Table 28 Value of households' agricultural production	59
Table 29 Agriculture Sales revenue in 2005PPP to 2010 USD	60

Table 30 Ag. Sales revenue household aggregate impact with alternative prices	61
Table 31 Land productivity: Yield (Kg/Ha) for Corn, Beans, Coffee	64
Table 32 Descriptive Statistics for Coffee Rust: Belief, Prevention, and Control	68
Table 33 Coffee Rust: Impact on Assistance and Practices.....	71
Table 34 Coffee Rust: Impact on Area, Production, and Costs	72
Table 35 Labor market participation	74
Table 36 Impacts on Income by Source	76
Table 37 Impact of Labor Income from primary activity	78
Table 38 Impact of Labor Income from three principal activities.....	79
Table 39 Impact on Expenditure Measures by type	82
Table 40 Impact on total expenditure	83
Table 41 Impact on Poverty.....	84
Table 42 Impacts of Food Security.....	86
Table 43 Impact on Breastfeeding Practices.....	88
Table 44 Impact on Children Dietary Diversity	90
Table 45 Impact on Consumption of iron-rich foods, vitamins and hydration solutions (6-23).....	91
Table 46 Impact on anthropometric z-scores	94
Table 47 Impact on children underweight, stunting, and wasting	95
Table 48 Impact on Child anemia and edema	96
Table 49: Impact on Maternal Nutrition Outcomes	98
Table 50 Impact on Maternal Anthropometric Outcomes	99
Table 51 Impact on anemia in women of reproductive age	100
Table 52 Empowerment in Agriculture Index by Year	109
Table 53 Empowerment in Agriculture Index by Treatment Status in 2013 and 2015	109
Table 54 Censored headcount from 5DE Index by Survey Year	116
Table 55 5DE Decomposed by Dimension and Indicator by Survey Year	117
Table 56 Censored headcount from 5DE Index by Treatment Status in 2013.....	118
Table 57 Decomposed by Dimension and Indicator by Treatment Status in 2013 and 2015	119
Table 58 Impact on household expenditures	120
Table 59 Impact on diarrhea in children under 5	121
Table 60 Impact on feeding practices for children 6-23 months.....	122
Table 61 Impact on birth control use.....	122

Table 62 Impact on female child anthropometric measures.....	123
Table 63 Impact on female child wasting, stunting and underweight indicators.....	124
Table 64 Impact on female child anemia	125
Table 65 Impact on male child anthropometric measures.....	126
Table 66 Impact on male child wasting, stunting and underweight indicators.....	127
Table 67 Impact on male child anemia	128
Table 68 Impact on log income and cost by type of work: principal activity	129
Table 69 Simulation Scenarios	133
Table 70 Simulation of a decrease of 30,000 households in ACCESO's area of influence.....	139
Table 71 Sample size calculations.....	140
<i>Table 72 ATT Impact on total expenditure</i>	<i>148</i>
<i>Table 73 ATT Impact on Poverty</i>	<i>149</i>
<i>Table 74 ATT Impact on Breastfeeding Practices</i>	<i>150</i>
<i>Table 75 ATT Impact on anthropometric z-scores.....</i>	<i>151</i>
<i>Table 76 ATT Impact on children underweight, stunting, and wasting</i>	<i>152</i>
<i>Table 77 ATT Impact on Child anemia</i>	<i>153</i>
<i>Table 78 ATT Impact on Women's dietary diversity</i>	<i>154</i>
<i>Table 79 ATT Land productivity: Yield (Kg/Ha) for Corn</i>	<i>155</i>
<i>Table 80 ATT Ag. Sales revenue household aggregate impact with alternative prices for Corn</i>	<i>156</i>
<i>Table 81 ATT Land productivity: Yield (Kg/Ha) for Beans</i>	<i>157</i>
<i>Table 82 ATT Ag. Sales revenue household aggregate impact with alternative prices for beans</i>	<i>158</i>
<i>Table 83 ATT Probability Household Grows Corn, Beans or Coffee</i>	<i>159</i>
<i>Table 84 ATT Prob. Household Grows vegetables, fruits, or tubers.....</i>	<i>160</i>
<i>Table 85 ATT Impact of Labor Income from three principal activities</i>	<i>161</i>
<i>Table 86 ATT Impacts on Income by Source</i>	<i>162</i>
<i>Table 87 ATT Labor Market Participation</i>	<i>163</i>
<i>Table 88 ATT Coffee Rust: Impact on Assistance and Practices</i>	<i>164</i>
<i>Table 89 Coffee Rust: Impact on assistance and area affected</i>	<i>165</i>

LIST OF FIGURES

Figure 1 Sample Power to Detect Changes in Expenditures.....	26
Figure 2 Sample Power to Detect Changes in Extreme Poverty	27
Figure 3 Sample Power to Detect Changes in Poverty	28
Figure 4 Distribution of Crops.....	32
Figure 5 Distribution of Area Planted	43
Figure 6 Prices paid to growers in exporting countries, Honduras, 2008-2015 (US cents/lb.)	66
Figure 7 Proportion of Production Affected by coffee rust	69
Figure 8 Proportion of Women that are EMPOWERED by indicator in the ZOI of FTF in Honduras by Year	105
Figure 9 Proportion of Women that are EMPOWERED by indicator in the ZOI of FTF in Honduras by Treatment Status in 2013 and 2015	106
Figure 10 5DE Decomposed by Dimension and Indicator for each year	107
Figure 11 5DE Decomposed by Dimension and Indicator for each Treatment Status in 2013and 2015 ..	108
Figure 12 Geographical Distribution of Beneficiary Households by Poverty Classification.....	134
Figure 13 Graphical Representation of the Sample Design	138
Figure 14 Standardized effect size (MDE) vs. Power under 3 survey waves for continuous outcomes...	142
Figure 15 Standardized effect size (MDE) vs. Power at baseline for continuous outcomes	143
Figure 16 Power vs. Number of cluster vs. power for binary outcome.....	143
Figure 17 Power vs. Prevalence of poverty in treatment group, binary outcome	144

EXECUTIVE SUMMARY

A. OVERVIEW OF COMPACT AND INTERVENTION(S) EVALUATED

USAID-ACCESO was a four-year activity funded by the USAID Office of Economic Growth in Honduras. Implemented by FINTRAC from 2011-2015, the project aimed to increase household income in order to directly lift 10,000 beneficiary households out of poverty (7,500 out of extreme poverty), as well as to decrease under-nutrition in these households. Until 2015, the activity represented the core investment by USAID-Honduras in the U.S. Government's (USG) Global Hunger and Food Security Initiative known as the Feed the Future Initiative (FTF) and worked through various components to enable economic development at the household level. These components included:

1. Providing agricultural and value-added technical assistance and training to enhance the capacity of Honduras's poorest households in production, management, and marketing skills;
2. Facilitating market access by linking farmers to market opportunities;
3. Increasing the availability of rural financial services through existing intermediaries, village banks, commercial banks, and other service and input providers;
4. Assisting in the elimination of policy barriers that impede rural households' access to market opportunities;
5. Engaging in dietary diversification and malnutrition prevention activities to enhance the capacity of rural households to improve their utilization and consumption of food; and
6. Establishing sound environmental and natural resource management.

USAID-ACCESO targeted six of Honduras's poorest departments in the Occidental region: Copán, Santa Bárbara, Ocotepeque, Lempira, Intibucá, and La Paz. The majority of farm households in these departments cultivate traditional crops on small plots, often on hillsides where access to markets is hindered by poor roads and long distances. In addition, the use of traditional cultural practices produces poor yields, depletes soil nutrients, and leads to forest encroachment.

B. EVALUATION TYPE, QUESTIONS, METHODOLOGY

This document lays out impact results using the 2012 baseline survey, the 2013 follow-up (midline) survey, and the second follow-up (endline) survey administered in 2015 for the impact evaluation of USAID-ACCESO. From May to July of 2012, 3,326 households in the Occidental part of Honduras were

interviewed to collect data on income and expenditures, access to credit, time use, use of land, agricultural practices, nutritional status of women of reproductive age and children under five years of age, and women's empowerment. The midline survey was conducted between May and July 2013 and administered the same questionnaire to all of the originally interviewed households that could be located. A third and final survey was implemented from May to July 2015, again to all of the originally interviewed households that could be located.

Our baseline sample included 1,256 households that were randomly selected from the beneficiary list available at baseline or that reported participating in the program; the remainder of the sample (2,070 households) were comparison households selected randomly from the 2001 Honduran Population Census. The comparison group was selected such that the sample was representative of living conditions in each of the six departments in which USAID-ACCESO operated.

The effect of USAID-ACCESO was evaluated with a rigorous, non-experimental design that incorporated matching, a panel survey, and econometric analysis. The impact of the program on various outcomes is estimated using a matched household-level difference-in-differences (DID) estimator with household fixed effects. The impact evaluation measures the changes in outcomes which can be attributed to a specific intervention. For example, the change in household expenditures in the comparison group (expenditures after the program minus expenditures before the program) is subtracted from the change in expenditures of the treatment households in order to estimate the impact of the program on household expenditures. Comparison households are used to account for changes in economic conditions in the area that could affect the income of treated households, even in the absence of treatment. The sample size for the panel survey was chosen to provide 80% power to detect a 10% increase in household expenditures, at a 5% significance level.

Our main goal is to assess the impact of participation in ACCESO's programs and extension services on a variety of outcomes, such as the households' agricultural production, health, and income. Simply comparing households that have participated in the program to households that have not may cause us to attribute differences to ACCESO that were either preexisting or were caused by something other than the program, and thus to either over- or underestimate the program's impact. Comparing households that participated in ACCESO before and after the program runs a similar risk; we do not know what the changes we observe would have happened in the absence of the program, and thus we may attribute changes to ACCESO that would have happened regardless. Comparing changes over time among households that were exposed to ACCESO's extension services to changes over time for households that

did not participate in the program allows us to account for initial differences between treated and comparison households, as well as changes over time that affect both groups. In order to ensure that there are no systematic differences between treatment and control households that would suggest they have different trends for our indicators of interest, we compare similar households between both groups. In order to match households appropriately, we construct a propensity score based on pre-treatment values of observable characteristics and use the inverse of the score to weight observations in our regression analysis. It is important to note that a household is classified as “treated” if it was on the initial list of beneficiaries at baseline or if it ever reported being part of USAID-ACCESO, regardless of whether the household actively participated in the program. The impact estimates in the report thus measures “intention to treat.”

C. FINDINGS

We estimate the impact of USAID-ACCESO across all the main outcomes monitored by Feed the Future and added other outcomes of interest to provide a narrative for the impact pathway when impacts are observed or to provide context when the expected impacts are not found. We find some significant effects across the main outcome measures; however, most are not robust across specifications. To summarize the main findings:

Agricultural Production: We find an increase in the coffee and corn area planted by endline, as well as some evidence that suggests an increase in crop diversification to vegetable and tuber production. We find changes in the use of inputs among treated farmers; households cultivating coffee, fruit, and tubers increased their use of chemical fertilizer, while households cultivating beans decreased their use of organic fertilizers. We see that vegetable farmers decreased their use of fungicides and herbicides, while fruit farmers increased their use of the same. We see that labor costs increased significantly for coffee and fruit farmers in the treatment group; for coffee, this is consistent with treatment farmers being able to spend more on labor to prevent coffee rust, while for fruit, this might be because treated farmers are able to cultivate long-term, high-value crops with the support given by ACCESO.

Overall, we find that agricultural production increased by 2000 kg overall, driven by a 4000 kg increase at endline. In general, the agricultural impact estimates suggest that negative impacts on the quantity of agricultural production and land productivity in the midline were mitigated by increases at endline. We find decreases in the value of production for most of our crops, but the results for corn, beans, coffee, and tubers are driven by wide variations in the reported price. When we test alternative specifications,

such as using an inverse hyperbolic sine transformation (to account for over dispersion) or using department median prices instead of reported prices (to correct for households misreporting prices), we do not see any effect.

Labor Markets: We find an 11 percentage point increase in participation in wage labor and a 9 percentage point increase in participation in the agricultural sector. This increase in participation in wage labor is coupled with a \$300 increase in household income from wage labor in treatment households.

Expenditures and Poverty: We find no impact on household expenditures or on the proportion of individuals falling below a variety of poverty lines. We detect no differential change in aggregate household expenditures among treated households, nor do we see any evidence of changing expenditures among subcategories of household expenditures (food, non-food, utilities, housing).

Hunger, Child Health & Nutrition, Maternal Health & Nutrition: We see an imprecisely measured 1.8 percentage point decrease in the proportion of households that experience moderate to severe hunger at endline, and we see that treated women have a slightly more diverse diet (scoring 0.17 higher on the 9 food dietary diversity scale). Aside from an imprecisely measured 1 cm change in the height of treated women, we find no impacts on women's health or nutrition; specifically, there is no change in women's BMI or prevalence of anemia. In addition, we find no evidence that children's nutrition improved; we detect no effects on breastfeeding, minimum meal frequency, or dietary diversity. We also see no corresponding changes in child anthropometrics, as the prevalence of stunting, wasting, and anemia in children remains the same. When we disaggregate children's outcomes by sex, we find a decrease in the proportion of underweight boys, but no impact on girls.

Women's Empowerment We find increasing trends in women's empowerment during the study period and similar dynamics between treatment and comparison households. The most salient result is that differences in empowerment between men and women in the ZOI are driven mostly from differences in access to resources, assets, and credit.

D. CONCLUSIONS

We detect no effects on some measures that a priori should be affected by USAID-ACCESO, including agricultural productivity, agricultural sales, nutritional status, household expenditures, and poverty. We expected that by the endline survey, households would have changed their agricultural and nutritional

practices as they participated in USAID-ACCESO for longer periods of time and that these changes in practice would translate into improvements in final outcomes. However, we observe no changes in intermediate outcomes or in the main final outcomes.

That the project had no significant effects on nutritional indicators is not surprising. USAID-ACCESO was designed to prevent malnutrition rather than to reverse damage caused by past under-nutrition. To detect impacts on intermediate indicators, such as dietary diversity and minimum meal frequency, the intervention should have specifically targeted undernourished children and provided households with extra resources to improve child nutrition rather than just relying on household production to meet nutritional needs. We also note that the nutrition component of USAID-ACCESO had limited coverage in the zone of influence; only 225 communities received the nutrition component, in comparison with over 2,700 communities that participated in the agricultural component.

In the case of agricultural productivity and market access indicators, we expect that longer exposure to treatment, through new interventions increasing access to irrigation, and reinforcement of good agricultural practices will improve these indicators. One possible explanation for why we did not observe significant impacts on these indicators may be the result of large aggregate shocks. Although aggregate shocks are not likely to differentially affect the treatment and comparison groups, such a shock (like that experienced with recent outbreaks of coffee rust) could decrease the resources available to all households and, as a consequence, the take-up of new practices and the introduction of more nutritious foods.

1 INTRODUCTION AND PROJECT BACKGROUND

USAID-ACCESO was a four-year activity funded by the USAID Office of Economic Growth in Honduras. Implemented by FINTRAC from 2011 – 2015, the activity aimed to increase household income in order to directly lift 10,000 beneficiary households out of poverty (7,500 out of extreme poverty), as well as to decrease under-nutrition in these households. Until 2015, the activity represented the core investment by USAID-Honduras in the U.S. Government's (USG) Global Hunger and Food Security Initiative known as the Feed the Future Initiative (FTF) and worked through various components to enable economic development at the household level. These components included:

1. Providing agricultural and value-added technical assistance and training to enhance the capacity of Honduras's poorest households in production, management, and marketing skills;
2. Facilitating market access by linking farmers to market opportunities;
3. Increasing the availability of rural financial services through existing intermediaries, village banks, commercial banks, and other service and input providers;
4. Assisting in the elimination of policy barriers that impede rural households' access to market opportunities;
5. Engaging in dietary diversification and malnutrition prevention activities to enhance the capacity of rural households to improve their utilization and consumption of food; and
6. Establishing sound environmental and natural resource management.

USAID-ACCESO targeted six of Honduras's poorest departments: Copán, Santa Bárbara, Ocotepeque, Lempira, Intibucá, and La Paz. The majority of farm households in these departments cultivate traditional crops on small plots, often on hillsides where access to markets is hindered by poor roads and long distances. In addition, the use of traditional cultural practices produces poor yields, depletes soil nutrients, and leads to forest encroachment.

Agricultural technical assistance and training first aimed to ensure that households were able to achieve food security through improvements in the production of basic crops (corn and beans); once food security was achieved, ACCESO encouraged households to cultivate income-generating horticultural crops. Technicians encouraged producers to adopt practices that had the potential to double or triple yields of basic grains, such as introducing soil preparation practices; increasing planting densities; implementing weed control; improving fertilization use, safe and correct use of pesticides, and record-

keeping; and providing farm certification options and post-harvest handling. These basic practices and associated skills not only had the potential to increase productivity but could also have reduced the risk of climate-related losses. Providing basic grain producers with low- or no-cost technologies enabled them to increase yields and reduce production and post-harvest losses. Increased yields would help households meet their consumption needs even while devoting less acreage to subsistence crops, thus freeing up land for the cultivation of more profitable horticultural crops. USAID-ACCESO's activities benefitted a large number of rural households, helping them meet their subsistence needs, reduce outside expenditures, and provide some supplemental income through surplus sales.

In addition, USAID-ACCESO engaged in collaborative activities to increase access to new technologies. Alliances with private companies that distribute irrigation and other agricultural equipment and clean energy production equipment, or that promote health, natural resource management, and disaster mitigation, complement assistance across USAID-ACCESO's agricultural and nutrition themes¹.

Value-added training topics included cost structure development; equipment recommendations; good manufacturing practices; legal issues; new market contacts; suppliers for packaging materials; and improvements in processing techniques to reduce the cost of production and to increase efficiencies.

The key need for credit among farmers was the purchase of inputs prior to planting and during production. Farmers often demand input amounts that are smaller than the quantity in which those inputs are sold, which makes it more difficult for them to get access to the appropriate inputs. USAID-ACCESO facilitated cooperation between input suppliers and community banks to create group buying schemes that address this constraint. These groups were directed by the community banks, to which farmers have a post-harvest liability.

The health and nutrition initiatives of USAID-ACCESO were directed at women and children in order to better affect households' overall nutritional status and dietary habits. The main activities under this component included:

1. The promotion and improvement of breast-feeding practices through female support groups during and after pregnancy. Breast milk has a high caloric content and strengthens children's immune systems, which prevents undernourishment and reduces infant mortality.

¹ Some examples include ecological coffee processing equipment, corn mills, water storage equipment, and food processing equipment.

2. The promotion and improvement of feeding practices for children less than 2 years of age, pregnant women, and lactating women in beneficiary households. This intervention targets the 1,000 day period in which stunted growth and undernourishment can affect a child's brain development, affecting them for the rest of their lives. This work aims to integrate children into the family diet smoothly and successfully, with breast-feeding and other high-nutrient food products acting as a complement during the transition period.
3. The introduction and support of production of high-nutrient fruits and vegetables produced both commercially and within kitchen gardens in order to increase the diversity of foods available to households and communities.
4. The promotion of "healthy household" practices aimed to reduce children's exposure to diarrheal and respiratory disease. Elements included covering dirt floors and walls with cement, introducing fuel-efficient wood-burning stoves, proper use of latrines, storage of clean water inside the home, and storage of animals outside the home.

USAID-ACCESO's environmental resource management activities included the development of community water infrastructure and water source protection. Protecting water sources from contamination is complementary to both the agricultural and health activities under USAID-ACCESO and included the promotion of better water management practices, better water quality monitoring, improved pesticide and fertilizer use, and decreased soil erosion through planting live barriers.

As can be seen from this brief description of the initiatives under USAID-ACCESO, there were many interlinked activities and hence many intermediate and final indicators that needed to be monitored to gauge the impact and effectiveness of the program. In what follows, we present the results of the three rounds of surveys conducted between June-July 2012 and May-September 2015 to measure the key indicators and outcomes in the FTF results framework and to evaluate the impact of the intervention on the economic wellbeing of the households in the zone of influence.

2 IMPACT EVALUATION METHODOLOGY

The benefits of USAID-ACCESO were measured using a rigorous, quasi-experimental impact evaluation methodology. An impact evaluation is a study that measures the changes in outcomes which can be attributed to a specific intervention. Impact evaluations require a credible and rigorously defined counterfactual which estimates what would have happened to beneficiaries had the project not taken place. The benefits of USAID-ACCESO were measured with a rigorous, non-experimental design that incorporates matching, a panel survey, difference in difference estimation, and econometric analysis.

Our main objective was to assess the impact of participation in USAID-ACCESO (T) on an outcome Y . Suppose that each household i would have an outcome equal to $Y_i(0)$ if that household did not participate in USAID-ACCESO ($T_i = 0$) or $Y_i(1)$ if it did ($T_i = 1$). Our parameter of interest would be the average impact of having participated, or the treatment:

$$\tau = E[Y_i(1) - Y_i(0)].$$

However, there are several problems with this basic framework. First, we are only able to observe one realization of the potential outcome for each household. In reality, households either participate in USAID-ACCESO (so we observe $Y_i(1)$) or do not (and we observe $Y_i(0)$ instead). Overall, our data is based on the observed data, $Y_i = T_i \cdot Y_i(1) + (1 - T_i) \cdot Y_i(0)$. Thus, it is convenient to express outcome Y in a regression framework:

$$Y_i = Y_i(0) + [Y_i(1) - Y_i(0)]T_i$$
$$Y_i = \beta_0 + \tau T_i \quad (1)$$

where:

Y_i is the observed outcome for household i and T_i is participation of household i in ACCESO, is equal to 1 for those that participate, and is equal to zero for those that do not participate.

We could use this framework to compare the outcomes of households that benefited from USAID-ACCESO and those that did not in order to find the impact of the program. However, this comparison assumes that the only differences between the two groups is their participation in the program. Had the treatment been randomly assigned, there would be no reason to believe there would be any systematic differences between beneficiaries and non-beneficiaries [(Duflo, Glennerster, & Kremer, 2008)]. Under randomization, the average outcome experienced by the control group provides a good counterfactual of what would have happened to beneficiaries in the absence of the program.

USAID-ACCESO, however, was not randomly assigned. Therefore, a second problem arises in our case because households that participated in USAID-ACCESO might be different than those that did not. For example, USAID-ACCESO's operations took place more in areas closer to paved roads (which made the transportation of extension agents easier) than in isolated areas. Differences between participants and non-participants create an additional problem in identifying the causal impact of USAID-ACCESO; it might be the case that beneficiaries experience better outcomes – such as higher income or productivity – merely because they were better off to begin with and not necessarily because the program had a positive effect.

There are different ways to address these evaluation problems. One is to use a regression framework to control for pre-treatment differences between participants and non-participants that may confound the effect of the program. In these lines, we would include control variables, X_i , in Equation (1) to estimate:

$$Y_i = \beta_0 + \tau T_i + \theta X_i + \varepsilon_i \quad (2)$$

Estimation of the τ parameter of interest relies on the “selection on observables” [(Heckman & Robb, 1984)] assumption: the potential outcomes are independent of the treatment status after controlling for a set of confounding variables X_i , i.e. $T_i \perp \{Y_i(0), Y_i(1)\} | X_i$. In other words, this assumes that controlling for the set of variables X_i eliminates the selection bias, such that the treatment would be assigned “as if random,” and that the outcomes of those in the control group would provide a valid counterfactual.

The dimension of X_i might be large and difficult to appropriately control for because of non-linearities. Because of this, subsequent research has been devoted to developing methods that avoid adjusting directly for all variables in X_i and instead focus on simpler adjustments. For example, Rosembaum and Rubin (1983) develop the propensity score methodology; their insight is that, if selection on observables holds when we account for X_i , then it still holds when we condition the outcomes on a scalar function of X_i . Denote $e(X_i)$ as the conditional probability of treatment assignment: the propensity score. Rosembaun and Rubin show that if the treatment is independent of the potential outcomes conditional on (X_i) , then it is also independent conditional on $e(X_i)$:

$$T_i \perp \{Y_i(0), Y_i(1)\} | X_i \rightarrow T_i \perp \{Y_i(0), Y_i(1)\} | e(X_i)$$

The econometric literature suggests different methods to condition on the propensity score. For example, (Rosenbaum & Rubin, 1983) propose dividing the sample into M blocks with fixed upper and lower limits and approximately equal probability of treatment. Their method proposes estimating the

program effect in each block and calculating the total effect as an average of the impact in each block. Another method is to control for $e(X_i)$ in a regression analysis; however, it is not clear why this would be better than directly controlling for X_i instead. An alternative option is to match observations with similar levels of propensity scores, as suggested by (Dehejia & Wahba, 1994), (Heckman, Ichimura, & Todd, Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program, 1997), and (Heckman, Ichimura, & Todd, Matching as an Econometric Evaluation Estimator, 1998). For each observation in the treatment group, we can find one or more observations in the comparison (non-treated) group with a very similar ex-ante probability of assignment. Finding a good match for each observation in the treatment group can be done in several ways: one-to-one (i.e. finding the observation in the control group with the closest propensity score), nearest k neighbors (i.e. finding the k control observations with the closest propensity score), radial (i.e. averaging all observations in the control group whose propensity score lies within an arbitrary radius r compared to that of the treatment), or kernel (i.e. constructing an weighted average using a kernel function based on the differences between the treatment's and each of the comparisons' propensity score).

More recently, (Hirano, Imbens, & Ridder, Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score, 2003) suggest an adjustment by weighting observations by the inverse of a non-parametric or flexible estimate of the propensity score. They show that this method achieves the semi-parametric efficiency bound. In general, they propose to weigh the treated group by $1/\hat{e}(x_i)$ (one over the probability of being treated conditional on X_i) and the comparison group by $1/(1 - \hat{e}(x_i))$ (one over the probability of not being treated conditional on X_i). Intuitively, this procedure balances the selection problem by placing larger weights on observations in the treatment group that had lower chances of program participation and on observations in the control group that had higher chances of participation. In particular, the authors suggest the following weighted estimator:

$$\hat{\tau} = \frac{\sum_{i=1}^N \frac{T_i y_i}{\hat{e}(X_i)}}{\sum_{i=1}^N \frac{T_i}{\hat{e}(X_i)}} - \frac{\sum_{i=1}^N \frac{(1 - T_i) y_i}{1 - \hat{e}(X_i)}}{\sum_{i=1}^N \frac{(1 - T_i)}{1 - \hat{e}(X_i)}} \quad (3)$$

where $\hat{e}(x_i)$ is an estimate of the propensity score or participation probability. The propensity score can be estimated by a logit model of the binary indicator of the treatment (T_i) on a set of pre-treatment covariates that capture how likely it is for a certain community or household to receive the treatment. In particular, we can estimate a logistic model for the probability to participate conditional of observables variables, X_i :

$$\Pr(T_i = 1|X_i) = \frac{\exp(X_i\alpha)}{1 + \exp(X_i\alpha)} \quad (4)$$

and calculate the propensity score for each household with the estimated parameters $\hat{\alpha}$ ².

In this impact evaluation, we combine two approaches proposed by (Hirano, Imbens, & Ridder, Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score, 2003) and (Hirano & Imbens, Estimation of Causal Effects using Propensity Score Weighting: An Application to Data on Right Heart Catheterization, 2002): we directly control for observable covariates X_i that might affect the decision to participate and weight our estimator by the inverse of the propensity score. (Robins & Rotnitzky, 1995) argue that this class of estimators is doubly robust because, as long as one of the models (either that conditioning on the observed $\{Y_i(0), Y_i(1)\}$ on X_i or that estimating the propensity score $\hat{e}(X_i)$) is correctly specified, the estimator will be consistent. We estimate Equation (1) controlling for potential confounding variables X_i and also weigh observations with the following slight variant of (3):

$$\lambda_i = T_i + (1 - T_i) \left[\frac{\hat{e}(X_i)}{1 - \hat{e}(X_i)} \right] \quad (5)$$

where observations in the treatment group receive a weight of 1, while those in the control group are weighted by $\frac{\hat{e}(X_i)}{1 - \hat{e}(X_i)}$, where $\hat{e}(X_i)$ is the propensity score estimated through Equation (4).

This framework ensures that program recipients are compared to non-recipients who are similar in terms of observable household characteristics. However, even after controlling for such observable differences, there may still be differences in recipients' and non-recipients' unobservable characteristics that could also bias the results (called 'selection on unobservables' in the literature; (Heckman, Ichimura, Smith, & Todd, 1998)). Therefore, we complement our analysis framework by considering household fixed effects in the estimation of the outcomes of interest. The idea is rather intuitive. If there are any time-invariant unobservables (e.g. ability, entrepreneurship, land quality, etc.) that have a linear additive effect, we can eliminate them by considering them as an additive and time-invariant effect in households' outcomes and the overall mean of the outcome or in general by including

² Note that the estimation of the propensity score is done on the "common support region." That is, the range of the estimated propensity score that is common between the treatment group and the comparison group. Households that are outside that range are excluded, as they have no similar counterpart in the other group (i.e. their X_i 's are too different).

household-level indicators in the regression. We call this estimator the household fixed effect difference in difference estimator (DID):

$$Y_i = \alpha_i + \psi_t + \tau \cdot T_i \cdot \psi_t + \varepsilon_i \quad (6)$$

where

α_i are the household time-invariant (fixed) effects,

ψ_t are common aggregate shocks or time effects, that affect both treatment and comparison groups,

T_i is participation of household i in USAID-ACCESO, is equal to 1 for those that participate, and is equal to zero for those who do not participate, as before, and

ε_i is a random error.

The impact evaluation is based on the following estimation procedure. First, we estimate the likelihood of receiving the treatment as a function of X_i : $\hat{e}(X_i)$ as defined in equation (4). X_i includes the variables that we need to account for to eliminate any selection on observables in the treatment assignment following the criteria delineated by USAID/Honduras and our discussions with USAID-ACCESO implementers for program allocation. Second, we include household fixed effects, α_i , in the households' outcome regression to eliminate any time-invariant unobservables that are not included in X_i . Third, we include year indicators to account for aggregate conditions in the ZOI that affect both treatment and comparison groups. Finally, we estimate a regression of Y_i on $T_i \cdot \psi_t$, the treatment assignment indicator, interacted with the post-baseline indicators, to estimate the impact of USAID-ACCESO (τ) on the outcome, Y_i . Our doubly robust estimator weighs each observation with the propensity score $\hat{e}(X_i)$ to adjust the means of the comparison group to matching that was done using baseline data³.

3 GEOGRAPHY, TREATMENT ASSIGNMENT, AND ATTRITION

To evaluate USAID-ACCESO, we designed a survey that sampled 3,326 households in the activity's area of influence across 207 villages within six departments of Honduras: Copán, Intibucá, La Paz, Lempira,

³ The propensity scores are determined at the household level, so the weights are at the household level. When the data is at a lower level, we scale the weights by the frequency of the household in the data. For example, if a household has three crops in the agricultural production survey module, the weights are divided by three. In addition, since at this level we do not have a balanced panel (households can change crops or households member can migrate), we are interested in the average change in the between outcomes in treatment groups vis-à-vis the change in the comparison group.

Ocotepeque, and Santa Bárbara. The baseline survey included 1,256 households in the treatment group, which were randomly selected from the roster of about 12,000 beneficiaries provided by the implementer, Fintrac, in April 2012. The comparison group was comprised of 2,005 households randomly selected from the 2001 Honduran Population Census. The comparison group was selected to preserve the representativeness of living conditions in each of the six departments in which USAID-ACCESO operated.

Our data results highlight an additional difficulty in determining the actual beneficiaries of USAID-ACCESO: as suggested by Table 1, 51% of households in the beneficiary roster do not self-report that they participated in the program in either the baseline or the two follow-up rounds of the survey. A plausible explanation for this is that USAID-ACCESO's assistance might not be sufficiently salient and thus households may not effectively recall access to the program even when they might have actually benefited from it. Throughout this report, we define households in the treatment group as those who: (a) were in the beneficiary roster provided by Fintrac in 2012 or (b) self-reported as participating in the USAID-ACCESO program in 2013 or 2015 survey. In Table 2, we identify the households in the treatment group with an updated list of beneficiaries at endline in order to gauge if households that did not report participating might have dropped out or become inactive. The table shows that the majority of treatment households located in the implementer's beneficiary list are active and that the probability of attrition is higher for those that were discarded or inactive in the updated list. That is, treatment households in the sample that are classified as inactive or discarded by the implementer are more likely to not respond to the survey; however, this attrition probability is much lower than the overall attrition for the survey. In addition, households classified as active by the implementer are more likely to report participating in USAID-ACCESO, as one would expect.

Table 1 Beneficiary Attribution

	Comparison	Treatment	Total
Does not self-report USAID-ACCESO	2005	674	2679
Self-reports USAID-ACCESO	0	647	647
Total	2005	1321	3326

Table 2 Composition of the Treatment Group by Status in Implementer's Beneficiary List

	Mean	SD	N of Households	Observations
Discarded Households				
Attrition Probability	5%	22%	1	21

Mentions USAID-ACCESO in any survey	33%	48%	7	21
Inactive Households				
Attrition Probability	3%	16%	1	39
Mentions USAID-ACCESO in any survey	23%	43%	9	39
Active Households				
Attrition Probability	2%	13%	14	869
Mentions USAID-ACCESO in any survey	46%	50%	397	869
Health and Nutrition Households				
Attrition Probability	0%	0%	0	19
Mentions USAID-ACCESO in any survey	16%	37%	3	19
Small-Medium Enterprise				
Attrition Probability	0%	0%	0	2
Mentions USAID-ACCESO in any survey	100%	0%	2	2
Total				
Attrition Probability	2%	13%	16	950
Mentions USAID-ACCESO in any survey	44%	50%	418	950

3.1 BALANCING WEIGHTS

As discussed in Section 2, the evaluation problem requires us to construct a counterfactual for what would have happened to the treatment group in the absence of the intervention. For this purpose, we need to estimate the probability of treatment (Equation 4) based on Fintrac's targeting criteria and to balance the characteristics so that the comparison group looks similar to the treatment group for all observable characteristics. We determine the probability of participation in USAID-ACCESO as a function of the following variables at baseline (X_i):

1. Time to the main and secondary roads, estimated through a GIS-accessibility model;
2. Size of the nearest market, measured through population in 2012;
3. Time and cost of transportation to the nearest market, estimated through a GIS-accessibility model;
4. Household characteristics: size; number of dependent members (younger than 14 or older than 65 years of age); number of working-aged members; number of reproductive-aged women (10-49 years old); and age, literacy, and gender of the head of household;

5. Living conditions at the village level: poverty rates in 2001 (from the Small Area Estimates Poverty and Inequality Database) and at baseline; children's nutrition status (percentage of children that are anemic, stunted, wasted, and underweight); and women's nutrition status (percentage of women that are stunted and underweight);
6. Households' crop composition: planted area of corn, beans, and coffee; and
7. Location: Indicator variables by department.

These should serve as good primary predictors of which households or communities are likely to be targeted. First note that these variables were collected either at baseline or in a previous year and, therefore, should not be affected by the treatment. Second, these variables correspond to the criteria set forth by USAID-Honduras and reflect the limitations that the implementers might have had when selecting the program beneficiaries. Table 3 presents the estimates for Equation 4 as a function of the variables described above. The first column presents the odds ratio of the indicated coefficient, while the second presents the standard error of the estimate. The last column shows the p-value, or whether the variable significantly explains the participation decision.

Table 3 Logit Model to Estimate the Probability of Treatment

	Coefficient	S.E.	p-value	
Market Access				
Cost to Market (Dollar per Kg.)	-4.244	1.088	0	***
Time to Market (Hrs.)	0.039	0.014	0.004	***
Distance to market (Kms.)	-0.001	0.002	0.678	
Time to Principal Road (Hrs.)	-0.006	0.007	0.353	
Time to Secondary Road (Hrs.)	0.019	0.011	0.071	*
2012 Population in market (Ths.)	0.002	0.001	0.217	
Household characteristics				
Household Size	0.03	0.007	0	***
Age of Head of HH	-0.002	0.001	0.001	***
No. women 10 to 49	0.005	0.01	0.621	
Head of HH is male	0.156	0.023	0	***
No. of dependent members	-0.01	0.008	0.205	
No. of productive members	0.08	0.019	0	***
Literacy of Head of HH	-0.002	0.001	0.001	***
HH's Crop Composition				
Corn Area Planted	0.107	0.017	0	***
Beans Area Planted	0.07	0.04	0.08	*
Coffee Area Planted	0.06	0.009	0	***
Village-level variables				
Poverty Rate 2001	0.473	0.156	0.002	***
Poverty Rate at baseline	0.033	0.04	0.415	
% Children 6-59 months with Anemia	0.189	0.044	0	***
% Children 0-59 months Stunted	0.058	0.129	0.655	
% Children 0-59 months Wasted	0.021	0.069	0.758	
%Children 0-59 months underweight	0.114	0.048	0.018	**
% Women 10-49 with anemia	0.002	0.075	0.983	
% Women 10-49 underweight	-0.224	0.107	0.036	**
Department Indicator Variables	YES			
Observations	3326			
Pseudo R-squared	0.109			

Significance level: * p<0.1, ** p<0.05, *** p<0.01

While we present the marginal effects of our coefficients to make it easier to discern the relationship between the variables of interest, we use the raw coefficients to estimate the probability of treatment $\hat{e}(X_i)$ and to construct the weights λ_i in Equation (5). The coefficients presented give us the change in probability of treatment given a one unit increase in the indicator of interest. For example, all else being equal, an additional productive member of the household increases the probability that the household is treated by 0.08 or 8 percentage points. Note that this estimation only looks to determine which variables and to what extent those variables explain the likelihood of participating in USAID-ACCESO. Table 4 highlights the effectiveness of our estimation strategy to construct a balanced comparison sample with similar characteristics to those in the treatment group. We compare the means for the set of observable baseline characteristics (X_i variables) between the beneficiaries and the comparison

group with and without weighing the observations by λ_i , the propensity score we constructed from the coefficients presented in Table 3. Column 1 shows the unweighted mean for the treatment group, column 2 shows the unweighted control mean, and column 3 shows the coefficient from the t-test of the difference between the two. Column 4 shows the p value, or whether the difference is significant or not. Columns 5-8 show the same statistics, but with the values adjusted based on the propensity score.

The first four columns of Table 4 show that there are considerable differences between the treated and comparison households before weighting, highlighting the need to rebalance the composition of both groups through propensity score methods. Because USAID-ACCESO was targeted toward extremely poor communities, those who benefited from the program lived in villages with larger shares of stunted, wasted, and underweight children and underweight women. Also, because the program provided agricultural extension services, those in the treatment group had larger planted areas of corn, beans, and coffee. Once we weight the observations by λ_i , as defined in equation (5), we see in columns 5-8 that the weighted differences between the treatment and comparison groups disappear for 23 out of 24 matching variables (the only exception being the acreage of corn planted, which remains significant at the 10% confidence level) compared to the unweighted differences in columns 1-4.

Table 4 Unweighted and Weighted Matching Variables

	Unweighted					Weighted			
	(1) Treat	(2) Control	(3) t-test	(4) p-value		(5) Treat	(6) Control	(7) t-test	(8) p-value
Cost to Market (Dollar per Kg.)	0.02	0.02	-0.39	0.70		0.02	0.02	0.18	0.85
Time to Market (Hrs.)	2.95	2.92	0.13	0.90		2.95	2.90	0.22	0.83
Distance to market (Kms.)	10.55	10.61	-0.11	0.91		10.55	10.38	0.34	0.74
Time to Principal Road (Hrs.)	2.67	2.62	0.27	0.79		2.67	2.65	0.11	0.91
Time to Secondary Road (Hrs.)	0.69	0.61	1.72	0.09	*	0.69	0.71	-0.27	0.79
2012 Pop. In Market (Thou)	13.62	13.46	0.32	0.75		13.62	13.42	0.49	0.63
Household Size	5.62	4.75	9.11	0.00	***	5.62	5.75	-0.84	0.40
Age of Head of HH	44.52	46.73	-4.03	0.00	***	44.52	45.65	-1.35	0.18
No. women 10 to 49	1.76	1.46	6.35	0.00	***	1.76	1.77	-0.22	0.82
Head of HH is male			10.9		***				
	0.90	0.75	9	0.00		0.90	0.90	-0.04	0.97
No. of dependent members	2.50	2.10	5.86	0.00	***	2.50	2.46	0.37	0.71
No. of productive members	3.13	2.65	8.08	0.00	***	3.13	3.29	-1.08	0.28
Literacy of Head of HH	0.77	0.67	5.74	0.00	***	0.77	0.73	1.07	0.29
Poverty rate at baseline	0.37	0.34	3.02	0.00	***	0.37	0.37	0.13	0.90
Poverty Rate 2001	0.40	0.39	1.83	0.07	*	0.40	0.39	0.57	0.57
% Child 6-59 months Anemic	0.24	0.24	-0.20	0.84		0.24	0.25	-0.73	0.46
% Child 0-59 months Stunted	0.43	0.38	3.80	0.00	***	0.43	0.42	0.42	0.68
% Child 0-59 months Wasted	0.03	0.03	0.55	0.58		0.03	0.03	-0.15	0.88
% Child 0-59 months Underweight					**				
	0.16	0.14	2.18	0.03		0.16	0.16	-0.02	0.99
% Women 10-49 with Anemia	0.11	0.12	-0.74	0.46		0.11	0.10	0.72	0.47
% Women 10-49 Underweight	0.07	0.08	-1.09	0.28		0.07	0.07	0.12	0.91
Corn Area Planted	0.44	0.26	6.29	0.00	***	0.44	0.62	-1.85	0.07
Beans Area Planted	0.11	0.06	5.26	0.00	***	0.11	0.14	-0.96	0.34
Coffee Area Planted	0.52	0.28	4.44	0.00	***	0.52	0.97	-1.14	0.25
Observations	3326								

Significance level: * p<0.1, ** p<0.05, *** p<0.01

Table 4 tests the differences in X_i between the treatment and comparison groups, weighing the observations by λ_i (which is a function of X_i). A stronger test for the weighting properties of our estimator is to test whether it is able to balance other variables that are not included in X_i . Table 5 tests differences between the treatment and comparison groups for some additional variables which were not used to create the propensity score: the number of children under five years old, the ethnicity of the head of household, whether the household has both a female and a male adult (or if it only has one of them), dependency ratio, housing expenses (imputed rent and utilities), food expenditures, and total annual per capita expenditures. As in Table 4, the first four columns present the treatment mean, the control mean, the statistic from the t-test of difference in means, and the p value, respectively, for the unweighted sample, while columns 5-8 do the same for the weighted sample. Initially, the means of the

treatment and control group observed in columns 1 and 2 were very different. Reassuringly, we see in columns 5-8 that our propensity score is also able to balance most of these variables.

Table 5 Unweighted and Weighted Comparison of Additional Variables at Baseline

	Unweighted					Weighted				
	(1) Treat	(2) Control	(3) t-test	(4) p-value		(5) Treat	(6) Control	(7) t-test	(8) p-value	
No. of Children under 5	0.70	0.63	2.68	0.01	***	0.70	0.76	-1.78	0.08	*
Head of HH is indigenous	2.60	2.80	-2.31	0.02	**	2.60	2.73	-1.09	0.28	
Female, No Male HH	0.04	0.14	-8.60	0.00	***	0.04	0.04	0.40	0.69	
Male, No Female HH	0.04	0.04	-0.29	0.77		0.04	0.03	1.15	0.25	
Male & Female HH	0.92	0.83	7.15	0.00	***	0.92	0.93	-1.06	0.29	
Dependency Ratio	0.96	1.03	-2.10	0.04	**	0.96	0.96	0.11	0.92	
Housing Expenses	2111.67	2398.37	-2.38	0.02	**	2111.67	2039.50	0.52	0.60	
Utilities per capita	1164.46	1520.75	-8.18	0.00	***	1164.46	1189.45	-0.59	0.56	
Housing Expenditures	3276.13	3919.12	-4.32	0.00	***	3276.13	3228.95	0.29	0.77	
Food Expenditures	4800.12	4939.22	-1.04	0.30		4800.12	4594.05	1.50	0.14	
Annual Exp PC USD (2011\$)	2.08	2.24	-2.39	0.02	**	2.08	1.97	1.78	0.08	*
Observations	3226									

Significance level: * p<0.1, ** p<0.05, *** p<0.01

3.2 ATTRITION WEIGHTS

While the propensity score and the weights were determined using variables at baseline, our analysis is further complicated because of attrition problems: 449 (13%) households surveyed at baseline could not be re-surveyed in the first follow-up. Some of these households had moved, were temporarily absent during the survey period, or refused to participate. However, the most likely problem was the composition of the team of enumerators. Due to logistical limitations, the survey was scheduled in two sub-periods. The first sub-period comprised surveys in Intibucá, La Paz, and Lempira. Local enumerators were hired in these departments to conduct the fieldwork, and the same teams of enumerators also collected the surveys in Copán, Ocotepeque, and Santa Bárbara during the second sub-period. Arguably, because they were not locals in the departments of the second sub-period, enumerators might have faced more difficulties in locating households in the roster. For example, the rate of attrition in Copán and Ocotepeque during the first follow-up survey were 21% and 19%, respectively. In our second follow-up survey, we were able to recover 313 of the households that were not surveyed during the first follow-up. Ultimately, we are able to observe 96% of our initial baseline sample in at least two rounds of surveying, as detailed in Table 6. The first column presents the number of attritive households in the

department, while the second column shows the total number of sampled households in the department. The third column shows the rate of attrition.

Table 6 Sample Attrition: Percent of Households Surveyed Only Once

		N Attriters	N Households	% Attrition
Copán		39	557	7%
	Treatment	5	207	2%
	Control	34	350	10%
Intibucá		29	540	5%
	Treatment	12	246	5%
	Control	17	294	6%
La Paz		3	556	1%
	Treatment	1	218	0%
	Control	2	338	1%
Lempira		22	563	4%
	Treatment	10	217	5%
	Control	12	346	3%
Ocotepeque		32	544	6%
	Treatment	8	214	4%
	Control	24	330	7%
Santa Bárbara		11	566	2%
	Treatment	0	219	0%
	Control	11	347	3%
Total		136	3326	4%
	Treatment	36	1321	3%
	Control	100	2005	5%

As laid out in Section 2, the dependent variables in most of our analysis will be the change that households experience before and after the program for a set of outcomes of interest (ΔY_i). Therefore, we can only estimate the impact of USAID-ACCESO on households that were interviewed both at the baseline and during at least one of the follow-up surveys. While we have shown that our methodology balances differences between the treatment and comparison groups for a broad range of variables at baseline, our results can still be affected by attrition if households that could not be re-surveyed in the follow-up are systematically different from those that remain in the sample. Thus, the final dataset for our estimations can be based on a population that does not necessarily correspond to the original design. Moreover, the estimates can be affected if those households that leave the sample are different within the treatment and the comparison groups. For example, if households in the treatment group that leave the sample are richer or more educated than those that leave the sample in the comparison group, such differential attrition can impact our estimates. Table 7 and Table 8 show the differences

between attrition and non-attrition households within the treatment and comparison groups for different dimensions: household characteristics and health status of women and children. We calculate these differences through a regression for each characteristic of interest on a dummy variable that takes on the value of zero if the household is present in at least two surveys and one if the household leaves the survey. This allows us to account for intra-cluster correlation in our sample at the village level. The first column in both Table 7 and Table 8 shows the difference in the means between attriters and non-attriters, which is captured by the coefficient on the dummy variable for attrition. The second column shows the standard error of this coefficient. Column three shows the mean for non-attriters, captured by the constant in our regression. The fourth column shows the number of observations. In each table, we first run these regressions for the entire sample (columns 1-4), then for the treatment group (columns 5-8), and then the control group (columns 9-12). While there are some differences in a few variables, these do not suggest a systematic pattern of differential attrition⁴.

⁴ **“Error! Reference source not found.”** shows the differences and means between attrition and non-attrition households within the treatment and comparison groups for different dimensions: household characteristics, health status of women and children, and agricultural variables.

Table 7 Household Characteristics of Attriters

	Total					Treatment					Comparison				
	Difference	SE	Mean of Non-Attriters	Obs		Difference	SE	Mean of Non-Attriters	Obs		Difference	SE	Mean of Non-Attriters	Obs	
Age	-8.795	46.221	1.242	3326	***	-4.640	44.668	1.930	1321	**	-10.779	47.269	1.504	2005	***
Male head of household	0.031	0.807	0.032	3326		0.015	0.901	0.047	1321		0.066	0.744	0.039	2005	*
No. children under 5	0.163	0.654	0.084	3326	*	-0.093	0.704	0.122	1321		0.271	0.619	0.093	2005	***
No. women 10 to 49	-0.091	1.583	0.093	3326		0.129	1.760	0.208	1321		-0.115	1.465	0.095	2005	
Dependency ratio	-0.079	1.004	0.093	3326		-0.152	0.964	0.115	1321		-0.064	1.030	0.114	2005	
No. of productive members	-0.362	2.855	0.103	3326	***	-0.078	3.134	0.223	1321		-0.376	2.666	0.111	2005	***
No. of dependent members	-0.253	2.267	0.161	3326		-0.223	2.501	0.290	1321		-0.190	2.110	0.158	2005	
Female only, No male	0.005	0.098	0.027	3326		0.014	0.041	0.038	1321		-0.016	0.136	0.032	2005	
Male only, no Female	0.046	0.035	0.024	3326	*	0.021	0.035	0.040	1321		0.055	0.035	0.028	2005	*
Male and Female	-0.051	0.867	0.038	3326		-0.035	0.924	0.055	1321		-0.039	0.829	0.044	2005	
Head of HH can read and write	-0.042	0.711	0.042	3326		-0.159	0.770	0.088	1321	*	0.019	0.671	0.048	2005	
Primary schooling or more	-0.027	0.724	0.039	3313		-0.113	0.779	0.082	1314		0.021	0.686	0.047	1999	
Secondary schooling or more	0.017	0.072	0.027	3313		0.032	0.079	0.053	1314		0.014	0.067	0.031	1999	
Ethnic Minority	-0.008	0.582	0.041	3326		0.004	0.607	0.087	1321		-0.004	0.564	0.047	2005	
Years living in the house	-7.042	20.375	1.645	2840	***	-2.484	19.640	2.580	1164		-9.469	20.891	2.055	1676	***
HH owns the House	-0.162	0.882	0.037	3326	***	0.010	0.907	0.047	1321		-0.216	0.866	0.047	2005	***
Expenditure PC in 2010 dollars	0.033	2.172	0.145	3326		-0.066	2.082	0.300	1321		0.039	2.233	0.136	2005	
Living on less than 1.25 USD	-0.029	0.352	0.047	3326		0.023	0.366	0.087	1321		-0.043	0.343	0.048	2005	
Moderate or severe hunger	0.023	0.033	0.020	3139		-0.030	0.030	0.007	1255	***	0.042	0.035	0.027	1884	

The Difference column shows the estimated difference between the attriters and non-attriters.

* p<0.10, ** p<0.05, *** p<0.01

Table 8 Child and Maternal Health Characteristics of Attriters

	(1)	(2)	(3)	(4)	(5)		(1)	(2)	(3)	(4)	(5)		(1)	(2)	(3)	(4)
Child Anthropometry	Difference	SE	Mean of Non Attriters	Obs	p-value		Difference	SE	Mean of Non Attriters	Obs	p-value		Difference	SE	Mean of Non Attriters	Obs
Age in Months	-0.225	2.135	0.127	2196	0.079	*	-0.110	2.156	0.277	927	0.691		-0.242	2.119	0.157	1269
Male	0.139	0.501	0.048	2196	0.004	***	0.138	0.498	0.074	927	0.065	*	0.138	0.503	0.051	1269
Has anthropometry info	-0.129	0.687	0.052	2196	0.013	**	-0.320	0.684	0.122	927	0.009	***	-0.083	0.690	0.053	1269
Child anemia (6-59 months)	0.047	0.236	0.062	1358	0.449		0.047	0.239	0.157	581	0.765		0.049	0.234	0.074	777
Child is Stunted	0.059	0.425	0.074	1489	0.426		0.281	0.469	0.156	626	0.074	*	0.054	0.391	0.069	863
Child is wasted	0.019	0.030	0.036	1443	0.607		-0.023	0.023	0.006	609	0.000	***	0.021	0.036	0.042	834
Child is underweight	0.048	0.165	0.063	1445	0.444		0.459	0.166	0.236	610	0.054	*	-0.013	0.164	0.048	835
Women Anthropometry																
Age	-0.923	24.652	0.541	5254	0.089	*	-2.444	24.297	0.811	2329	0.003	***	-0.266	24.940	0.672	2925
Has anthropometry info	-0.084	0.424	0.032	5254	0.009	***	-0.152	0.402	0.057	2329	0.009	***	-0.056	0.442	0.041	2925
Are you pregnant?	-0.032	1.949	0.027	3759	0.230		0.048	1.952	0.005	1592	0.000	***	-0.060	1.947	0.037	2167
Woman's anemia status	0.034	0.109	0.035	2675	0.329		-0.075	0.111	0.035	1150	0.035	**	0.078	0.108	0.042	1525
Weight (Kgs.)	-1.786	54.501	1.411	2798	0.207		-2.771	54.081	2.854	1190	0.333		-1.546	54.818	1.423	1608
Height (cm)	-0.887	151.020	1.377	2796	0.520		-3.612	150.154	4.007	1189	0.369		-0.096	151.675	0.887	1607
BMI	-0.432	23.620	0.410	2748	0.292		-1.563	23.723	0.673	1171	0.021	**	0.051	23.542	0.427	1577
BMI < 18.5	-0.020	0.076	0.024	2576	0.410		0.038	0.069	0.051	1104	0.452		-0.049	0.081	0.024	1472

Significance level: * p<0.1, ** p<0.05, *** p<0.01

Even when there is no systematic pattern suggesting that those households that remained in our estimation sample were either better or worse off, we construct attrition adjustment factors based on (Fitzgerald, Gottschalk, & Moffitt, 1998)'s approach. If attrition is determined by observables, these authors suggest that we can: (a) use statistical models to determine the probability of attrition based on observables, (b) predict the probability of attrition for each observation, and (c) weight the observations by the inverse of this probability to adjust for any changes in the estimation sample.

Similar to how we modeled the probability of treatment, we model the probability of full attrition as a function of Z_i variables that include household characteristics, village attachment, enumerators' subjective perceptions, and location variables. For example, households with younger heads of household are more likely to migrate. We also include variables to capture attachment to the village of origin (e.g. time of residency and home ownership) and the subjective perception of the enumerator about the household's participation in the follow-up (i.e. enumerators were asked at baseline whether they believed that the household would agree to participate in a future round of the survey). As previously explained, we hired local enumerators in half of the departments in our sample; this appears to be an important factor that determined households' participation in the follow-up survey. Therefore, we also included department indicator variables to capture these differences.

We estimate a probit model to determine the probability of attrition at endline based on baseline and midline characteristics:

$$P(Attrition) = \Phi(\hat{\theta}Z_i).$$

The marginal effects of the coefficients of this model are presented in

Table 9, with the standard errors below in brackets. These should not be interpreted as causal relationship between the variables, but rather as factors that are associated with lower or higher attrition in our sample. For example, while age of the household head does not *cause* attrition, a one-year increase in the age of the household head at baseline is associated with a 0.006 increase in the probability of attrition. Our estimated $\hat{\alpha}$ coefficients are used to construct predicted probabilities $\Phi(\hat{\theta}Z_i)$ and attrition weights $1/\Phi(\hat{\theta}Z_i)$. Finally, we construct sample adjustment weights that account for both imbalances in baseline characteristics and attrition probabilities, $\lambda_i/\Phi(\hat{\theta}Z_i)$.

In the same spirit as Table 4, Table 10 presents a balancing exercise for our estimation sample (3,190 households with completed surveys in at least two rounds⁵). In particular, we check whether our final attrition adjusted weights $\lambda_i/\Phi(\hat{\theta}Z_i)$ are able to adjust for differences in baseline characteristics between the treatment and comparison groups. Column one shows the unweighted mean of treatment households, while the second column shows the unweighted mean for control households. The third column shows the coefficient for the t-test of the treatment and control means, and the fourth column shows the p value of the t-test. Columns 5-8 repeat the same, but this time correct the values by the sample adjustment weights. The values in column 4 indicate that prior to including weights, there are significant differences between the treatment and control groups. The results in column 8 suggest that, after accounting for our final weights, observations in both groups are very similar in terms of a broad set of observable characteristics.

⁵ Note that the baseline sample is 3,326 that but some households are not found in the midline(2013) or the endline(2015), thus the difference in the sample size discussed.

Table 9 Probit for the Probability of being in both surveys: Correlates of attrition

	(1)
Age	0.006 [0.002]***
Male	-0.022 [0.054]
No. children under 5	-0.014 [0.031]
Household Size	0.008 [0.011]
Head of HH can read and write	0.04 [0.045]
Lived in house fewer than 10 years	-0.077 [0.051]
HH owns the House	0.132 [0.062]**
Enumerator believes will participate	0.014 [0.064]
2 or more revisits	0.015 [0.053]
Household Attrits in Midline	0 [.]
Observations	3326
Department Indicator Variables	YES

Standard errors in brackets

* p<0.10, ** p<0.05, *** p<0.01

Table 10 Unweighted and Weighted (for Treatment Assignment and Attrition) Characteristics at Baseline

	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
	Unweighted					Weighted			
	Treat	Control	t-test	p-value		Treat	Control	t-test	p-value
Cost to Market (Dollar per Kg.)	0.022	0.024	-0.511	0.610		0.022	0.022	0.078	0.938
Time to Market (Hrs.)	2.888	2.887	0.001	0.999		2.888	2.860	0.130	0.897
Distance to market (Kms.)	10.405	10.570	-0.292	0.770		10.405	10.307	0.193	0.847
Time to Principal Road (Hrs.)	2.610	2.570	0.200	0.842		2.610	2.593	0.093	0.926
Time to Secondary Road (Hrs.)	0.680	0.596	1.753	0.081		0.680	0.682	-0.032	0.974
2012 Pop. in Market (Thou)	13.650	13.459	0.413	0.680		13.650	13.433	0.540	0.589
Household Size	5.632	4.776	8.641	0.000	***	5.632	5.803	-1.138	0.256
Age of Head of HH	44.646	47.269	-4.685	0.000	***	44.646	46.230	-1.845	0.066
No. women 10 to 49	1.756	1.465	6.000	0.000	***	1.756	1.795	-0.576	0.565
Head of HH is male	0.901	0.744	11.075	0.000	***	0.901	0.900	0.080	0.936
No. of dependent members	2.501	2.110	5.557	0.000	***	2.501	2.473	0.275	0.783
No. of productive members	3.129	2.666	7.524	0.000	***	3.129	3.330	-1.303	0.194
Literacy of Head of HH	0.770	0.671	6.016	0.000	***	0.770	0.734	1.171	0.243
Poverty rate at baseline	0.370	0.334	2.998	0.003	***	0.370	0.368	0.214	0.831
Poverty Rate 2001	0.397	0.389	1.573	0.117		0.397	0.394	0.411	0.681
% Child 6-59 months Anemic	0.241	0.242	-0.135	0.893		0.241	0.251	-0.696	0.487
% Child 0-59 months Stunted	0.425	0.374	3.746	0.000	***	0.425	0.417	0.521	0.602
% Child 0-59 months Wasted	0.033	0.030	0.801	0.424		0.033	0.032	0.286	0.775
% Child 0-59 months Underweight	0.160	0.139	2.129	0.034	**	0.160	0.158	0.217	0.828
% Women 10-49 with Anemia	0.111	0.115	-0.494	0.621		0.111	0.102	1.304	0.193
% Women 10-49 Underweight	0.071	0.076	-1.181	0.239		0.071	0.071	0.008	0.994
Corn Area Planted	0.443	0.260	5.951	0.000	***	0.443	0.636	-1.920	0.056
Beans Area Planted	0.112	0.061	5.220	0.000	***	0.112	0.141	-0.991	0.323
Coffee Area Planted	0.516	0.281	4.252	0.000	***	0.516	0.995	-1.153	0.250
Observations	3326								

Significance level: * p<0.1, ** p<0.05, *** p<0.01

3.3 EX-POST POWER ANALYSIS AND SIMULATION

Given the changes in the sample each year, it is important to verify the kind of effects that we are able to detect with the panel sample obtained at endline, that is, after recuperating some of the households that were not surveyed in 2013 and losing some others in 2015. In “Appendix C – Sample Methodology,” we included the original sample design and the power analysis that was done before implementing the baseline survey. For the ex-post power analysis, instead of relying on the formulas and estimated means and variances from secondary data, we estimate the power of the sample using the empirical distribution at baseline and simulated data that follows the model/specification presented above.

The procedure consists of the following:

1. Use the underlying model $Y_i = \alpha_i + \lambda_t + \tau_t \cdot T_i \cdot \lambda_t + \varepsilon_i$ to generate random data with the observed sample size and the distribution of the outcome variable at baseline.
 - Our complete panel sample is comprised of 9,055 observations representing 3,190 households over three years (for which we have at least two observations). Households are grouped in 301 geographical clusters (villages).
2. Use the parameter values that express the difference the hypothesis test needs to detect.
 - Our treatment effects are given τ_t , the difference in difference term.
 - The generated data also needs to follow the correlation structure of the panel. Thus we include all the nuisance parameters: the additive household fixed effects (α_i), year fixed effects (λ_t), and the household level error term ε_i
3. Estimate the model with the generated data and repeat the procedure many times to obtain the empirical distribution of the tests statistic.

With this generated data, we can estimate the power for testing the null hypothesis that $\tau_t = 0$, against any specified alternative. In summary, we can generate data that has a real treatment effect of τ_t and then estimate the proportion of times the test does not reject the alternative hypothesis when it is the truth.

We explore the power of the design under two different indicators and a plausible range of treatment effects: continuous outcomes (per capita expenditure) and dichotomous outcomes (poverty prevalence). Figure 1 shows the estimated power to detect treatment effects in the range of 0 to 15 percent in total per capita expenditure for each year and the combined effect in the year after the

program. In the initial power analysis, our preferred sample detected a 12% difference between the treatment and control group with 80% power. In the figure, we can see that the power of the sample is better than expected; we can detect effects as low as a 7.5 % difference in expenditures with our sample.

Figure 2 and Figure 3 show the power analysis for binary outcomes such as poverty indicators. In this case, the prevalence of extreme poverty and relative poverty as established by Honduras are presented. This simulation is done under the assumptions that the prevalence of poverty follows the distribution observed at baseline and that the effect on poverty is driven from the treatment effects in expenditure simulated previously. The baseline power analysis calculated that the sample could detect a 7 percentage point difference in the prevalence of poverty with 80% power. From these figures, we can see that this calculation is very similar to the results of the simulation. We can detect changes in poverty of just under 7 percentage points; this change in poverty would imply a 15% increase in the annual income of treatment households.

In summary, the results from the simulation power analysis show that the sample design is well powered to detect economically significant differences in expenditures and poverty rates across the treatment and comparison groups in the sample.

Figure 1 Sample Power to Detect Changes in Expenditures

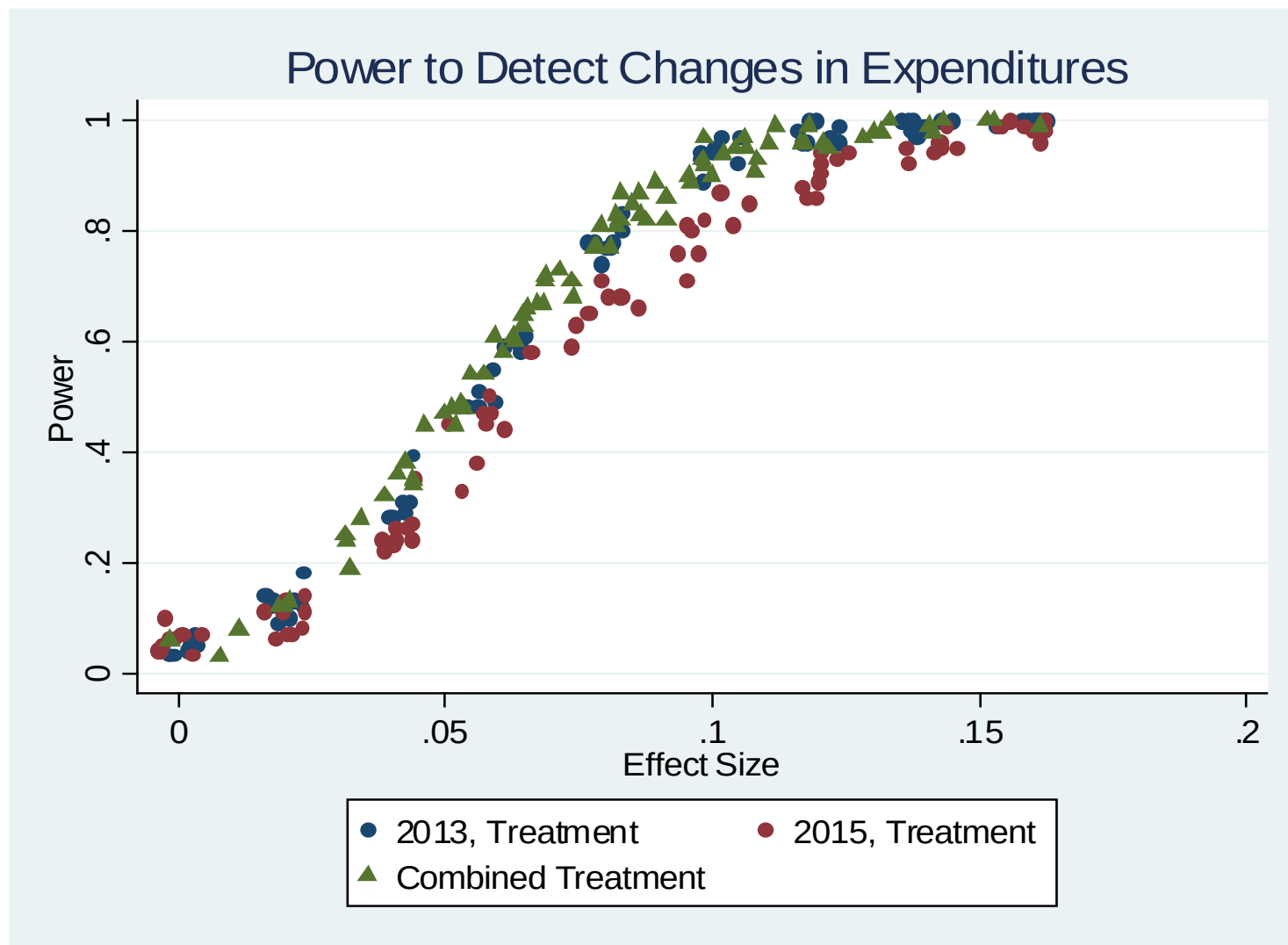


Figure 2 Sample Power to Detect Changes in Extreme Poverty

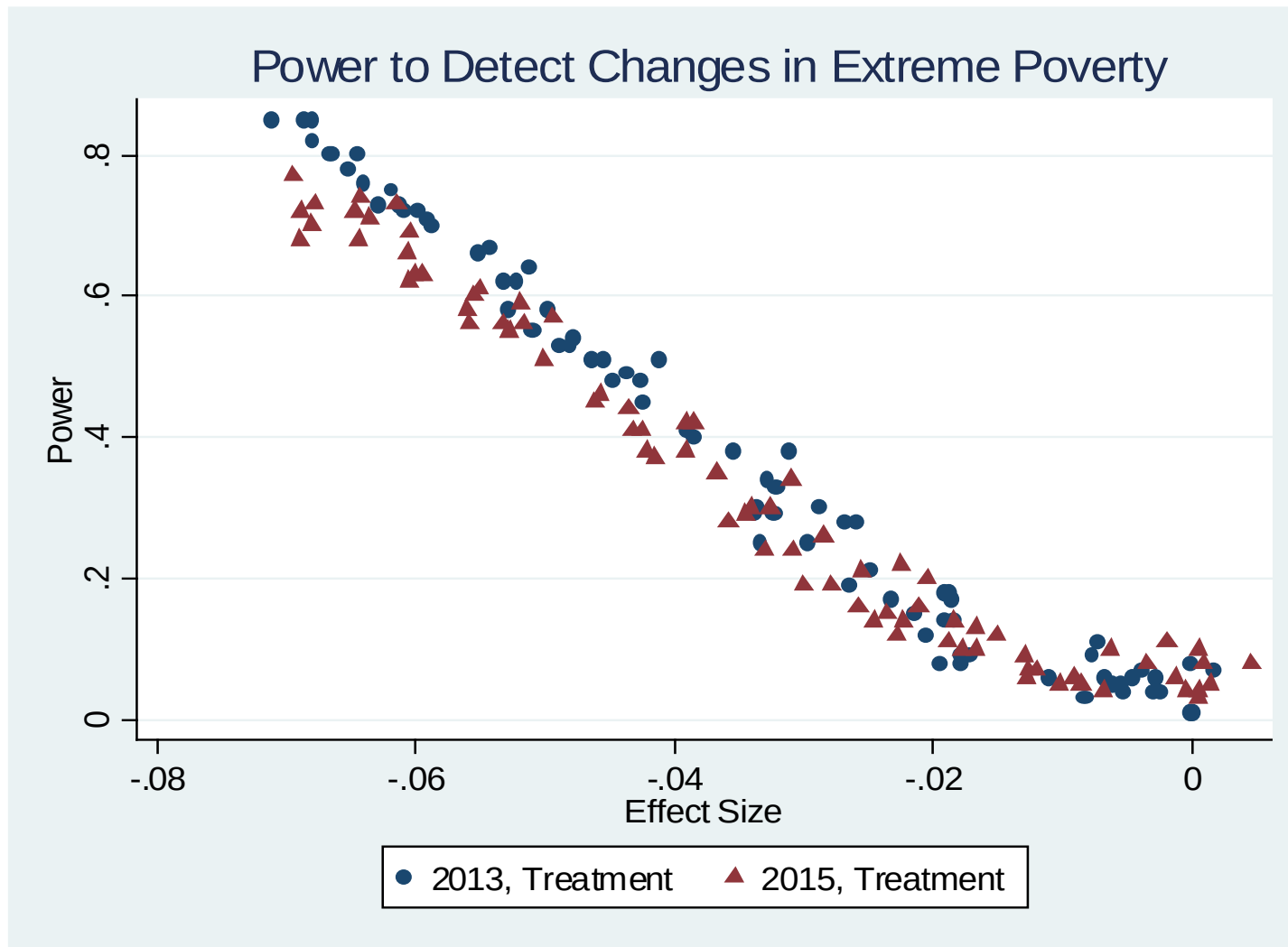
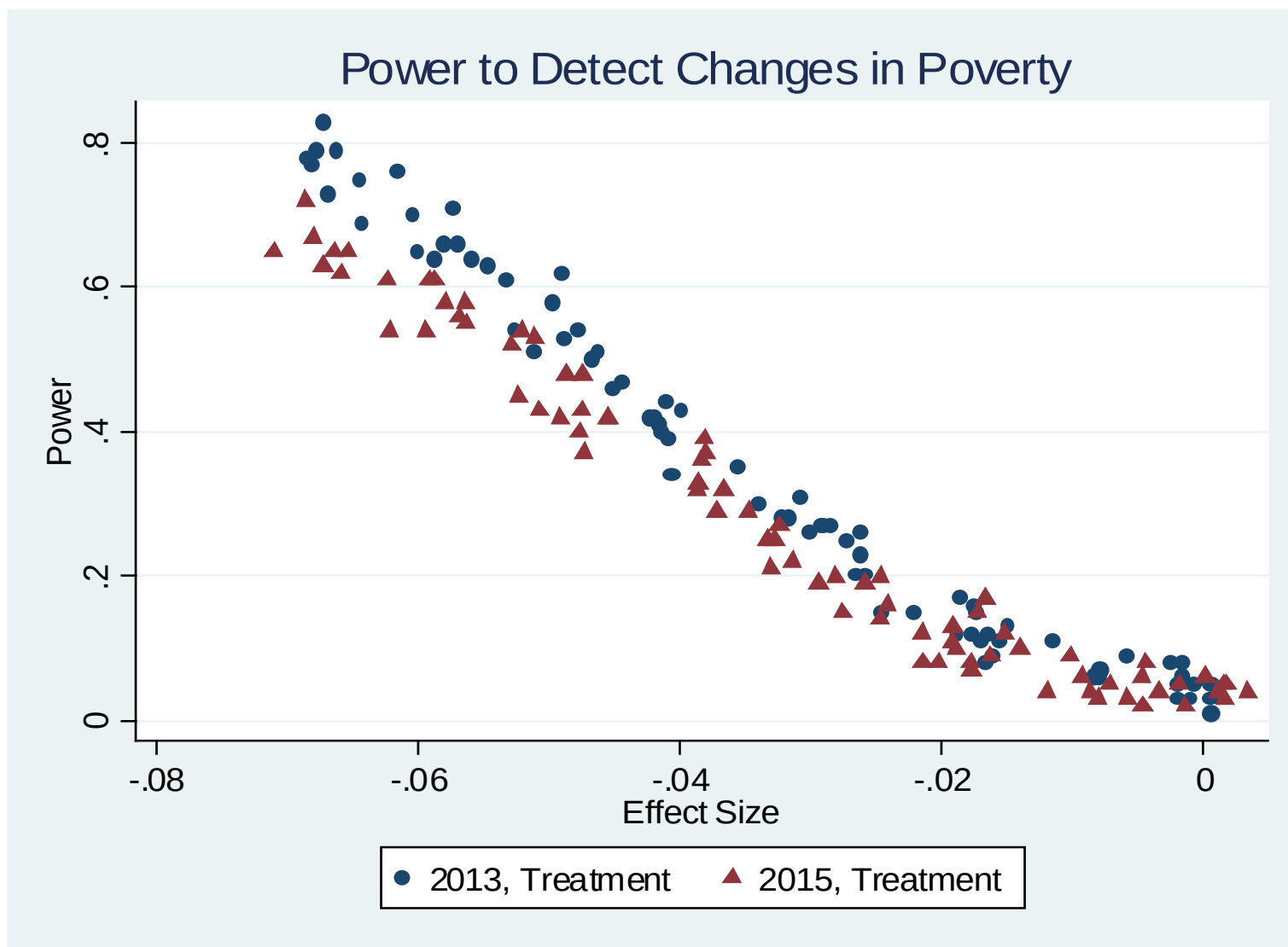


Figure 3 Sample Power to Detect Changes in Poverty



4 RESULTS

In this section, we present the impact estimates for the primary outcome indicators of the USAID-ACCESO activity. The data consists of a detailed baseline and two follow-up surveys collected from May to July of 2012, 2013, and 2015, respectively. The surveys allow us to calculate the indicators/outcomes for each year of the survey and compare the outcomes of the groups of households that participated in USAID-ACCESO to those that did not participate, both before and after the program went into full force.

The impact estimation consists of a household fixed effects difference in differences (DID) regression, weighted by propensity scores. The impact estimates exploit within-household variation in the outcome across time between the treatment and comparison group, where the comparison group is constructed using the weights defined in the methodology section. In the interest of space, we present only the relevant treatment effects; that is, the coefficients on the interaction terms between treatment indicator and the post-treatment time indicator variables. The estimates in the “Post-2013 & 2015” row are the average across both years, as described above. Namely, the impact estimates are the rows:

Treatment x 2013- Midline

Treatment x 2015- Endline

Treatment x Post- 2013 & 2015

In the first two columns (marked 1 and 2) of each table, we look at the treatment interaction in 2013 (midline) and 2015 (endline) separately; these are the yearly impacts. The last two columns in each table show the average impact in the 2013 and 2015 (midline and endline) surveys, indicated by a “post” treatment period (i.e. after the baseline). These “post” estimates present the combined effect of the program over the two survey years. We present the mean and standard deviation of the comparison group at baseline at the bottom of each table to contextualize our results.

Referring to the methodology section, the columns marked 1 and 3 in every table present results weighted with the propensity score weights (λ_i), while the second and fourth columns present results weighted with the attrition adjusted propensity score $\lambda_i/\Phi(\hat{\theta}Z_i)$. This is to show that our results are robust to correcting for attrition in our sample. We include household fixed effects in all regressions, as well as cluster standard errors at the village level. Bold letters in the table describe the outcome variable, and any additional notes are included in each table.

4.1 AGRICULTURAL PRODUCTION AND MARKET PARTICIPATION

In this sub-section, we discuss the main agricultural indicators. One of the main activities within USAID-ACCESO is the provision of agricultural extension services to promote good agricultural practices in order to improve yields and provide opportunities for participation in high-value cash crop markets.

With this in mind, we present results on crop selection, agricultural practices, and agricultural productivity.

Distribution of crops

The three main crops reported by the households in the surveyed area are beans, corn, and coffee. As USAID-ACCESO's primary objective was to ensure food security in farming households, our analysis of this agricultural indicator will focus on these three crops. However, as USAID-ACCESO also encouraged households to cultivate high-value cash crops after establishing food security, we will also explore some indicators on the production of vegetables, fruits, and tubers.

Figure 4 describes the number of farmers cultivating each major crop in the comparison and treatment groups for each year. Over all years of our survey, corn is the most popularly cultivated crop, followed by coffee and then beans. Other crops are less frequently cultivated; we group these into vegetables, fruits, and tubers. While the number of households cultivating corn and beans appear similar among treatment and comparison groups at baseline, midline, and endline, we see that more households in the treatment group cultivate coffee, fruits, vegetables, and tubers in each year. However, the overall number of households cultivating fruits, vegetables, and tubers is very low.

In Table 11 and Table 12, we use the household fixed effects difference in differences framework to estimate the impact of USAID-ACCESO on the probability that a farmer cultivated each of these major crops in the two agricultural seasons before each survey. Table 11 shows the results for corn, beans, and coffee, and Table 12 presents results for vegetables, fruits, and tubers.

Our results show that households that benefited from USAID-ACCESO's programs are between 5 and 7 percentage points more likely to have reported producing corn and beans, although these differences are only significant at the 10% range in some specifications. For coffee, vegetables, fruits, and tubers, we find no differential probability of production between treatment and comparison groups. We see that the program's impact on corn cultivation is larger — although less precisely estimated — at endline than at midline. In contrast, we see the program had a 6 percentage point increase in the probability that a

household cultivates beans at midline (this is again not significant at conventional levels) and no differential impact on cultivation between treatment and control households at endline. To further contextualize the program's impact, note that 80% of households in the control group farmed corn at baseline and 38% of the control group farmed beans at baseline. Our results suggest that USAID-ACCESO was responsible for a 13% increase in the number of households farming corn at midline and a 20% increase in the number of households farming corn at endline, in addition to a 16% increase in the number of households cultivating beans at midline.

USAID-ACCESO's agricultural technical assistance activities prioritized improvements in the cultivation of basic grains to establish households' food security before encouraging households to adopt and intensify production of income-generating crops. Given the proportion of households in our sample that fall below the national extreme poverty threshold (88%), it is unsurprising that we only observe increases in the number of households cultivating corn and beans, two popular staple crops; there are likely few food-secure households able to cultivate income-generating crops that were not already doing so. Furthermore, this can also explain the diminished effects we find at endline; there are fewer farmers who have yet to cultivate staple crops and improve their household food security.

Figure 4 Distribution of Crops

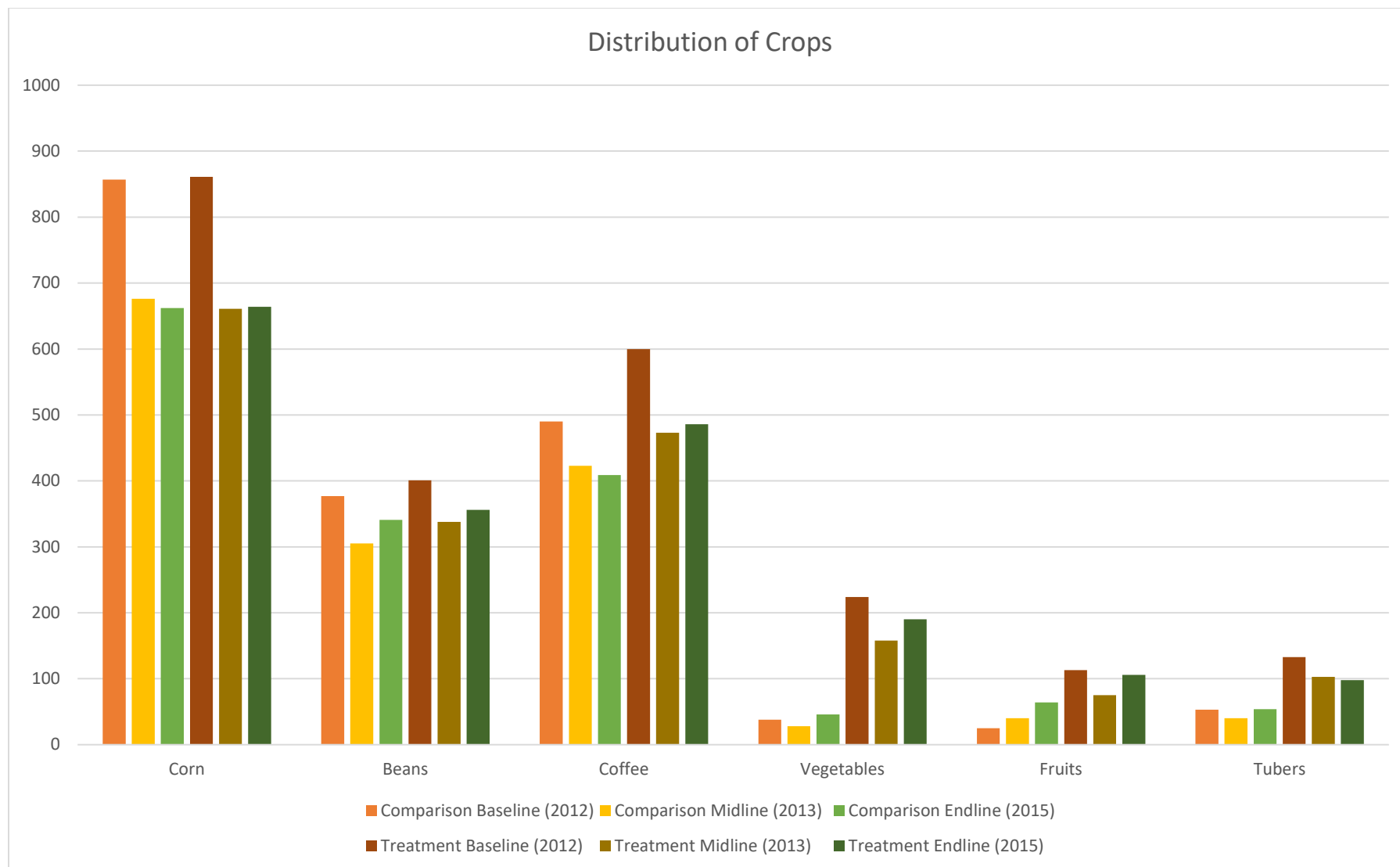


Table 11 Probability Household Grows Corn, Beans or Coffee

	(1)	(2)	(3)	(4)
Probability of growing corn				
Treatment x 2013- Midline	0.06 [0.030]**	0.059 [0.031]*		
Treatment x 2015- Endline	0.081 [0.045]*	0.082 [0.047]*		
Treatment x Post- 2013 & 2015			0.071 [0.034]**	0.071 [0.036]**
Mean of Control. at Baseline	0.8	0.8	0.8	0.8
SD of Comp. at Baseline	0.4	0.4	0.4	0.4
Observations	6252	6252	6252	6252
Probability of growing beans				
Treatment x 2013- Midline	0.063 [0.036]*	0.062 [0.035]*		
Treatment x 2015- Endline	0.043 [0.038]	0.043 [0.037]		
Treatment x Post- 2013 & 2015			0.053 [0.029]*	0.052 [0.029]*
Mean of Comp. at Baseline	0.38	0.38	0.38	0.38
SD of Comp. at Baseline	0.49	0.48	0.49	0.48
Observations	6252	6252	6252	6252
Probability of growing coffee				
Treatment x 2013- Midline	0.0023 [0.025]	0.0028 [0.024]		
Treatment x 2015- Endline	-0.0059 [0.026]	-0.0012 [0.025]		
Treatment x Post- 2013 & 2015			-0.0023 [0.024]	0.00036 [0.023]
Mean of Comp. at Baseline	0.49	0.49	0.49	0.49
SD of Comp. at Baseline	0.5	0.5	0.5	0.5
Number of Clusters	286	286	286	286
Number of Households	2673	2673	2673	2673
Observations	6252	6252	6252	6252

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 12 Prob. Household Grows vegetables, fruits, or tubers after ACCESO

	(1)	(2)	(3)	(4)
Probability of growing vegetables				
Treatment x 2013- Midline	-0.028 [0.018]	-0.028 [0.018]		
Treatment x 2015- Endline	-0.0042 [0.019]	-0.0046 [0.019]		
Treatment x Post- 2013 & 2015			-0.016 [0.013]	-0.016 [0.013]
Mean of Comp. at Baseline	0.027	0.026	0.027	0.026
SD of Comp. at Baseline	0.16	0.16	0.16	0.16
Observations	6252	6252	6252	6252
Probability of growing fruits				
Treatment x 2013- Midline	-0.014 [0.015]	-0.013 [0.015]		
Treatment x 2015- Endline	-0.018 [0.022]	-0.016 [0.021]		
Treatment x Post- 2013 & 2015			-0.016 [0.015]	-0.015 [0.015]
Mean of Comp. at Baseline	0.021	0.02	0.021	0.02
SD of Comp. at Baseline	0.14	0.14	0.14	0.14
Observations	6252	6252	6252	6252
Probability of growing tubers				
Treatment x 2013- Midline	-0.012 [0.012]	-0.012 [0.012]		
Treatment x 2015- Endline	-0.016 [0.011]	-0.016 [0.011]		
Treatment x Post- 2013 & 2015			-0.014 [0.010]	-0.014 [0.010]
Mean of Comp. at Baseline	0.067	0.065	0.067	0.065
SD of Comp. at Baseline	0.25	0.25	0.25	0.25
Number of Clusters	286	286	286	286
Number of Households	2673	2673	2673	2673
Observations	6252	6252	6252	6252

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Use of Inputs: Fertilizer, Fungicides, Labor, and Land

To explore effects on agricultural practices, we estimate the impact of participating in USAID-ACCESO on the probability of using fertilizers and fungicides and on the days of paid labor that are used. To allow for these outcomes to differ by the type of crop, we estimate separate equations for corn, beans, and coffee.

The specific agricultural inputs and practices promoted by USAID-ACCESO varied depending on the crop cultivated. In general, adopting most or all of the basic recommendations for a particular crop would theoretically double or triple yields from baseline. For farmers cultivating corn and beans, ACCESO recommended:

- 1) Basic land preparation and contouring where necessary;
- 2) Seed selection (with existing types);
- 3) Specific planting distance with one seed per hole (high density);
- 4) Liming of soils;
- 5) Fertilizer applications applied in solutions with more frequent applications (where fertilizers were used, the total volume of fertilizer used was equal to or less than that used traditionally); and
- 6) Weed control

For farmers already cultivating coffee, USAID-ACCESO recommended:

- 1) Liming of soils;
- 2) Diluted and more frequent fertilizer applications (where fertilizers were used);
- 3) Plant pruning (after harvest); and
- 4) Control of soil-borne insects

Only farmers who had established and/or made improvements in basic grain cultivation were introduced to high-value crops as a means of income generation. USAID-ACCESO recommended a diverse array of practices and inputs for production of these crops, from constructing improved irrigation infrastructure to adopting improved record-keeping practices. Table 13 shows the project's impact on the use of chemical fertilizer, organic fertilizer, and fungicides/herbicides by agricultural households. We find no significant differences in these variables, likely because aggregating the positive and negative changes in input intensity required to achieve an optimal level of use among all crops results in a net zero average change in input use. This masks any shifts that farmers might have made in

their farming practices. Table 14 disaggregates the estimation by crop to better identify the impact of USAID-ACCESO's extension activities on particular crops. We find that bean growers in the treatment group have a slightly lower probability of using organic fertilizer (14 percentage points) and that coffee growers are slightly more likely to use chemical fertilizers (12 percentage points). Both of these results are only apparent in the midline and are only significant at the 10% level. The results for coffee are consistent with USAID-ACCESO recommendations; although our results are not significant at conventional levels, the low variance and high level of baseline chemical fertilizer use make it more likely that our findings are a causal effect of USAID-ACCESO. The lack of negative impact on corn and beans may be because farmers' current levels of input use were optimal, or it may be due to the fact that convincing farmers to cut down on fertilizer application is more difficult than encouraging increased use. The drop in use of organic fertilizer is similarly puzzling, although given the extremely low levels of organic fertilizer use at baseline, and the high variance of that use, our imprecisely estimated results may not indicate a causal effect of the program.

We find that after USAID-ACCESO's program activities, vegetable growers and tuber farmers are much more likely to use chemical fertilizers (25 and 30-50 percentage points more likely, respectively). We do not find any effect among vegetable farmers, which are robust to different specifications. The strength of our results suggests that USAID-ACCESO encouraged farmers to increase fertilizer use for high-value crops.

Surprisingly, we see a corresponding 29 percentage drop in the use of organic fertilizer for tubers, as well as a large drop in the use of organic fertilizers by farmers cultivating vegetables. In the case of tubers, it may be that farmers are substituting chemical fertilizer for organic fertilizers, as the cost of labor required for frequent chemical fertilizer applications may make additional applications of additional organic fertilizer prohibitively expensive. Alternatively, USAID-ACCESO's specific recommendations for tuber farmers may have prompted this change in behavior. Similarly, the highly significant, large magnitude drop in organic fertilizer use by vegetables farmers may be a consequence of USAID-ACCESO's extension recommendations for vegetable crops. Finally, we see that farmers cultivating vegetables use 32 percentage points less fungicide/herbicide at endline, while farmers cultivating fruits use 13 percentage points more at endline. While our results measure changes in input use, we are unable to determine whether farmers are using inputs at the optimal level given the crop they cultivate, the plot's soil quality, climate conditions, and other variables.

Table 13 Inputs: Use of chemical and organic fertilizers, herbicides, and fungicides

	(1)	(2)	(3)	(4)
Use of chemical fertilizer				
Treatment x 2013- Midline	0.062 [0.040]	0.061 [0.040]		
Treatment x 2015- Endline	-0.018 [0.035]	-0.017 [0.036]		
Treatment x Post- 2013 & 2015			0.019 [0.032]	0.019 [0.033]
Mean of Comp. at Baseline	0.69	0.69	0.69	0.69
SD of Comp. at Baseline	0.46	0.46	0.46	0.46
Observations	6267	6267	6267	6267
Use of organic fertilizer				
Treatment x 2013- Midline	-0.032 [0.027]	-0.03 [0.027]		
Treatment x 2015- Endline	-0.039 [0.031]	-0.035 [0.031]		
Treatment x Post- 2013 & 2015			-0.035 [0.025]	-0.033 [0.025]
Mean of Comp. at Baseline	0.083	0.085	0.083	0.085
SD of Comp. at Baseline	0.28	0.28	0.28	0.28
Observations	6267	6267	6267	6267
Uses fungicides/herbicides				
Treatment x 2013- Midline	-0.019 [0.044]	-0.019 [0.043]		
Treatment x 2015- Endline	0.016 [0.044]	0.016 [0.043]		
Treatment x Post- 2013 & 2015			-0.00079 [0.040]	-0.001 [0.039]
Mean of Comp. at Baseline	0.3	0.31	0.3	0.31
SD of Comp. at Baseline	0.46	0.46	0.46	0.46
Number of Clusters	286	286	286	286
Number of Households	2675	2675	2675	2675
Observations	6267	6267	6267	6267

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 14 Inputs for Corn, Beans and Coffee: Use of fertilizer, herbicides, and fungicides

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Use of chemical fertilizer	Corn				Beans				Coffee			
Treatment x 2013- Midline	0.03 [0.054]	0.031 [0.054]			0.11 [0.11]	0.11 [0.11]			0.12 [0.060]*	0.12 [0.060]*		
Treatment x 2015- Endline	-0.048 [0.047]	-0.047 [0.048]			0.0055 [0.093]	0.0062 [0.093]			0.069 [0.057]	0.073 [0.059]		
Treatment x Post- 2013 & 2015			-0.011 [0.040]	-0.01 [0.041]			0.057 [0.086]	0.056 [0.086]			0.091 [0.051]*	0.092 [0.052]*
Mean of Comp. at Baseline	0.64	0.64	0.64	0.64	0.38	0.37	0.38	0.37	0.57	0.58	0.57	0.58
SD of Comp. at Baseline	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.49	0.49	0.49	0.49
Observations	4586	4586	4586	4586	2208	2208	2208	2208	3047	3047	3047	3047
Use of organic fertilizer												
Treatment x 2013- Midline	0.018 [0.028]	0.019 [0.028]			-0.14 [0.073]*	-0.14 [0.073]*			-0.039 [0.038]	-0.039 [0.037]		
Treatment x 2015- Endline	-0.025 [0.028]	-0.024 [0.028]			-0.088 [0.061]	-0.091 [0.060]			-0.00043 [0.023]	0.002 [0.022]		
Treatment x Post- 2013 & 2015			-0.0045 [0.023]	-0.0037 [0.023]			-0.11 [0.061]*	-0.12 [0.060]*			-0.018 [0.025]	-0.016 [0.024]
Mean of Comp. at Baseline	0.025	0.025	0.025	0.025	0.018	0.018	0.018	0.018	0.1	0.1	0.1	0.1
SD of Comp. at Baseline	0.15	0.16	0.15	0.16	0.13	0.13	0.13	0.13	0.3	0.3	0.3	0.3
Observations	4586	4586	4586	4586	2208	2208	2208	2208	3047	3047	3047	3047
Uses fungicides/herbicides												
Treatment x 2013- Midline	-0.07 [0.051]	-0.073 [0.049]			0.0077 [0.12]	0.0043 [0.11]			0.031 [0.052]	0.031 [0.052]		
Treatment x 2015- Endline	0.021 [0.050]	0.016 [0.049]			0.026 [0.073]	0.025 [0.072]			0.018 [0.059]	0.02 [0.060]		
Treatment x Post- 2013 & 2015			-0.023 [0.039]	-0.026 [0.038]			0.017 [0.082]	0.015 [0.081]			0.024 [0.050]	0.025 [0.051]
Mean of Comp. at Baseline	0.26	0.26	0.26	0.26	0.21	0.21	0.21	0.21	0.14	0.15	0.14	0.15
SD of Comp. at Baseline	0.44	0.44	0.44	0.44	0.41	0.41	0.41	0.41	0.35	0.35	0.35	0.35
Number of Clusters	274	274	274	274	235	235	235	235	223	223	223	223
Number of Households	2262	2262	2262	2262	1424	1424	1424	1424	1504	1504	1504	1504
Observations	4586	4586	4586	4586	2208	2208	2208	2208	3047	3047	3047	3047

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household-crop level and all regressions include household fixed effects, and crop indicators. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 15: Inputs for Vegetables, Fruits and Tubers: Use of fertilizer, herbicides, and fungicides

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)
Use of chemical fertilizer	Vegetables				Fruits				Tubers		
Treatment x 2013- Midline	-0.2 [0.12]*	-0.2 [0.12]			0.26 [0.076]***	0.25 [0.075]***			0.53 [0.22]**		
Treatment x 2015- Endline	-0.27 [0.23]	-0.25 [0.24]			0.28 [0.090]***	0.28 [0.090]***			0.07 [0.10]		
Treatment x Post- 2013 & 2015			-0.25 [0.19]	-0.24 [0.20]			0.27 [0.072]***	0.26 [0.072]***		0.3 [0.13]**	0.29 [0.13]**
Mean of Comp. at Baseline	0.34	0.37	0.34	0.37	0.049	0.049	0.049	0.049	0.81	0.81	0.81
SD of Comp. at Baseline	0.48	0.49	0.48	0.49	0.22	0.22	0.22	0.22	0.4	0.39	0.4
Observations	719	719	719	719	457	457	457	457	500	500	500
Use of organic fertilizer											
Treatment x 2013- Midline	-0.44 [0.16]***	-0.43 [0.16]***			-0.19 [0.22]	-0.19 [0.22]			-0.28 [0.15]*		
Treatment x 2015- Endline	-0.38 [0.18]**	-0.37 [0.18]**			-0.42 [0.26]	-0.43 [0.27]			-0.3 [0.13]**		
Treatment x Post- 2013 & 2015			-0.39 [0.13]***	-0.38 [0.13]***			-0.31 [0.27]	-0.32 [0.27]		-0.29 [0.14]**	-0.29 [0.13]**
Mean of Comp. at Baseline	0.15	0.15	0.15	0.15	0.049	0.049	0.049	0.049	0.098	0.098	0.098
SD of Comp. at Baseline	0.36	0.36	0.36	0.36	0.22	0.22	0.22	0.22	0.3	0.3	0.3
Observations	719	719	719	719	457	457	457	457	500	500	500
Uses fungicide/herbicide											
Treatment x 2013- Midline	-0.48 [0.30]	-0.49 [0.29]			0.058 [0.071]	0.053 [0.070]			-0.28 [0.22]		
Treatment x 2015- Endline	-0.32 [0.12]***	-0.32 [0.13]**			0.13 [0.062]**	0.13 [0.060]**			-0.24 [0.19]		
Treatment x Post- 2013 & 2015			-0.35 [0.12]***	-0.35 [0.13]***			0.099 [0.053]*	0.095 [0.051]*		-0.27 [0.19]	-0.26 [0.19]
Mean of Comp. at Baseline	0.27	0.27	0.27	0.27	0	0	0	0	0.28	0.28	0.28
SD of Comp. at Baseline	0.45	0.45	0.45	0.45	0	0	0	0	0.45	0.45	0.45
Number of Clusters	81	81	81	81	108	108	108	108	65	65	65
Number of Households	311	311	311	311	310	310	310	310	278	278	278
Observations	719	719	719	719	457	457	457	457	500	500	500

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household-crop level (Vegetables) and all regressions include household fixed effects, and controls for the number of crops in each group. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 16 shows the impact on the input and labor costs for these crops. We find significant results for the labor costs to produce coffee. From columns 3 and 4 of the bottom panel in Table 16, on average, treatment households spend nearly \$750 more per year on labor than similar households in the comparison group. This increase in labor costs could be related to the incidence of coffee rust in the region in follow-up years, which increased the amount of labor needed in a season to prune the plant⁶. While the incidence of coffee rust likely affected treatment and control households equally, the increased expenditures on labor in the treatment group suggests that ACCESO's package of extension and financial services helped households in the treatment group mitigate the effects of coffee rust, either by providing them financial assistance (loans) to hire labor or by helping them substitute labor for other inputs. Table 17 shows the same indicators for vegetables, fruits, and tubers; we find that labor costs increase by around \$925 for members of the treatment group who cultivate fruits, which is significant at conventional levels. As ACCESO encouraged farmers to begin cultivation of fruit trees, the increase in labor costs is consistent with the labor-intensive process of beginning fruit cultivation (land preparation and planting likely require large amounts of labor, depending on the scale of cultivation), compared to farmers in the control group who likely had established orchards requiring less labor. Farmers who cultivate tubers in our treatment group spend more on inputs and less on labor at midline, but this effect is not present at endline.

Figure 5 shows the distribution of hectares planted by the major type of crops in the sample. Most of the mass of the distribution is to the left, with the area planted below two hectares for each crop. We note that most of the farmers in the sample have small farms and that coffee farms tend to be larger.

Table 18 shows the descriptive statistics for area planted for each type of crop at baseline. We can see that the average and median-sized plots tend to be under one hectare and that coffee plots have larger values, with a 1.15 hectares for the comparison group and 1.22 hectares for the treatment group (unweighted) mean at baseline.

⁷ We estimate the same regressions valuing production using the median price of the crop in the department and the results are qualitatively the same.

Table 16 Input costs for Corn, Beans and Coffee: Products and Labor

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Total Input cost in 2005 PPP to 2010 USD	Corn				Beans				Coffee			
Treatment x 2013- Midline	-20.6 [20.7]	-20 [20.6]			-17 [12.6]	-18 [12.7]			-106.7 [419.9]	-136.8 [434.0]		
Treatment x 2015- Endline	-27.1 [14.7]*	-26.7 [14.7]*			-19.3 [16.8]	-21 [16.6]			-399.2 [817.6]	-444.1 [857.4]		
Treatment x Post- 2013 & 2015			-24 [15.3]	-23.6 [15.2]			-18.2 [12.7]	-19.6 [12.6]			-268.4 [637.5]	-309.7 [671.3]
Mean of Comp. at Baseline	210.4	208.7	210.4	208.7	63.2	63.4	63.2	63.4	909.8	904.9	909.8	904.9
SD of Comp. at Baseline	228.6	227.3	228.6	227.3	82.6	82.5	82.6	82.5	1561.1	1550.7	1561.1	1550.7
Observations	4421	4421	4421	4421	1904	1904	1904	1904	2771	2771	2771	2771
Total labor cost in 2005 PPP to 2010 USD												
Treatment x 2013- Midline	-10.6 [17.4]	-11 [16.9]			-7.24 [19.4]	-6.61 [19.2]			607.2 [227.3]***	611.7 [224.6]***		
Treatment x 2015- Endline	-25.4 [15.4]	-25.7 [15.0]*			18.2 [24.0]	19.2 [23.8]			853.4 [299.4]***	871.8 [309.5]***		
Treatment x Post- 2013 & 2015			-18.4 [15.9]	-18.9 [15.4]			5.49 [21.2]	6.29 [21.0]			742.8 [259.0]***	757.4 [264.2]***
Mean of Comp. at Baseline	137.9	137.8	137.9	137.8	80.7	81	80.7	81	1449	1462.8	1449	1462.8
SD of Comp. at Baseline	209.9	211.3	209.9	211.3	141.2	141	141.2	141	2559.9	2566.6	2559.9	2566.6
Number of Clusters	231	231	231	231	154	154	154	154	183	183	183	183
Number of Households	1356	1356	1356	1356	504	504	504	504	940	940	940	940
Observations	2997	2997	2997	2997	883	883	883	883	2163	2163	2163	2163

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household-crop level and all regressions include household fixed effects, and crop indicators. . Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 17: Input costs for Vegetables, Fruits, and Tubers: Products and Labor

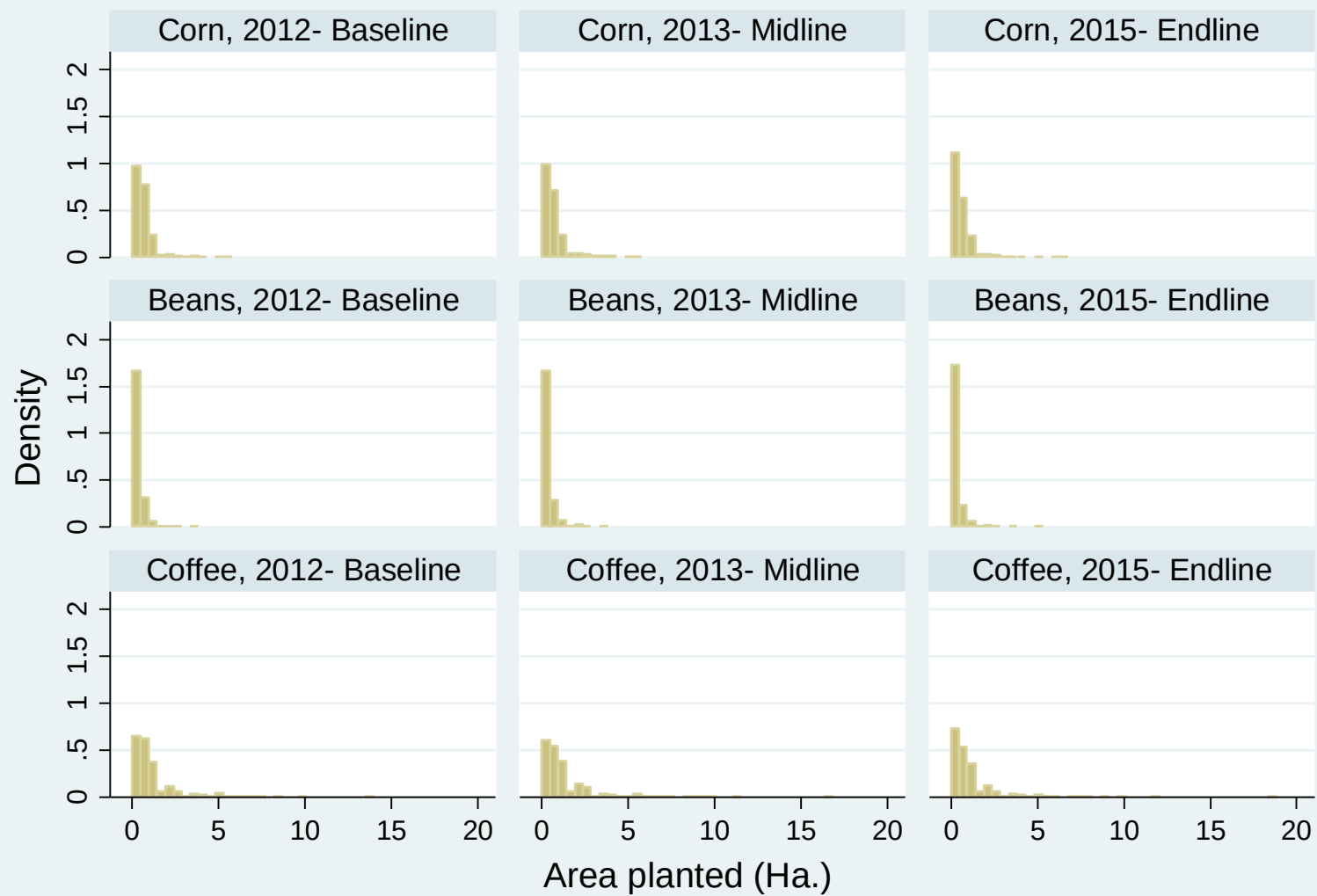
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Total Input cost in 2005 PPP to 2010 USD	Vegetables				Fruits				Tubers			
Treatment x 2013- Midline	-424.5 [296.0]	-424.5 [298.6]			-66.8 [156.1]	-96.2 [170.2]		652.4	631.4 [340.9]*		[339.0]*	
Treatment x 2015- Endline	-79.4 [107.7]	-63.4 [107.8]			981 [1026.8]	979.3 [1042.1]			-26.8 [489.6]	-18.2 [490.0]		
Treatment x Post- 2013 & 2015			-152.1 [127.0]	-137.8 [127.6]			966.2 [1008.8]	964.6 [1024.1]			297.9 [370.3]	292 [370.3]
Mean of Comp. at Baseline	166.9	164.2	166.9	164.2	50	49.4	50	49.4	766.2	768.3	766.2	768.3
SD of Comp. at Baseline	282.1	277.7	282.1	277.7	50	49.7	50	49.7	1149.1	1152.8	1149.1	1152.8
Number of Clusters	70	70	70	70	65	65	65	65	51	51	51	51
Number of Households	246	246	246	246	141	141	141	141	233	233	233	233
Observations	609	609	609	609	211	211	211	211	448	448	448	448
Total labor cost in 2005 PPP to 2010 USD												
Treatment x 2013- Midline	-319 [205.1]	-312.2 [202.4]			925.8 [191.0]***	927.7 [190.0]***			-266.4 [129.5]**	-252.7 [122.9]**		
Treatment x 2015- Endline	-248.8 [208.5]	-236.4 [198.2]			904.2 [261.8]***	904.3 [260.7]***			-158 [181.4]	-143.4 [175.5]		
Treatment x Post- 2013 & 2015			-259.7 [204.4]	-248.1 [195.4]			925.5 [190.3]***	927.4 [189.2]***			-209.8 [153.4]	-195.5 [147.3]
Mean of Comp. at Baseline	207.7	194.5	207.7	194.5	581.8	570	581.8	570	300.1	300.3	300.1	300.3
SD of Comp. at Baseline	458.7	442.9	458.7	442.9	849.3	749.3	849.3	749.3	1027.3	1019.9	1027.3	1019.9
Number of Clusters	45	45	45	45	32	32	32	32	43	43	43	43
Number of Households	161	161	161	161	67	67	67	67	198	198	198	198
Observations	343	343	343	343	104	104	104	104	395	395	395	395

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household-crop level (Vegetables) and all regressions include household fixed effects, and controls for the number of crops in each group. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Figure 5 Distribution of Area Planted



Graphs by (firstnm) major_crop and Year

Table 18 Descriptive Statistics of Area Planted (ha.)

		2012		2013		2015	
		Comparison	Treatment	Comparison	Treatment	Comparison	Treatment
Corn	Mean	0.60	0.61	0.55	0.68	0.75	0.68
	Median	0.49	0.46	0.35	0.65	0.70	0.52
	Observations	857	676	771	861	661	760
Beans	Mean	0.32	0.31	0.30	0.37	0.38	0.38
	Median	0.27	0.23	0.17	0.35	0.28	0.35
	Observations	377	305	383	401	338	404
Coffee	Mean	1.15	1.22	1.03	1.15	1.45	1.29
	Median	0.70	0.70	0.70	0.70	0.85	0.70
	Observations	490	423	495	600	473	566
Vegetables	Mean	0.13	0.24	0.31	0.32	0.25	0.49
	Median	0.09	0.17	0.17	0.14	0.13	0.13
	Observations	38	28	49	224	158	222
Fruits	Mean	0.31	0.31	0.45	0.66	1.14	0.39
	Median	0.35	0.16	0.17	0.35	0.70	0.17
	Observations	25	40	79	113	75	125
Tubers	Mean	0.31	0.45	0.38	0.38	0.62	0.42
	Median	0.21	0.35	0.17	0.21	0.35	0.17
	Observations	53	40	57	133	103	114

Table 19 estimates the impact for all types of crops planted in total and for corn, beans, and coffee separately. We see that cultivated areas for coffee among treatment households increase by 0.35 ha despite the prevalence of coffee rust, which suggests that USAID-ACCESO's activities allowed farmers to expand their cultivation of a profitable crop and mitigate their losses due to coffee rust, while farmers in the control group were hesitant to take on more risk by increasing their acreage of coffee cultivated without the safety net provided by USAID-ACCESO. While we observe increases in the area of corn and tubers planted, these are only significant at the 10% significance level.

Table 20 shows the impact estimates for vegetables, fruits, and tubers. We find decreases in vegetable area planted in 2013 and no effects in 2015. The changes in area planted to fruits and tubers are not significant at conventional levels.

In Table 21 and Table 22, we present the impact estimates using an inverse hyperbolic sine transform that allows us to correct for over-dispersion area planted; the estimated effects can be interpreted as a percentage increase in the baseline area planted. We present the results separately for each crop. We find that participation in USAID-ACCESO led to a 12 to 15% increase in the acreage of corn and coffee planted, but we do not find any other impacts significant at conventional levels.

To summarize our results for agricultural production, input use, and area cultivated, we find that:

- Inputs Used
 - Significant increase in chemical fertilizer use among households cultivating coffee, fruit, and tubers;
 - Decrease in organic fertilizer use among households cultivating beans;
 - Decreased in fungicide/herbicide use by vegetable farmers; and
 - Increase in fungicide/herbicide use by fruit farmers
- Labor and Input Costs
 - Increase in labor costs for fruit, coffee;
 - Decrease in input costs for tubers; and
 - Increase in input costs for fruit
- Area Cultivated
 - Significant increase in area of coffee and corn cultivated;
 - Imprecisely measured increase in the area of tubers cultivated; and
 - Decrease in area of vegetables cultivated at midline

Table 19 Area Planted (Ha.): All, Corn, Beans, and Coffee

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Area planted (Ha.)	All crops				Corn				Beans				Coffee			
Treatment x 2013- Midline	0.11	0.1			0.27	0.26			0.12	0.12			0.27	0.28		
	[0.13]	[0.13]			[0.15]*	[0.15]*			[0.10]	[0.099]			[0.12]**	[0.12]**		
Treatment x 2015- Endline	0.057	0.034			0.32	0.3			0.12	0.11			0.44	0.45		
	[0.18]	[0.19]			[0.16]*	[0.16]*			[0.14]	[0.13]			[0.22]**	[0.23]**		
Treatment x Post- 2013 & 2015			0.084	0.068			0.29	0.28			0.12	0.11			0.35	0.37
			[0.15]	[0.15]			[0.15]*	[0.15]*			[0.11]	[0.11]			[0.16]**	[0.17]**
Mean of Comp. at Baseline	1.29	1.3	1.29	1.3	1.02	1.01	1.02	1.01	0.48	0.47	0.48	0.47	2.65	2.7	2.65	2.7
SD of Comp. at Baseline	2.29	2.34	2.29	2.34	1.25	1.25	1.25	1.25	0.57	0.56	0.57	0.56	3.76	3.83	3.76	3.83
Number of Clusters	285	285	285	285	271	271	271	271	230	230	230	230	222	222	222	222
Number of Households	2649	2649	2649	2649	2189	2189	2189	2189	1335	1335	1335	1335	1462	1462	1462	1462
Observations	10972	10972	10972	10972	4347	4347	4347	4347	2061	2061	2061	2061	2866	2866	2866	2866

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household-crop level and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 20 Area Planted (ha.): Vegetables, Fruits, and Tubers

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Area planted (Ha.)	Vegetables				Fruits				Tubers			
Treatment x 2013- Midline	-0.63 [0.26]**	-0.6 [0.25]**			-1.01 [0.71]	-0.99 [0.71]			0.53 [0.33]	0.52 [0.32]		
Treatment x 2015- Endline	-0.08 [0.16]	-0.066 [0.14]			-1.15 [0.68]*	-1.14 [0.68]*			0.48 [0.25]*	0.47 [0.25]*		
Treatment x Post- 2013 & 2015			-0.19 [0.21]	-0.17 [0.19]			-0.93 [0.71]	-0.93 [0.70]			0.53 [0.27]*	0.52 [0.27]*
Mean of Comp. at Baseline	0.12	0.12	0.12	0.12	0.37	0.37	0.37	0.37	0.54	0.54	0.54	0.54
SD of Comp. at Baseline	0.14	0.14	0.14	0.14	0.21	0.21	0.21	0.21	0.57	0.56	0.57	0.56
Number of Clusters	79	79	79	79	88	88	88	88	64	64	64	64
Number of Households	289	289	289	289	235	235	235	235	259	259	259	259
Observations	645	645	645	645	359	359	359	359	466	466	466	466

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level (Vegetables) and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 21 Percentage area Planted (Ha) Vegetables, Fruits, and Tubers

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
aSinH-Log - Area planted (Ha.)	Vegetables				Fruits				Tubers			
Treatment x 2013- Midline	-0.45	-0.43			-0.34	-0.35			0.36	0.35		
	[0.18]**	[0.17]**			[0.38]	[0.38]			[0.27]	[0.27]		
Treatment x 2015- Endline	-0.02	-0.016			-0.64	-0.63			0.37	0.37		
	[0.075]	[0.071]			[0.36]*	[0.36]*			[0.19]*	[0.19]*		
Treatment x Post- 2013 & 2015			-0.1	-0.091			-0.52	-0.51			0.38	0.37
			[0.14]	[0.13]			[0.36]	[0.36]			[0.22]*	[0.22]
Mean of Comp. at Baseline	0.12	0.12	0.12	0.12	0.35	0.35	0.35	0.35	0.46	0.46	0.46	0.46
SD of Comp. at Baseline	0.14	0.13	0.14	0.13	0.2	0.2	0.2	0.2	0.44	0.44	0.44	0.44
Number of Clusters	79	79	79	79	88	88	88	88	64	64	64	64
Number of Households	289	289	289	289	235	235	235	235	259	259	259	259
Observations	645	645	645	645	359	359	359	359	466	466	466	466

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level (Vegetables) and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. * p<0.10, ** p<0.05, *** p<0.01

Table 22 Percentage area planted (Ha) Corn, Beans, and Coffee

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
aSinH-Log - Area planted (Ha.)	Corn				Beans				Coffee			
Treatment x 2013- Midline	0.12	0.12			0.093	0.09			0.12	0.12		
	[0.061]**	[0.060]*			[0.077]	[0.077]			[0.044]***	[0.042]***		
Treatment x 2015- Endline	0.15	0.14			0.077	0.074			0.15	0.15		
	[0.068]**	[0.067]**			[0.088]	[0.085]			[0.036]***	[0.035]***		
Treatment x Post- 2013 & 2015			0.13	0.13			0.085	0.082			0.13	0.13
			[0.062]**	[0.061]**			[0.075]	[0.073]			[0.034]***	[0.033]***
Mean of Comp. at Baseline	0.73	0.72	0.73	0.72	0.41	0.41	0.41	0.41	1.22	1.23	1.22	1.23
SD of Comp. at Baseline	0.58	0.58	0.58	0.58	0.38	0.37	0.38	0.37	0.92	0.93	0.92	0.93
Number of Clusters	271	271	271	271	230	230	230	230	222	222	222	222
Number of Households	2189	2189	2189	2189	1335	1335	1335	1335	1462	1462	1462	1462
Observations	4347	4347	4347	4347	2061	2061	2061	2061	2866	2866	2866	2866

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household-crop level and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. * p<0.10, ** p<0.05, *** p<0.01

Agriculture Production and Sales

In the previous section, we found limited effects from the use of inputs (chemical and organic fertilizer, fungicide, and labor) that could increase productivity for each of the major crop types targeted by USAID_ACCESO. Additionally, many of the positive impacts we observed from input use, labor, and acreage planted were concentrated in coffee production. As we acknowledged, observing the levels of input use does not give us any information regarding whether farmers are using inputs at the optimal level or whether they are using the appropriate combination of inputs required by their particular farming conditions. USAID-ACCESO may have provided information to each farmer in our treatment group that would allow them to make specific adjustments to increase the productivity of their crops. Thus, in this section we examine the program's impact on agricultural productivity; if the program's extension services were effective, we would expect to see agricultural productivity increasing more in the treatment group than in the comparison group.

Table 23 shows the impact of USAID-ACCESO on the quantity of agricultural output in kilograms, as well as the value of this production in 2005 PPP dollars inflated to 2010 US prices, using both reported prices and the median prices for each year⁷. The regression combining outcomes for all crops includes crop fixed effects to account for differences in the characteristics (e.g. densities) of different crops. We find an increase in production by about 4,000 kg at endline, significant at conventional levels, although we fail to find any effect at midline. As in other extension programs, many farmers may delay adoption of new technologies or practices until they see results among their peers, which may explain the lack of results at midline.

It is important to disaggregate the production values by crop, since the farmers who obtain the major part of their agricultural income from coffee are likely very different than those who farm corn primarily for household consumption. Table 24 and Table 25 disaggregate the impact estimate by crop for production. In general, we do not find coherent patterns in the few impacts we find for specific crops. Corn production in the treatment group remains unchanged in the midline and endline, although we see that the reported value of production falls at midline. Treatment farmers of other crops (namely, beans and fruits) also report lower prices at midline. Our midline results for corn are not robust to using median prices; as there is no difference in the quantity of production, this would suggest that treatment households are systematically reporting lower prices for their crop than their control counterparts. If we

⁷ We estimate the same regressions valuing production using the median price of the crop in the department and the results are qualitatively the same.

do not think there is a reporting bias correlated with treatment, and if we accept that treatment households receive lower prices for their output, this could be for several reasons. Treatment households may produce lower quality varieties of corn or may sell in markets that give a lower price. Treatment household may also devote less time to post-harvest value added activities, resulting in a lower price.

On the other hand, our midline results for the value of production of beans and fruit are robust to use of the department median price. This is consistent with the fact that we observe a slight drop in the quantity of production for both but are not able to precisely estimate the drop at conventional levels of significance. However, when we look at the value of production using reported and median prices, we find that treated households produce less than their control counterparts. One explanation for the drop in the value of fruit production may be because more treatment households began cultivating fruits after USAID-ACCESO's training but were yet to harvest their crops, reporting a value of 0 for production. As the fruit farmers in the control group were likely well established and already harvesting their fruit crops, this would explain our result. However, it also may simply be that the 2012-2013 season was bad for fruit cultivation and thus affected treatment farmers disproportionately.

We observe the opposite for the production of tubers; the quantity produced increases by 2500 kg at midline and, while we do not detect any increase in the value of production using reported prices, we find a significant increase when using midline prices. This pattern suggests that treated households may have received systematically lower prices than their control counterparts (perhaps due to varieties cultivated or time spent on post-harvest activities), which would suggest there was no difference in the total value of production. Using the median price masks this variation in reported prices and suggests that treated households are cultivating more than their control counterparts.

We see no other midline results, but we do find a significant increase for the production of vegetables by the treatment group at endline. This is consistent with the fact that USAID-ACCESO encouraged both cultivation of high-value horticultural crops and improvement in staple crop yields to ensure households' food security. Our results show a 25,000 kg increase in the quantity of vegetables cultivated; we do not see any impact on the value of production when using the reported price, but we do detect an increase in value when using the median price, although this is not significant at conventional levels. This suggests high variance in reported prices, which makes sense as we group a variety of horticultural prices in a single category. We would expect there to be systematic differences in the horticultural crops cultivated by treatment and control households after the intervention, depending

on the specific crops targeted by USAID-ACCESO. This would suggest that using the median price to correct for inaccuracies in reporting would mask differences in the different values of crops produced and suggest that treated households are not as well off as they actually are.

Lastly, contrary to our midline findings, we find that there is no difference in the quantity of tuber production between treatment and control farmers at endline, but that treated farmers report significantly lower prices than their control counterparts and have a lower value of production. This effect disappears when using median prices, suggesting that treatment households systematically report lower prices. As before, this may be due to the quality of output, systematic differences in the varieties cultivated, time spent on post-harvest value added activities, or the markets in which treated and control households sell their produce. It seems that the gains in outputs experienced at midline did not persist until endline.

Table 26 and Table 27 report the same outcomes using an inverse hyperbolic sine transform of quantity and value outcomes. Our results generally echo the findings in Table 24 and Table 25. We find no impacts for corn, coffee, fruits, or tubers. As we found impacts while using the raw (untransformed) data, we theorize that our results may have been driven by a small handful of outliers.

We do find consistent effects for beans and vegetables, however. On average, we find the value of production for treated households that farm beans is 1% less than that of their control counterparts. We see this change in value when using reported prices (we expect no change while using the median price because there is little variation in the prices in our dataset), but do not detect a change in the quantity. This suggests that the treatment households' systematically lower reported prices are not driven by outliers but may instead be due to differences in production. In contrast, we see that treated households cultivate a nearly 3% higher quantity of vegetables than their control counterparts and that this result is not driven by outliers. We do not observe a higher value of producing using reported prices, but that is likely due to the fact that the inverse hyperbolic sine transform remedies outliers and not a high variance of data. We see a similar 3% increase in the value of production when using the median price, confirming that treatment group vegetable production is higher.

In Table 28, we aggregate the value of all crops produced by each household under alternative pricing schemes to obtain the total value of agricultural production in each household. We find no systematic difference in the value of the agricultural production of households in the treatment group versus the comparison group. Aggregation of production values of all crops masks systematic differences between

the types of crops that treatment and control households cultivate, as well as systematic variations in price, and may be less useful in evaluating the program's success.

In the survey, we distinguish between total agricultural production (and its value) and the amount of production sold in the market in order to differentiate the crops that are mainly used for households' own consumption. As we did with the value of production, we value the quantity sold with the median price for each crop in the zone of influence. We do not find significant effects in the value of sales. Table 30 estimates the impact on households' total agricultural income from sales using different pricing schemes and does not find significant effects. As we do find significant increases in production for vegetables and tubers, our results suggest that households may retain their production for household use or, in the case of corn, beans, and coffee (which store well), may intend to sell their output at a later time. As before, we note that these results do not say anything definitive about the welfare of treated households compared to the welfare of their counterparts in the control group.

Table 23 Total Production (Kgs.)

	(1)	(2)	(3)	(4)
Total Production (Kgs.)	All crops			
Treatment x 2013- Midline	326.4 [741.6]	371.2 [759.7]		
Treatment x 2015- Endline	3917.2 [1772.3]**	4021.5 [1827.4]**		
Treatment x Post- 2013 & 2015			2127.3 [1138.5]*	2224.5 [1185.7]*
Mean of Comp. at Baseline	1074.3	1073.5	1074.3	1073.5
SD of Comp. at Baseline	2066.7	2052.3	2066.7	2052.3
Observations	10325	10325	10325	10325
Production values with reported price - in 2005 PPP to 2010 USD				
Treatment x 2013- Midline	4136.6 [3496.6]	4122.2 [3418.4]		
Treatment x 2015- Endline	7246.1 [6151.2]	7223.4 [6036.3]		
Treatment x Post- 2013 & 2015			5706.9 [4846.0]	5699.1 [4758.9]
Mean of Comp. at Baseline	9282.9	9179	9282.9	9179
SD of Comp. at Baseline	175632	173230	175632	173230
Observations	8793	8793	8793	8793
Production values with median price - in 2005 PPP to 2010 USD				
Treatment x 2013- Midline	-1204.1 [929.1]	-1270.3 [988.6]		
Treatment x 2015- Endline	-1844.2 [2213.1]	-1972.1 [2365.8]		
Treatment x Post- 2013 & 2015			-1521.3 [1571.2]	-1620 [1686.0]
Mean of Comp. at Baseline	2509.4	2523.9	2509.4	2523.9
SD of Comp. at Baseline	6572.4	6613.5	6572.4	6613.5
Number of Clusters	285	285	285	285
Number of Households	2557	2557	2557	2557
Observations	10305	10305	10305	10305

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household-crop level and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 24 Total Production (Kgs) and Value: Corn, Beans, and Coffee

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Total Production (Kgs.)	Corn				Beans				Coffee			
Treatment x 2013- Midline	-86.9 [77.1]	-82.3 [76.3]			-49.8 [29.7]*	-49.1 [29.9]			-668.4 [614.3]	-706.1 [636.5]		
Treatment x 2015- Endline	62.7 [85.0]	61.7 [82.8]			-6.63 [34.7]	-8.22 [35.1]			-1695.1 [1469.2]	-1776.3 [1533.3]		
Treatment x Post- 2013 & 2015			-18.7 [72.9]	-16.8 [71.8]			-29 [29.6]	-29.5 [30.0]			-1178.8 [1082.4]	-1250.4 [1143.0]
Mean of Comp. at Baseline	843.1	832.4	843.1	832.4	185.8	185.2	185.8	185.2	1877	1889.6	1877	1889.6
SD of Comp. at Baseline	779.9	766.7	779.9	766.7	206.5	206.4	206.5	206.4	2516.7	2528.7	2516.7	2528.7
Observations	4292	4292	4292	4292	2020	2020	2020	2020	2641	2641	2641	2641
Production values with reported price - in 2005 PPP to 2010 USD												
Treatment x 2013- Midline	-123.4 [54.6]**	-120.4 [53.2]**			-124.9 [60.4]**	-122.8 [59.4]**			600.9 [1947.1]	736.8 [1976.0]		
Treatment x 2015- Endline	30.3 [70.2]	27.8 [68.5]			-4.55 [66.7]	-2.19 [65.8]			4041.7 [3061.4]	4307.1 [3252.5]		
Treatment x Post- 2013 & 2015			-50.2 [53.8]	-49.5 [52.2]			-67 [54.2]	-64.2 [53.3]			2294.3 [2384.3]	2532 [2534.1]
Mean of Comp. at Baseline	284.9	284.1	284.9	284.1	234.4	240.2	234.4	240.2	8536.6	8697	8536.6	8697
SD of Comp. at Baseline	682	678.3	682	678.3	455	458	455	458	12316.5	12462.7	12316.5	12462.7
Observations	3667	3667	3667	3667	1604	1604	1604	1604	2396	2396	2396	2396
Production values with median price - in 2005 PPP to 2010 USD												
Treatment x 2013- Midline	-53.8 [51.8]	-50.8 [51.5]			-113.7 [40.6]***	-113.6 [40.2]***			-2375.6 [2280.8]	-2502.5 [2351.1]		
Treatment x 2015- Endline	97.1 [61.3]	94.8 [59.9]			41.1 [59.5]	38.7 [59.2]			-6062.2 [5277.3]	-6339.2 [5499.5]		
Treatment x Post- 2013 & 2015			17.9 [50.2]	18.8 [49.7]			-39.3 [45.1]	-39.6 [44.8]			-4216.8 [3958.6]	-4464 [4170.6]
Mean of Comp. at Baseline	501.7	495.4	501.7	495.4	284.4	283.4	284.4	283.4	8198.3	8253	8198.3	8253
SD of Comp. at Baseline	464.1	456.3	464.1	456.3	316.5	316.2	316.5	316.2	10991.7	11044.4	10991.7	11044.4
Number of Clusters	272	272	272	272	225	225	225	225	208	208	208	208
Number of Households	2137	2137	2137	2137	1295	1295	1295	1295	1283	1283	1283	1283
Observations	4292	4292	4292	4292	2020	2020	2020	2020	2641	2641	2641	2641

Standard errors in brackets. Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household-crop level and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. * p<0.10, ** p<0.05, *** p<0.01

Table 25 Total Production (Kgs) and Value: Vegetables, Fruits, and Tubers

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Total Production (Kgs.)	Vegetables				Fruits				Tubers			
Treatment x 2013- Midline	-926.7 [4938.2]	153 [4869.1]			-63776.6 [37347.2]*	-65927.9 [39161.2]*			2532.6 [986.8]**	2579.2 [983.1]**		
Treatment x 2015- Endline	24852.3 [8588.2]***	26789.8 [9267.0]***			87657.7 [49271.4]*	88163.3 [49195.5]*			672.9 [1027.6]	669.7 [1016.2]		
Treatment x Post- 2013 & 2015			11887.6 [4712.6]**	13464.1 [5093.2]**			10987.5 [17176.6]	13342.2 [18406.4]			1728.2 [891.6]*	1747.6 [886.4]*
Mean of Comp. at Baseline	2210.1	2202.4	2210.1	2202.4	2754.9	2760	2754.9	2760	2694.3	2709.5	2694.3	2709.5
SD of Comp. at Baseline	4447.4	4432.4	4447.4	4432.4	11481.9	10952.7	11481.9	10952.7	4171	4168	4171	4168
Observations	522	522	522	522	208	208	208	208	452	452	452	452
Production values with reported price - in 2005 PPP to 2010 USD												
Treatment x 2013- Midline	21852.2 [16131.7]	21185.5 [15533.9]			-15045.3 [3853.9]***	-14986 [3823.8]***			-1940.1 [1375.8]	-1819.6 [1402.3]		
Treatment x 2015- Endline	-2715.5 [3688.5]	-2404.2 [3412.4]			5091.5 [2971.3]*	5281 [2981.2]*			-2307.3 [635.9]***	-2286.5 [647.4]***		
Treatment x Post- 2013 & 2015			2751.7 [3308.1]	2776.9 [3112.4]			375 [3752.8]	696.3 [3723.6]			-2114.1 [767.6]***	-2043.1 [780.5]**
Mean of Comp. at Baseline	352004	341605	352004	341605	151.2	146.4	151.2	146.4	2073.3	2084.8	2073.3	2084.8
SD of Comp. at Baseline	1283513	1262564	1283513	1262564	411.6	396	411.6	396	3056.9	3049.3	3056.9	3049.3
Observations	445	445	445	445	153	153	153	153	393	393	393	393
Production values with median price - in 2005 PPP to 2010 USD												
Treatment x 2013- Midline	-750.6 [1395.3]	-673.2 [1355.1]			-15510.7 [3595.8]***	-15734 [3612.2]***			2547.9 [704.4]***	2582.2 [698.5]***		
Treatment x 2015- Endline	1173.1 [659.6]*	1173.9 [648.8]*			629.5 [1623.8]	447.4 [1644.2]			880.3 [744.9]	873.5 [729.0]		
Treatment x Post- 2013 & 2015			761.4 [573.4]	787 [563.5]			-1013.6 [2096.3]	-1119.1 [2042.9]			1828 [622.2]***	1839.6 [613.1]***
Mean of Comp. at Baseline	940.8	942.9	940.8	942.9	371.8	367.3	371.8	367.3	1835.6	1842.9	1835.6	1842.9
SD of Comp. at Baseline	1623.1	1615.5	1623.1	1615.5	1272.7	1213.5	1272.7	1213.5	2791.9	2786.8	2791.9	2786.8
Number of Clusters	67	67	67	67	71	71	71	71	57	57	57	57
Number of Households	234	234	234	234	154	154	154	154	241	241	241	241
Observations	514	514	514	514	204	204	204	204	447	447	447	447

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level (Vegetables) and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. * p<0.10, ** p<0.05, *** p<0.01

Table 26: Log Total Production (kgs) and Values: Corn, Beans, and Coffee

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
aSinH-Log - Total Production (Kgs.)	Corn				Beans				Coffee			
Treatment x 2013- Midline	-0.26 [0.20]	-0.24 [0.20]			-0.4 [0.25]	-0.4 [0.25]			-0.25 [0.33]	-0.22 [0.33]		
Treatment x 2015- Endline	0.24 [0.24]	0.26 [0.24]			0.26 [0.26]	0.24 [0.26]			-0.12 [0.33]	-0.084 [0.34]		
Treatment x Post- 2013 & 2015			-0.029 [0.18]	-0.01 [0.18]			-0.085 [0.23]	-0.093 [0.23]			-0.19 [0.29]	-0.16 [0.29]
Mean of Comp. at Baseline	6.79	6.79	6.79	6.79	5.23	5.22	5.23	5.22	6.66	6.67	6.66	6.67
SD of Comp. at Baseline	1.79	1.76	1.79	1.76	1.52	1.52	1.52	1.52	2.76	2.76	2.76	2.76
Observations	4292	4292	4292	4292	2020	2020	2020	2020	2641	2641	2641	2641
aSinH-Log - Production values with reported price - in 2005 PPP to 2010												
Treatment x 2013- Midline	-0.63 [0.37]*	-0.62 [0.37]*			-0.89 [0.41]**	-0.91 [0.40]**			0.65 [0.73]	0.72 [0.75]		
Treatment x 2015- Endline	-0.42 [0.36]	-0.43 [0.35]			-1.27 [0.43]***	-1.26 [0.42]***			1.87 [1.43]	1.99 [1.51]		
Treatment x Post- 2013 & 2015			-0.53 [0.30]*	-0.53 [0.29]*			-1.07 [0.37]***	-1.07 [0.36]***			1.23 [0.98]	1.33 [1.05]
Mean of Comp. at Baseline	1.82	1.83	1.82	1.83	2.1	2.15	2.1	2.15	7.33	7.34	7.33	7.34
SD of Comp. at Baseline	3.24	3.24	3.24	3.24	3.3	3.32	3.3	3.32	3.84	3.86	3.84	3.86
Observations	3667	3667	3667	3667	1604	1604	1604	1604	2396	2396	2396	2396
aSinH-Log - Production values with median price - in 2005 PPP to 2010 US												
Treatment x 2013- Midline	-0.25 [0.19]	-0.23 [0.19]			-0.43 [0.26]*	-0.43 [0.26]*			-0.28 [0.38]	-0.25 [0.38]		
Treatment x 2015- Endline	0.24 [0.24]	0.26 [0.23]			0.29 [0.28]	0.28 [0.27]			-0.099 [0.38]	-0.049 [0.39]		
Treatment x Post- 2013 & 2015			-0.02 [0.18]	-0.0017 [0.17]			-0.084 [0.24]	-0.09 [0.24]			-0.2 [0.33]	-0.16 [0.34]
Mean of Comp. at Baseline	6.3	6.3	6.3	6.3	5.63	5.63	5.63	5.63	7.97	7.98	7.97	7.98
SD of Comp. at Baseline	1.68	1.66	1.68	1.66	1.59	1.58	1.59	1.58	3.18	3.17	3.18	3.17
Number of Clusters	272	272	272	272	225	225	225	225	208	208	208	208
Number of Households	2137	2137	2137	2137	1295	1295	1295	1295	1283	1283	1283	1283
Observations	4292	4292	4292	4292	2020	2020	2020	2020	2641	2641	2641	2641

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household-crop level and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. * p<0.10, ** p<0.05, *** p<0.01

Table 27: Log Total Production (kgs) and Values: Vegetables, Fruits, Tubers

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
aSinH-Log - Total Production (Kgs.)	Vegetables				Fruits				Tubers			
Treatment x 2013- Midline	-0.04 [1.15]	0.039 [1.19]			-3.8 [2.71]	-3.98 [2.71]			0.69 [0.61]	0.7 [0.63]		
Treatment x 2015- Endline	2.87 [1.25]**	2.96 [1.30]**			-1.66 [2.70]	-1.59 [2.67]			0.087 [0.46]	0.082 [0.46]		
Treatment x Post- 2013 & 2015			2.03 [1.05]*	2.12 [1.06]**			-5.04 [2.81]*	-4.93 [2.81]*			0.46 [0.47]	0.46 [0.48]
Mean of Comp. at Baseline	7.5	7.5	7.5	7.5	6.07	6.2	6.07	6.2	7.55	7.55	7.55	7.55
SD of Comp. at Baseline	1.73	1.72	1.73	1.72	2.28	2.3	2.28	2.3	2.16	2.19	2.16	2.19
Observations	522	522	522	522	208	208	208	208	452	452	452	452
aSinH-Log - Production values with reported price - in 2005 PPP to 2010												
Treatment x 2013- Midline	5 [2.24]**	4.82 [2.22]**			-1.36 [3.66]	-1.5 [3.58]			-1.96 [2.32]	-1.85 [2.35]		
Treatment x 2015- Endline	1.57 [2.10]	1.62 [2.13]			-0.61 [2.08]	-0.47 [2.00]			-0.84 [1.48]	-0.84 [1.47]		
Treatment x Post- 2013 & 2015			2.13 [1.84]	2.11 [1.86]			-2.25 [1.85]	-2.15 [1.82]			-1.41 [0.79]*	-1.36 [0.80]*
Mean of Comp. at Baseline	6.49	6.45	6.49	6.45	3.21	3.23	3.21	3.23	7.35	7.36	7.35	7.36
SD of Comp. at Baseline	3.85	3.83	3.85	3.83	3.12	3.07	3.12	3.07	1.93	1.92	1.93	1.92
Observations	445	445	445	445	153	153	153	153	393	393	393	393
aSinH-Log - Production values with median price - in 2005 PPP to 2010 US												
Treatment x 2013- Midline	0.41 [1.36]	0.49 [1.39]			-4.81 [3.24]	-5.05 [3.25]			0.66 [0.60]	0.67 [0.62]		
Treatment x 2015- Endline	2.38 [0.84]***	2.45 [0.86]***			-1.27 [2.89]	-1.35 [2.84]			0.048 [0.45]	0.043 [0.45]		
Treatment x Post- 2013 & 2015			2.48 [1.02]**	2.55 [1.00]**			-3.2 [2.78]	-3.25 [2.75]			0.43 [0.47]	0.43 [0.48]
Mean of Comp. at Baseline	6.67	6.67	6.67	6.67	4.17	4.29	4.17	4.29	7.19	7.19	7.19	7.19
SD of Comp. at Baseline	1.65	1.64	1.65	1.64	2.61	2.6	2.61	2.6	2.12	2.14	2.12	2.14
Number of Clusters	67	67	67	67	71	71	71	71	57	57	57	57
Number of Households	234	234	234	234	154	154	154	154	241	241	241	241
Observations	514	514	514	514	204	204	204	204	447	447	447	447

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level (Vegetables) and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 28 Value of households' agricultural production

	(1)	(2)	(3)	(4)
Production values with reported price - in 2005 PPP to 2010 USD				
Treatment=1 # 2013- Midline	10593.4 [10325.3]	10455.7 [10084.2]		
Treatment=1 # 2015- Endline	11204.7 [9749.9]	11036.1 [9398.8]		
Treatment=1 # Post- 2013 & 2015=1			10916.5 [10010.0]	10766.8 [9704.6]
Mean of Comp. at Baseline	11881.4	11489.1	11881.4	11489.1
SD of Comp. at Baseline	198692.2	193803.6	198692.2	193803.6
Observations	6267	6267	6267	6267
Production values with median price - in 2005 PPP to 2010 USD				
Treatment=1 # 2013- Midline	-1481 [1250.7]	-1482.3 [1273.2]		
Treatment=1 # 2015- Endline	-1242.7 [2333.8]	-1201.9 [2461.1]		
Treatment=1 # Post- 2013 & 2015=1			-1359.1 [1708.2]	-1336.7 [1802.4]
Mean of Comp. at Baseline	4353.7	4343.3	4353.7	4343.3
SD of Comp. at Baseline	8400.3	8423.7	8400.3	8423.7
Observations	6267	6267	6267	6267
Production values with department median price - in 2005 PPP to 2010 USD				
Treatment=1 # 2013- Midline	-716.2 [923.0]	-700.3 [954.4]		
Treatment=1 # 2015- Endline	-649.5 [1940.6]	-525.4 [2049.3]		
Treatment=1 # Post- 2013 & 2015=1			-686.5 [1412.5]	-612.9 [1493.2]
Mean of Comp. at Baseline	4439.6	4443.1	4439.6	4443.1
SD of Comp. at Baseline	9084.9	9125.9	9084.9	9125.9
Number of Clusters	286	286	286	286
Number of Households	2675	2675	2675	2675
Observations	6267	6267	6267	6267

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 29 Agriculture Sales revenue in 2005PPP to 2010 USD

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	Corn				Beans				Coffee			
Sales revenue with reported price - in 2005 PPP to 2010 USD												
Treatment x 2013- Midline	-130.1	-127.2			-136.3	-131.7			976.3	1111.4		
	[67.6]*	[66.2]*			[88.2]	[86.7]			[1970.8]	[2006.4]		
Treatment x 2015- Endline	40.8	39.8			-88	-78.9			4270.4	4542.4		
	[91.0]	[90.3]			[155.0]	[153.3]			[3082.6]	[3273.2]		
Treatment x Post- 2013 & 2015			-47.9	-46.2			-92	-86			2582.4	2819.5
			[72.2]	[71.3]			[108.5]	[106.7]			[2396.5]	[2546.3]
Mean of Comp. at Baseline	335.9	337.7	335.9	337.7	306.1	316.3	306.1	316.3	8897.2	9070.2	8897.2	9070.2
SD of Comp. at Baseline	526.6	531	526.6	531	378.4	382	378.4	382	13146.6	13318.1	13146.6	13318.1
Observations	1303	1303	1303	1303	647	647	647	647	2303	2303	2303	2303
Sales revenue with median price - in 2005 PPP to 2010 USD												
Treatment x 2013- Midline	-117.5	-112.1			-122.3	-120.1			-2355.9	-2469.1		
	[72.1]	[70.4]			[88.8]	[86.7]			[2469.7]	[2540.6]		
Treatment x 2015- Endline	57.9	57.8			-102.2	-100.5			-6028	-6320.6		
	[111.4]	[108.5]			[152.6]	[148.1]			[5553.9]	[5777.1]		
Treatment x Post- 2013 & 2015			-31.3	-27.9			-90.3	-88.6			-4174.7	-4424.9
			[78.1]	[76.4]			[101.9]	[99.4]			[4190.3]	[4406.1]
Mean of Comp. at Baseline	311.9	311.9	311.9	311.9	274.8	284.2	274.8	284.2	8634	8706.4	8634	8706.4
SD of Comp. at Baseline	487.4	492.3	487.4	492.3	344.1	346.8	344.1	346.8	12067.4	12161.9	12067.4	12161.9
Number of Clusters	160	160	160	160	118	118	118	118	195	195	195	195
Number of Households	629	629	629	629	379	379	379	379	1189	1189	1189	1189
Observations	1426	1426	1426	1426	694	694	694	694	2504	2504	2504	2504

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household-crop level and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 30 Ag. Sales revenue household aggregate impact with alternative prices

	(1)	(2)	(3)	(4)
Sales revenue with reported price - in 2005 PPP to 2010 USD				
Treatment=1 # 2013- Midline	7844.3 [10691.3]	7867.2 [10416.3]		
Treatment=1 # 2015- Endline	8404.6 [10125.1]	8425.8 [9736.6]		
Treatment=1 # Post- 2013 & 2015=1			8139.7 [10381.8]	8165.9 [10040.7]
Mean of Comp. at Baseline	11823.6	11438.1	11823.6	11438.1
SD of Comp. at Baseline	198710.9	193823.4	198710.9	193823.4
Observations	6267	6267	6267	6267
Sales revenue with median price - in 2005 PPP to 2010 USD				
Treatment=1 # 2013- Midline	-3610.2 [3123.9]	-3458.6 [2989.3]		
Treatment=1 # 2015- Endline	-4720.2 [3561.9]	-4571.1 [3493.2]		
Treatment=1 # Post- 2013 & 2015=1			-4201.2 [3311.9]	-4060.1 [3213.6]
Mean of Comp. at Baseline	3953.9	3957.5	3953.9	3957.5
SD of Comp. at Baseline	8980	9026.4	8980	9026.4
Observations	6267	6267	6267	6267
Sales revenue with department median price - in 2005 PPP to 2010 USD				
Treatment=1 # 2013- Midline	-3503.9 [3095.8]	-3311.7 [2952.6]		
Treatment=1 # 2015- Endline	-4029.6 [3340.6]	-3797.2 [3241.9]		
Treatment=1 # Post- 2013 & 2015=1			-3786.1 [3192.1]	-3577.1 [3071.9]
Mean of Comp. at Baseline	4006	4019.9	4006	4019.9
SD of Comp. at Baseline	9329.5	9368.3	9329.5	9368.3
Number of Clusters	286	286	286	286
Number of Households	2675	2675	2675	2675
Observations	6267	6267	6267	6267

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

To summarize our results for agricultural productivity and sales:

- Production in KG
 - ~2000 kg increase in production overall, driven by ~4000 kg increase at endline;
 - Increase in production of vegetables and tubers;
 - Decrease in production of beans; and
 - In general, significant effects for high value crops appear only at endline, which reflects ACCESO's prioritization of food security over high-value crop cultivation.
- Value of Production
 - Decrease in midline value of production of corn and beans; effect for beans robust to use of median price;
 - Results for corn, beans, coffee driven by wide variations in reported price; effects not robust to inverse hyperbolic sine transformation (to correct for over-dispersion) and use of department median prices;
 - 2% increase in quantity of vegetables produced and 2.5% increase in value of vegetables produced after correcting for over-dispersion and when using median price; and
 - Treatment farmers report lower prices for tubers but see larger output, likely driven by outliers as this finding is not robust to corrections for over-dispersion.
- No aggregate effect on sales value, regardless of whether we use reported price, median price, or department median price.

Land Productivity or Crop Productivity

We use the estimated production and the area planted discussed in the previous sub-sections to estimate the productivity of land for the households in the sample. We measure the productivity of land in kilograms per hectares (Kg/ha) for each major crop category individually. For corn, beans, and coffee, we estimate the yield (Kg/ha), imputing⁸ and trimming productivity levels that fall outside of the levels that one would expect when using good practices and technologies. The levels used are: 3,584 Kgs. /ha for beans, 10,426 kgs. /ha for corn, and 2,997 Kgs./ha for coffee⁹. Table 31 shows the results for the productivity measure. We find consistent significant negative impacts estimates on yields at midline,

⁸ The values outside the range given were imputed with the top of the range. In the case of trimming, the values was defined as missing and not included in the calculation.

⁹ We use the limits described in the Fintrac Technical Proposal for Corn and Beans (slightly higher) and the high maximum yield for Arabica coffee given by the International Coffee Organization (http://www.ico.org/botanical.asp?section=About_Coffee).

suggesting that yields were worse in the treatment group than the control group; the significant effects are only at the 10 percent level. We see no difference between treatment and comparison households at endline of these crops. These results are similar to other studies which suggest that it takes time for farmers to see benefits to technology adoption. This may be for several reasons. First, adopting new practices may not occur smoothly; farmers may imperfectly adapt new techniques to their plots and thus see lower returns, even though there may be significant gains from learning-by-doing in subsequent seasons. Second, early adopters may be significantly different from the average farmer, in ways that are correlated with worse initial outcomes when adopting new technology. As technology diffuses to households closer to the average farmer, average outcomes will improve. Lastly, we see that treatment farmers significantly expanded the area planted for coffee and corn and detect insignificant increases for other crops. It may be that the newly sown plots were of lower quality in ways that affect land productivity (i.e. further away from a water source or unirrigated, lower soil quality, or less sun exposure, among other factors). Lower quality plots would be less productive on average, even though farmers may have been able to increase their total production quantities or values and see positive returns to their expansion.

Table 31 Land productivity: Yield (Kg/Ha) for Corn, Beans, Coffee

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Yield (Kg/Ha)	Corn				Beans				Coffee			
Treatment x 2013- Midline	-504.8 [389.3]	-472.5 [367.1]			-505.9 [256.9]*	-476.9 [246.2]*			-1255.7 [516.6]**	-1296.1 [528.4]**		
Treatment x 2015- Endline	-750.6 [788.4]	-688 [734.7]			-67.1 [197.9]	-86.5 [203.8]			-1827.6 [1307.7]	-1913 [1365.8]		
Treatment x Post- 2013 & 2015			-627.2 [571.8]	-582.3 [535.7]			-310.7 [165.4]*	-305.5 [163.6]*			-1571 [932.2]*	-1644.5 [986.5]*
Mean of Comp. at Baseline	1661.8	1682.6	1661.8	1682.6	721.5	722.6	721.5	722.6	1245.9	1244.9	1245.9	1244.9
SD of Comp. at Baseline	2316.9	2402.9	2316.9	2402.9	1771.8	1730.2	1771.8	1730.2	1806.7	1855.2	1806.7	1855.2
Observations	3960	3960	3960	3960	1824	1824	1824	1824	2289	2289	2289	2289
Yield Kg/Ha Med. Impute-Out of range												
Treatment x 2013- Midline	-214.9 [117.5]*	-208.1 [116.3]*			-288.9 [141.2]**	-284.2 [141.1]**			-0.36 [75.6]	-6.1 [76.3]		
Treatment x 2015- Endline	14.2 [134.9]	23 [136.0]			-158 [128.7]	-157.8 [128.8]			-97.1 [107.1]	-103 [109.1]		
Treatment x Post- 2013 & 2015			-109.9 [106.5]	-102 [106.7]			-229.4 [128.6]*	-226.8 [128.7]*			-48.6 [87.3]	-55.3 [89.5]
Mean of Comp. at Baseline	1478.5	1479.8	1478.5	1479.8	621.8	626.9	621.8	626.9	899.7	887.1	899.7	887.1
SD of Comp. at Baseline	1351.5	1355.7	1351.5	1355.7	592	592.3	592	592.3	691.3	687.6	691.3	687.6
Observations	3961	3961	3961	3961	1825	1825	1825	1825	2289	2289	2289	2289
Yield Kg/Ha trimmed-Out of range												
Treatment x 2013- Midline	-194.8 [117.4]*	-187.6 [116.4]			-295.9 [139.6]**	-291.3 [139.5]**			120.1 [67.7]*	116.7 [68.2]*		
Treatment x 2015- Endline	31.7 [131.4]	40.7 [132.6]			-155.8 [130.5]	-155.7 [130.7]			0.62 [86.5]	-3.47 [86.8]		
Treatment x Post- 2013 & 2015			-90.1 [105.0]	-81.9 [105.3]			-232.2 [128.0]*	-229.6 [128.2]*			64.9 [65.9]	60.9 [66.4]
Mean of Comp. at Baseline	1480.4	1481.9	1480.4	1481.9	623.1	628.2	623.1	628.2	898.7	885.2	898.7	885.2
SD of Comp. at Baseline	1358.7	1363.6	1358.7	1363.6	595.6	595.8	595.6	595.8	716.6	713.1	716.6	713.1
Number of Clusters	271	271	271	271	222	222	222	222	204	204	204	204
Number of Households	2099	2099	2099	2099	1222	1222	1222	1222	1180	1180	1180	1180
Observations	3909	3909	3909	3909	1767	1767	1767	1767	2009	2009	2009	2009

Standard errors in brackets. Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household-crop level and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

To summarize the effects for land productivity:

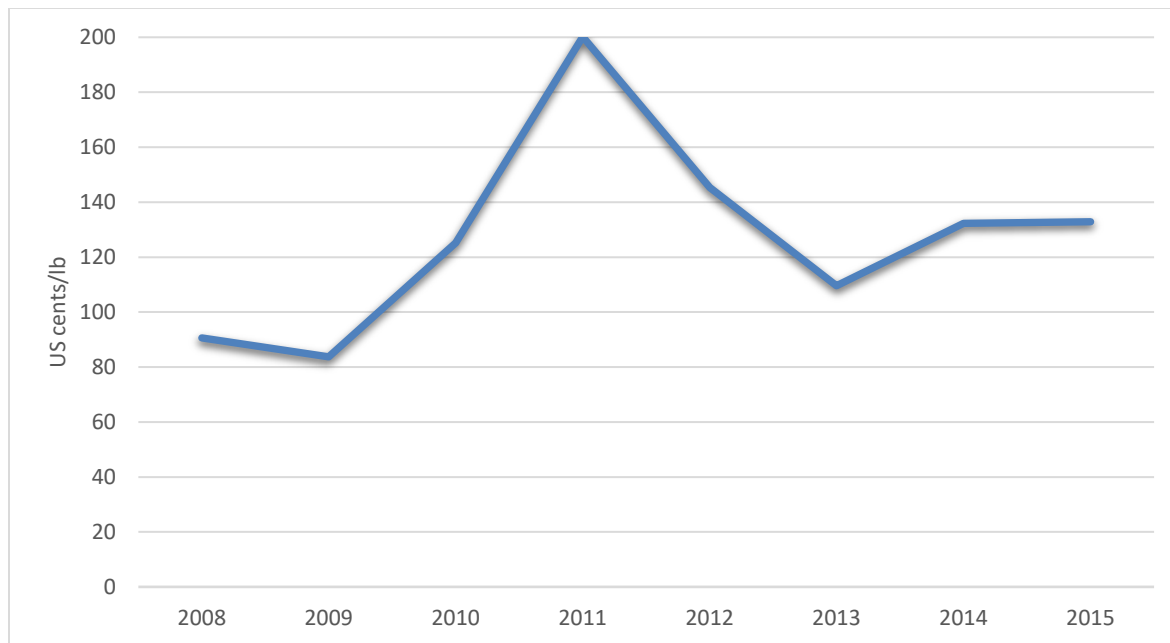
- We see a drop in the midline yield of beans, which is robust to imputing or excluding outliers
 - Note that effects are not precisely measured when looking at the “post” period.

Coffee rust

Coffee rust is a disease caused by the *Hemileia vastatrix* fungus, which attacks the coffee bush and causes defoliation that significantly decreases the production of grain and leads to the death of the plant. The disease first appeared in the Western Hemisphere in 1970 in Brazil and has quickly spread throughout Central and South America. The coffee rust fungus spreads by wind and can be controlled with the use of fungicides.

The rust devastated coffee production in Central America in 2013, causing a 30% decline in coffee production in the 2013-2014 season compared to the 2011-2012 season [(FEWS NET, February 2014)]. This posed a significant problem for Honduras’s economy, as coffee represents an important portion of the country’s export revenues (around 30% of GDP). In addition, the reduction in coffee production affected labor demand and the wages received by coffee day laborers and decreased the household income of coffee producers (mostly small farmers), endangering the food security of many vulnerable workers and farmers. This represents a significant change in the income dynamic for many households in the area, which have cut their revenue and maintenance of coffee in order to supplement their income and provide basic food during the lean season. The vulnerability of coffee-producing households is exacerbated by the lack of other sources of employment and income, as well as by a fall in coffee prices in the international market, as can be seen in Figure 6. Coffee prices experienced a sharp decline in 2013 and only started to recuperate after November 2013. This increase was short-lived, however; the price of coffee has again been in decline since the second quarter of 2014, arriving at prices at the end of 2015 similar to those seen in 2013.

Figure 6 Prices paid to growers in exporting countries, Honduras, 2008-2015 (US cents/lb.)



Source: International Coffee Organization ¹⁰

Over 40% of the households in the sample that participate in agricultural activities grew coffee at baseline. Given the potential for coffee rust to affect agriculture in the zone of influence, for the follow-up surveys, we expanded the survey instrument to characterize the situation of coffee producers affected by rust in the 2013-2014 season. We do not have baseline data for these measures, and thus we present a descriptive analysis of how coffee farmers in the sample were affected by the coffee rust and the measures that they took to mitigate the effects. To expand on this analysis, we also present cross-sectional impact estimates for some measures by pooling the follow-up surveys and estimating the difference between treatment and control coffee growers after controlling for the year of the survey. The purpose is to gauge if the treatment group was more successful in mitigating the effects of coffee rust than the comparison group. We compare outcomes using the propensity score amounts to assume that after accounting for observables included in the propensity scores, unobservables are not differentially related to the outcomes related to coffee rust.

Table 32 shows the descriptive statistics for coffee growers in the sample. These estimates suggest that around 38% of the coffee area in the zone of influence was affected by coffee rust and that the average area affected was 1.4 hectares per farmer. Figure 7 shows the proportion of the production that was

¹⁰ http://www.ico.org/new_historical.asp

lost due to coffee rust. In the majority of cases, farmers lost over 75% of their expected production. On average, coffee growers in the treatment group seem to be more knowledgeable about the disease, with 66% reporting that they believe that coffee rust can be prevented and reporting more practices that can be used to prevent and control coffee rust. Treatment households also received more assistance on average and had a smaller proportion and a smaller extent of their farms affected by the rust.

To explore if these differences are systematic between treatment and control after constructing a counterfactual based on the propensity score, we proceed to estimate a cross-section regression using the propensity scores and controlling for the survey year.

Table 32 Descriptive Statistics for Coffee Rust: Belief, Prevention, and Control

	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
	Comparison			Treated			Total		
	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs
Believes Coffee Rust can be prevented	0.63	0.48	811	0.66	0.48	929	0.64	0.48	1740
Number of Ways Farmer knows to prevent Coffee Rust	1.53	0.85	512	1.67	1.09	610	1.60	0.99	1122
Affected by Coffee Rust in 2012-2013 Season	0.68	0.47	811	0.65	0.48	929	0.67	0.47	1740
Farm received any assistance with coffee rust	0.07	0.26	551	0.17	0.37	607	0.12	0.33	1158
Number of Ways Farmer knows to control Coffee Rust	0.90	0.67	627	0.91	0.68	698	0.91	0.67	1325
Average area per household affected by Coffee Rust (Ha.)	1.81	15.72	544	1.05	3.37	596	1.41	11.13	1140
Proportion Coffee area affected by Rust	0.41	0.40	745	0.36	0.38	857	0.38	0.39	1602
Cost of inputs associated with rust (2005PP/10 USD)	116.38	395.90	531	122.02	456.44	578	119.32	428.34	1109
Cost of labor associated with rust, Maintenance + Cut (2005PP/2010 USD)	146.08	541.70	533	245.24	799.26	575	197.54	689.00	1108
Total Costs associate with rust, Labor+Inputs (2005PP/10 USD)	259.58	761.76	538	363.47	1101.00	582	313.57	954.23	1120
Proportion Coffee production lost	0.60	0.33	552	0.56	0.33	606	0.58	0.33	1158

Figure 7 Proportion of Production Affected by coffee rust

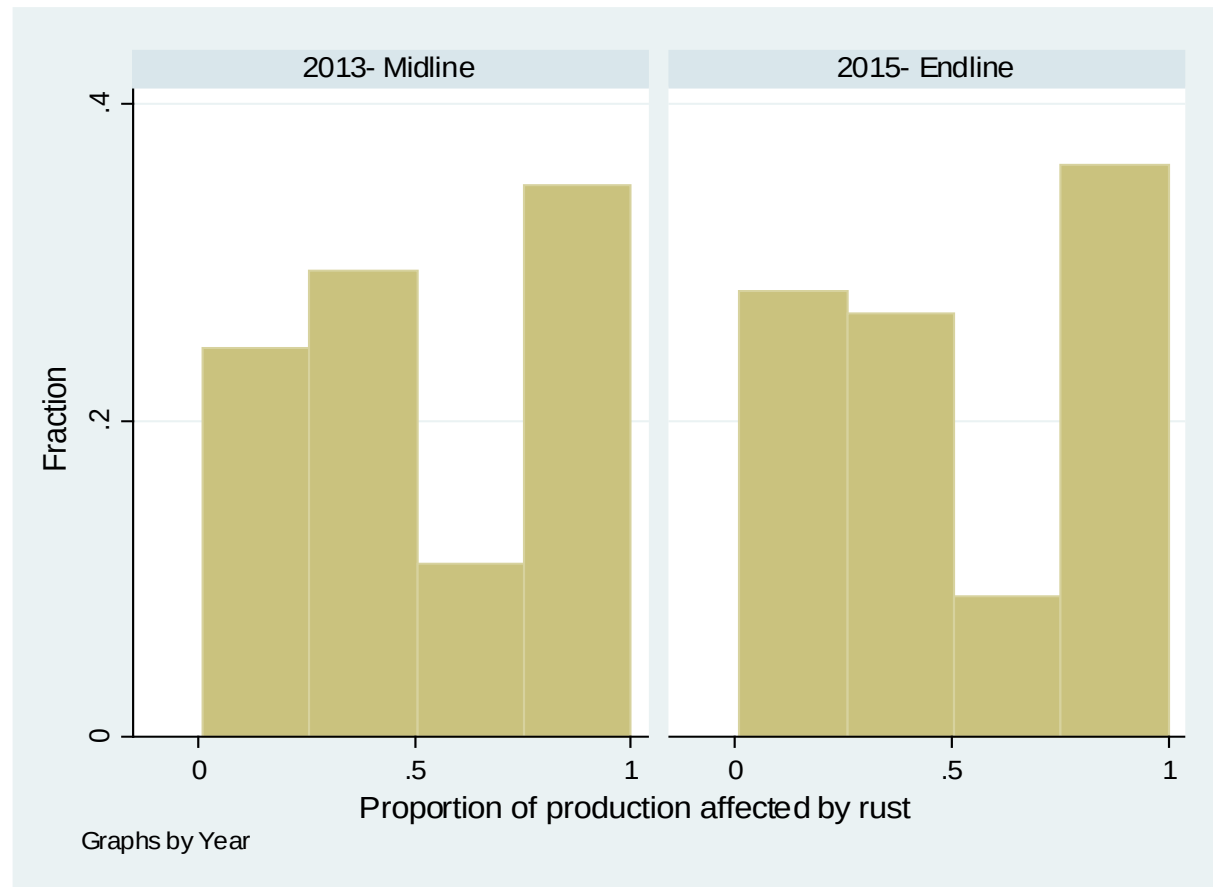


Table 33 confirms that coffee growers in the treatment group are marginally more knowledgeable about coffee rust control and that treatment farmers are 9.5 percentage points more likely to have received technical assistance to handle coffee rust than the comparison group. The estimate on the endline indicators suggest that the situation improved in the 2014-2015 season; both groups increased their knowledge regarding how to prevent coffee rust, and a smaller proportion of cultivation was affected.

The coefficient “2015-Endline” in Table 34 shows that both groups were less affected by the rust in the 2014-2015 season as a proportion of their total coffee area; in the second panel, the effect on the percentage of coffee-growing land is negative and significant. However, treatment households do not fare better than comparison households in the proportion of area affected. The table also explores whether there are differences in the inputs and labor costs employed by farmers to mitigate the effects of the disease. The estimates suggest that both groups increased their expenditures to mitigate the effects of the rust and that the treatment households spent less, although this result is not significant at conventional levels. Overall, we find that treatment and control households face similar experiences with coffee rust, although treatment households appear to spend marginally less on inputs associated with coffee rust and receive more assistance in mitigating the effects of coffee rusts. Combined with our earlier results, which show that treated households that farm coffee spend more on labor and chemical inputs compared to control households, we surmise that USAID-ACCESO gave farmers the resources they needed to mitigate losses from coffee rust.

Table 33 Coffee Rust: Impact on Assistance and Practices

	(1)	(2)
Farmer believes coffee rust can be prevented		
Treatment	-0.015 [0.026]	-0.014 [0.026]
2015- Endline	0.19 [0.032]***	0.19 [0.033]***
Mean of Comp. at Baseline	0.56	0.56
SD of Comp. at Baseline	0.5	0.5
Observations	1740	1740
Number of Ways Farmer knows to prevent Coffee Rust		
Treatment	0.039 [0.12]	0.034 [0.13]
2015- Endline	0.31 [0.10]***	0.33 [0.10]***
Mean of Comp. at Baseline	1.42	1.42
SD of Comp. at Baseline	0.8	0.8
Observations	1122	1122
Farmer Affected by Coffee Rust in Season of Survey		
Treatment	-0.045 [0.037]	-0.047 [0.038]
2015- Endline	-0.15 [0.031]***	-0.15 [0.031]***
Mean of Comp. at Baseline	0.78	0.78
SD of Comp. at Baseline	0.41	0.42
Observations	1740	1740
Farm received any assistance with coffee rust		
Treatment	0.095 [0.024]***	0.094 [0.023]***
2015- Endline	-0.0063 [0.025]	-0.0098 [0.025]
Mean of Comp. at Baseline	0.081	0.081
SD of Comp. at Baseline	0.27	0.27
Observations	1158	1158
Number of Ways farmer knows to control Coffee Rust		
Treatment	-0.14 [0.079]*	-0.15 [0.082]*
2015- Endline	0.35 [0.069]***	0.35 [0.071]***
Mean of Comp. at Baseline	0.83	0.82
SD of Comp. at Baseline	0.68	0.68
Number of Clusters	189	189
Number of Households	985	985
Observations	1325	1325

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and impact estimation strategy consists of single difference controlling for survey year. Column (1) presents the PSM weighted impact estimates and columns (2) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 34 Coffee Rust: Impact on Area, Production, and Costs

	(1)	(2)
Area affected by Coffee Rust (Ha.)		
Treatment	-3.65 [3.38]	-3.48 [3.20]
2015- Endline	-2.99 [2.98]	-2.86 [2.86]
Mean of Comp. at Baseline	7.78	7.5
SD of Comp. at Baseline	47.1	46.2
Observations	1140	1140
% of coffee-growing land affected by coffee rust		
Treatment	-0.023 [0.025]	-0.023 [0.025]
2015- Endline	-0.11 [0.022]***	-0.11 [0.022]***
Mean of Comp. at Baseline	0.45	0.45
SD of Comp. at Baseline	0.38	0.38
Observations	1602	1602
Cost of inputs associated with rust - 2005 PPP to 2010 USD		
Treatment	-127.1 [74.8]*	-127.5 [72.5]*
2015- Endline	148.4 [70.4]**	146.2 [67.0]**
Mean of Comp. at Baseline	126	124.7
SD of Comp. at Baseline	398.6	392.8
Observations	1109	1109
Total Cost of labor associated with rust, Maintenance. + Cut - 2005 PPP to 2010 USD		
Treatment	-242.1 [260.2]	-267.5 [277.8]
2015- Endline	384.7 [257.7]	400.9 [270.8]
Mean of Comp. at Baseline	168.1	166.5
SD of Comp. at Baseline	665.8	658.3
Observations	1108	1108
Total costs associate with rust (Labor+Inputs) - 2005 PPP to 2010 USD		
Treatment	-364.8 [303.7]	-389.8 [321.1]
2015- Endline	529.7 [294.4]*	543.2 [306.8]*
Mean of Comp. at Baseline	290.2	287.3
SD of Comp. at Baseline	807.9	798.7
Observations	1120	1120
Proportion of production lost		
Treatment	0.015 [0.045]	0.016 [0.047]
2015- Endline	-0.046 [0.042]	-0.047 [0.044]
Mean of Comp. at Baseline	0.58	0.58
SD of Comp. at Baseline	0.32	0.32
Number of Clusters	183	183
Number of Households	907	907
Observations	1158	1158

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and impact estimation strategy consists of single difference controlling for survey year. Column (1) presents the PSM weighted impact estimates and columns (2) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

To summarize our findings on coffee rust:

- Treated households are about 10 percentage points more likely to receive assistance in dealing with coffee rust; and
- Both treatment and control households are less affected by coffee rust by endline.

4.2 ECONOMIC WELL-BEING

Through USAID-ACCESO activities, USAID and Feed the Future expected to increase household income in order to directly lift beneficiary households out of poverty and/or extreme poverty. In this impact evaluation, household expenditure was used as a proxy for income and is the principal indicator of household welfare. We also measure income, but we focus on aspects of income that we believe would be most directly related to agricultural practices and rural labor markets.

We present the impact estimates on household income from different sources and expenditures across different categories, including food and non-food expenditures. This section presents the impact estimates using the household-level data and the households fixed effects with propensity score weights discussed in the methodology section. The section on income aggregates different sources from which households derive their income, e.g., from agricultural, salary, and entrepreneurial income. As is well known, income measures from surveys tend to be understated and the consensus is that they are unreliable; thus, we focus on households' consumption expenditures, which more realistically reflect economic well-being.

Income and Labor Supply

To determine the labor market and transfer dynamics in these households, we construct a measure of income that incorporates the main categories from which households may derive a regular flow of income. This includes income from salaries, independent work/businesses, and agricultural production. We define total household income as the income from the three main labor activities of every member of the household between the ages of 15 and 65. To explore if the program affected income from different activities, we estimate the impact by type of activity (dependent or independent) and by sector (agricultural, commerce, professional, etc.).

We first start by exploring the impacts on the income from primary labor activities. Table 35 shows the impact on the likelihood that someone in the household is employed in specific activities. ACCESO encouraged households to diversify their income sources in a variety of ways; the most common method was to encourage household food production (milk and eggs), the excess of which would be sold to generate income. As such, we estimate the likelihood of participating in: wage labor,

commerce/market activities, agriculture, and professional services. The results in Table 35 are consistent with the program's design and suggest that treatment households are over 11 percentage points more likely to have members who work for wages and 9 percentage points more likely to have their own agriculture production activities. We find no significant impacts on the probability that treatment households have members engaged in commerce (small shops, etc.) or professional work, which is unsurprising, as USAID-ACCESO did not actively promote either.

Table 35 Labor market participation

	(1)	(2)	(3)	(4)
Participation 1 - Wage-Labor				
Treatment x 2013- Midline	0.1 [0.076]	0.1 [0.074]		
Treatment x 2015- Endline	0.12 [0.066]*	0.11 [0.065]*		
Treatment x Post- 2013 & 2015			0.11 [0.054]**	0.11 [0.053]**
Mean of Comp. at Baseline	0.83	0.83	0.83	0.83
SD of Comp. at Baseline	0.96	0.95	0.96	0.95
Observations	5086	5086	5086	5086
Participation 1 - Commerce				
Treatment x 2013- Midline	-0.068 [0.11]	-0.068 [0.11]		
Treatment x 2015- Endline	0.09 [0.096]	0.094 [0.097]		
Treatment x Post- 2013 & 2015			0.0098 [0.090]	0.013 [0.091]
Mean of Comp. at Baseline	0.51	0.52	0.51	0.52
SD of Comp. at Baseline	0.57	0.58	0.57	0.58
Observations	1377	1377	1377	1377
Participation 1 - Agriculture				
Treatment x 2013- Midline	0.054 [0.051]	0.051 [0.051]		
Treatment x 2015- Endline	0.12 [0.054]**	0.12 [0.054]**		
Treatment x Post- 2013 & 2015			0.09 [0.044]**	0.089 [0.044]**
Mean of Comp. at Baseline	1.25	1.25	1.25	1.25
SD of Comp. at Baseline	1.03	1.03	1.03	1.03
Observations	6877	6877	6877	6877
Participation 1 - Professional				
Treatment x 2013- Midline	-0.16 [0.084]*	-0.16 [0.084]*		
Treatment x 2015- Endline	-0.086 [0.11]	-0.087 [0.11]		
Treatment x Post- 2013 & 2015			-0.12 [0.086]	-0.12 [0.086]
Mean of Comp. at Baseline	0.61	0.61	0.61	0.61
SD of Comp. at Baseline	0.69	0.68	0.69	0.68
Number of Clusters	158	158	158	158
Number of Households	601	601	601	601

Observations	1625	1625	1625	1625
Standard errors in brackets				
Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects, and controls for the number of members age 15 to 65 and the household size. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.				
* p<0.10, ** p<0.05, *** p<0.01				

In Table 36, we estimate the impact on the income from each of the aforementioned activities. We find differential increases in the treatment group for the endline survey. Treatment households have more income from wage labor (\$387 dollars per household in the past year), or 17% of the average wage income of the comparison at baseline; this is likely due to the increased labor force participation we observe in Table 35 rather than to higher wages due to higher productivity. For other categories of income, we do not find statistically significant differences. In the case of agriculture, these two results suggest that more people seem to participate in the household's independent agriculture activities, but there was no significant increase in the income from household farms.

Table 36 Impacts on Income by Source

	(1)	(2)	(3)	(4)
Annual income for -Wage-Labor - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	203.5 [209.1]	217.4 [205.5]		
Treatment x 2015- Endline	391.8 [192.6]**	386.9 [193.5]**		
Treatment x Post- 2013 & 2015			298.9 [162.2]*	304.7 [161.4]*
Mean of Comp. at Baseline	2264.5	2287.8	2264.5	2287.8
SD of Comp. at Baseline	3520.1	3515.2	3520.1	3515.2
Observations	5686	5686	5686	5686
Annual income for -Commerce - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	1534 [953.3]	1549.3 [939.0]		
Treatment x 2015- Endline	379.5 [719.3]	372.9 [730.1]		
Treatment x Post- 2013 & 2015			959.5 [714.8]	957.4 [713.7]
Mean of Comp. at Baseline	1886.1	1902.7	1886.1	1902.7
SD of Comp. at Baseline	6031.5	6006.4	6031.5	6006.4
Observations	1531	1531	1531	1531
Annual income for -Agriculture - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	20.9 [864.3]	152.9 [860.5]		
Treatment x 2015- Endline	2064.3 [1534.4]	2205.4 [1505.3]		
Treatment x Post- 2013 & 2015			1045 [1159.0]	1194.1 [1144.7]
Mean of Comp. at Baseline	7104.5	7425.5	7104.5	7425.5
SD of Comp. at Baseline	20139.6	20679.1	20139.6	20679.1
Observations	6345	6345	6345	6345
Annual income for -Professional - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	192.6 [607.9]	271 [629.7]		
Treatment x 2015- Endline	192.6 [661.0]	246.6 [658.9]		
Treatment x Post- 2013 & 2015			181 [566.7]	242.6 [575.2]
Mean of Comp. at Baseline	3482	3569.4	3482	3569.4
SD of Comp. at Baseline	9225.8	9268.8	9225.8	9268.8
Number of Clusters	165	165	165	165
Number of Households	636	636	636	636
Observations	1720	1720	1720	1720

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003).

Observations are at the household level and all regressions include household fixed effects, and controls for the number of members age 15 to 65 and the household size. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Next we aggregate the income of the primary activity of each able member of the household depending on whether they work for others (dependent) or independently. Table 37 shows the impact on annual household income and the costs derived from dependent and independent labor activities. We find positive estimates for income from wages but not significant. The effects on independent income are not significant; however, the sign of the impact estimate for 2013 is negative, suggesting depressed income in general in 2013. We do not find that treatment households invested more in their independent business when compared to similar households in the comparison group Table 38 shows the impact of the program when we aggregate earnings and business investment costs for up to three labor activities per member of the household. We find no differences between treatment and control for the years after the baseline.

To summarize our results for household labor activities, we find that the program led to:

- An 11 percentage point increase in wage labor participation;
- A 9 percentage point increase in participation in agriculture;
- An imprecisely estimated \$300 increase in come from wage labor; and

Table 37 Impact of Labor Income from primary activity

	(1)	(2)	(3)	(4)
Annual wages from dependent - Job1 - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	78 [214.0]	75.4 [215.2]		
Treatment x 2015- Endline	174.7 [281.3]	156.5 [281.1]		
Treatment x Post- 2013 & 2015			119.7 [222.5]	109.6 [224.6]
Mean of Comp. at Baseline	2654.6	2676.8	2654.6	2676.8
SD of Comp. at Baseline	3560.5	3562.8	3560.5	3562.8
Observations	5759	5759	5759	5759
Annual Income from independent - Job1 - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-293.3 [354.2]	-244.1 [354.0]		
Treatment x 2015- Endline	174.6 [428.6]	224.5 [422.3]		
Treatment x Post- 2013 & 2015			-55.2 [356.2]	-2.31 [353.0]
Mean of Comp. at Baseline	2817.5	2869.6	2817.5	2869.6
SD of Comp. at Baseline	5153.4	5197.1	5153.4	5197.1
Observations	5697	5697	5697	5697
Invest./Cost for independent - Job1 - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-30 [114.2]	-29.1 [111.5]		
Treatment x 2015- Endline	119.5 [108.2]	123.6 [105.5]		
Treatment x Post- 2013 & 2015			47.3 [105.5]	51 [102.4]
Mean of Comp. at Baseline	836.4	834.8	836.4	834.8
SD of Comp. at Baseline	1520.1	1511.3	1520.1	1511.3
Number of Clusters	272	272	272	272
Number of Households	2151	2151	2151	2151
Observations	5839	5839	5839	5839

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects, and controls for the number of members age 15 to 65 and the household size. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 38 Impact of Labor Income from three principal activities

	(1)	(2)	(3)	(4)
Annual Income from dependent activity - Total - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	25.3 [203.8]	25.5 [204.0]		
Treatment x 2015- Endline	69.6 [257.7]	64.4 [259.2]		
Treatment x Post- 2013 & 2015			44.2 [213.2]	41.8 [214.7]
Mean of Comp. at Baseline	2571.1	2593.8	2571.1	2593.8
SD of Comp. at Baseline	3426.1	3433.5	3426.1	3433.5
Observations	6516	6516	6516	6516
Annual Income from independent activity - Total - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-373.7 [338.0]	-331.5 [340.1]		
Treatment x 2015- Endline	-43.3 [401.2]	-13.6 [402.7]		
Treatment x Post- 2013 & 2015			-203.4 [333.3]	-165 [335.8]
Mean of Comp. at Baseline	2820.8	2890	2820.8	2890
SD of Comp. at Baseline	5286.3	5373.3	5286.3	5373.3
Observations	6258	6258	6258	6258
Annual Invest./Cost for independent activity - Total - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	12.7 [117.9]	10.9 [115.4]		
Treatment x 2015- Endline	102.7 [111.4]	111 [108.6]		
Treatment x Post- 2013 & 2015			59.5 [104.6]	63.3 [101.4]
Mean of Comp. at Baseline	670.5	675.3	670.5	675.3
SD of Comp. at Baseline	1448.9	1437.8	1448.9	1437.8
Number of Clusters	270	270	270	270
Number of Households	2157	2157	2157	2157
Observations	5879	5879	5879	5879

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects, and controls for the number of members age 15 to 65 and the household size. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Expenditures

Next we present impact estimates for the indicators of households' economic well-being. We present the measures of expenditures across different categories, including food and non-food expenditures. Since individuals tend to misreport their own income, the consensus is that measures relying on self-reported income are unreliable; thus, we focus on households' consumption expenditures, which more faithfully reflect economic well-being.

Our expenditure measure is composed of the following items: food (including food purchased, food consumed outside of the household, and self-supplied food products), non-food items (including transportation, non-durable household products, educational expenses, clothes, shoes, and travel expenses), and housing expenditures (including rent for renters, expected rental value for owners, utility payments, and fuel expenses). Expenditures on durable goods and tax payments are excluded [(Deaton & Zaidi, 1999)].

Table 39 disaggregates the annual per capita (PC) expenditures by the categories mentioned above, while Table 40 shows the impact for total expenditure. We present the expenditure in 2005 PPP dollars inflated to 2010 US prices. The log panel takes the logarithm of the outcome to adjust for over-dispersion/skewedness, and the coefficients can be interpreted as a percentage increase in income.

We also present the same per capita expenditure regressions transformed into logs in order to account for the large range of per capita expenditures observed in our sample. In these tables, the coefficients presented represent the *percent* change in per capita expenditures, rather than the *quantity* change between treatment and control households; this allows us to better contextualize the impact of the program. When we estimate the impact on the measures for food and non-food expenditures, rent, and utilities expenses, we do not find significant impacts in the treatment households versus the comparison households.

Note that this does not mean that treatment households experience no changes in welfare compared to their control counterparts; as we observed changes in investment in household businesses in the previous section and do not see a corresponding drop in household expenditures, it may be that treatment households reinvest additional income into productive activities rather than increasing consumption.

USAID-ACCESO targeted poor households in the western departments of Honduras. To gauge poverty levels of the households in the sample, we calculate daily consumption per capita and convert this figure to 2005 PPP (Purchasing Power Parity) dollars to use the international poverty line of 1.25 2005 PPP

dollars. The 1.25 2005 PPP poverty line is a measure of extreme poverty in the developing world and as such, baseline levels in our sample are relatively low. To better understand how the population in the ZOI fares relative to the rest of the country, we calculate the poverty under the baseline poverty line in Honduras (May 2012). The poverty line in Honduras is derived from the cost of a basic basket of goods (food and shelter) at current prices. To account for the different cost of living in urban and rural areas, the calculation is done separately for each. In addition to the national poverty line, the government calculates the national extreme poverty threshold, which is calculated based on the cost of the basic basket of food (it does not include the cost of shelter). We also look at the proportion of households that fall between the national poverty threshold and the extreme poverty threshold (i.e. households that can afford a basic basket of food, but not shelter); Table 41 shows impact estimates for the international poverty line and for extreme, relative, and total poverty using the Honduras poverty line. We find no significant differences between treatment and control households after treatment began in expenditure per capita or in the prevalence of poverty using any threshold. Note that this does not take into account improvements in the quality of goods consumed within each category. For example, households in the treatment group may not spend more on food, but they may consume smaller amounts of higher quality food (i.e., eating more vegetables, dairy, and meat rather than basic grains). USAID-ACCESO encouraged household production and consumption of eggs, poultry, and vegetables, and households may have been able to increase consumption of these higher quality foods without increasing expenditures. Thus, the lack of significant changes in expenditures does not mean that treatment households did not experience any welfare gains.

Table 39 Impact on Expenditure Measures by type

	(1)	(2)	(3)	(4)
Annual PC Food Expenditures in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-8.63 [10.4]	-7.96 [10.4]		
Treatment x 2015- Endline	-0.27 [13.5]	-0.33 [13.6]		
Treatment x Post- 2013 & 2015=1			-4.76 [10.5]	-4.46 [10.6]
Mean of Comp. at Baseline	355	354.9	355	354.9
SD of Comp. at Baseline	218.1	218.9	218.1	218.9
Observations	9055	9055	9055	9055
Annual PC Non-Food Expenditures in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-19.9 [16.8]	-20.9 [17.1]		
Treatment x 2015- Endline	-26.6 [28.2]	-27.8 [29.3]		
Treatment x Post- 2013 & 2015=1			-23.4 [21.4]	-24.5 [22.3]
Mean of Comp. at Baseline	165.7	166	165.7	166
SD of Comp. at Baseline	205.5	206.2	205.5	206.2
Observations	9055	9055	9055	9055
Annual Rent Expenditures PC in 2005 PPP 2010 USD				
Treatment x 2013- Midline	0.11 [9.35]	0.62 [9.51]		
Treatment x 2015- Endline	14.6 [12.2]	14.2 [12.6]		
Treatment x Post- 2013 & 2015=1			7.45 [9.61]	7.62 [9.95]
Mean of Comp. at Baseline	158.3	160	158.3	160
SD of Comp. at Baseline	175	177.6	175	177.6
Observations	9055	9055	9055	9055
Annual Utilities Expenditures PC in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-1.18 [2.29]	-1.25 [2.30]		
Treatment x 2015- Endline	-3.06 [2.58]	-3.03 [2.64]		
Treatment x Post- 2013 & 2015=1			-2.08 [2.03]	-2.11 [2.07]
Mean of Comp. at Baseline	92.3	92.9	92.3	92.9
SD of Comp. at Baseline	68.3	68.9	68.3	68.9
Observations	9055	9055	9055	9055
Annual Housing Related Expenditure PC in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-1.07 [10.2]	-0.63 [10.3]		
Treatment x 2015- Endline	11.6 [12.3]	11.2 [12.5]		
Treatment x Post- 2013 & 2015=1			5.37 [9.83]	5.51 [10.1]
Mean of Comp. at Baseline	250.6	252.9	250.6	252.9
SD of Comp. at Baseline	211.3	214.1	211.3	214.1
Number of Clusters	301	301	301	301
Number of Households	3322	3322	3322	3322
Observations	9055	9055	9055	9055

Standard errors in brackets. Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 40 Impact on total expenditure

	(1)	(2)	(3)	(4)
Expenditure PC per day in 2005PPP to 2010 Prices				
Treatment x 2013- Midline	-0.074 [0.055]	-0.072 [0.056]		
Treatment x 2015- Endline	-0.035 [0.078]	-0.037 [0.079]		
Treatment x Post- 2013 & 2015=1			-0.055 [0.058]	-0.055 [0.060]
Mean of Comp. at Baseline	1.97	1.97	1.97	1.97
SD of Comp. at Baseline	1.32	1.32	1.32	1.32
Observations	9055	9055	9055	9055
Expenditure PC per day in 2005 PPP USD				
Treatment x 2013- Midline	-0.066 [0.049]	-0.065 [0.050]		
Treatment x 2015- Endline	-0.031 [0.070]	-0.033 [0.071]		
Treatment x Post- 2013 & 2015=1			-0.049 [0.052]	-0.049 [0.054]
Mean of Comp. at Baseline	1.77	1.77	1.77	1.77
SD of Comp. at Baseline	1.18	1.19	1.18	1.19
Observations	9055	9055	9055	9055
Log-Expenditure PC in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-0.019 [0.027]	-0.018 [0.028]		
Treatment x 2015- Endline	0.022 [0.035]	0.02 [0.035]		
Treatment x Post- 2013 & 2015=1			0.0013 [0.026]	0.0011 [0.027]
Mean of Comp. at Baseline	0.49	0.49	0.49	0.49
SD of Comp. at Baseline	0.61	0.61	0.61	0.61
Observations	9055	9055	9055	9055
Log-Expenditure PC in 2005 PPP USD				
Treatment x 2013- Midline	-0.019 [0.027]	-0.018 [0.028]		
Treatment x 2015- Endline	0.022 [0.035]	0.02 [0.035]		
Treatment x Post- 2013 & 2015=1			0.0013 [0.026]	0.0011 [0.027]
Mean of Comp. at Baseline	0.38	0.38	0.38	0.38
SD of Comp. at Baseline	0.61	0.61	0.61	0.61
Number of Clusters	301	301	301	301
Number of Households	3322	3322	3322	3322
Observations	9055	9055	9055	9055

Standard errors in brackets. Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. * p<0.10, ** p<0.05, *** p<0.01

Table 41 Impact on Poverty

	(1)	(2)	(3)	(4)
Poverty <= 1.25 2005 PPP USD PC per Day				
Treatment x 2013- Midline	0.029 [0.028]	0.03 [0.028]		
Treatment x 2015- Endline	-0.015 [0.028]	-0.011 [0.028]		
Treatment x Post- 2013 & 2015=1			0.0068 [0.025]	0.0094 [0.025]
Mean of Comp. at Baseline	0.4	0.4	0.4	0.4
SD of Comp. at Baseline	0.49	0.49	0.49	0.49
Observations	9055	9055	9055	9055
Extreme Poverty - <= 2.84 2005 PPP 2010 USD				
Treatment x 2013- Midline	0.023 [0.022]	0.023 [0.023]		
Treatment x 2015- Endline	0.04 [0.040]	0.042 [0.041]		
Treatment x Post- 2013 & 2015=1			0.032 [0.030]	0.033 [0.031]
Mean of Comp. at Baseline	0.82	0.82	0.82	0.82
SD of Comp. at Baseline	0.38	0.38	0.38	0.38
Observations	9055	9055	9055	9055
Local Poverty <= 3.79 2005 PPP 2010 USD				
Treatment x 2013- Midline	0.011 [0.020]	0.011 [0.020]		
Treatment x 2015- Endline	0.033 [0.035]	0.036 [0.037]		
Treatment x Post- 2013 & 2015=1			0.023 [0.026]	0.024 [0.028]
Mean of Comp. at Baseline	0.91	0.91	0.91	0.91
SD of Comp. at Baseline	0.28	0.28	0.28	0.28
Observations	9055	9055	9055	9055
Relative Poverty - (2.84,3.79] 2005 PPP 2010 USD				
Treatment x 2013- Midline	-0.012 [0.019]	-0.012 [0.018]		
Treatment x 2015- Endline	-0.0067 [0.018]	-0.0067 [0.019]		
Treatment x Post- 2013 & 2015=1			-0.009 [0.017]	-0.0092 [0.017]
Mean of Comp. at Baseline	0.09	0.09	0.09	0.09
SD of Comp. at Baseline	0.29	0.29	0.29	0.29
Number of Clusters	301	301	301	301
Number of Households	3322	3322	3322	3322
Observations	9055	9055	9055	9055

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

To summarize the effects on household expenditure:

Expenditure

- ACCESO did not change household expenditure, even after correcting for over-dispersion of the data. This applies to aggregate household expenditure as well as to categories of household expenditures (rent, utilities, food and non-food expenditures).

Poverty

- ACCESO did not change proportion of individuals below the \$1.25/day poverty line, the local poverty line, the local extreme poverty line, or the local relative poverty threshold.

4.3 NUTRITION AND HEALTH

USAID-ACCESO's health and nutrition interventions aimed to decrease poverty and under-nutrition by encouraging crop diversification and increased agricultural productivity and thus improvements in the diets of beneficiary households.

In this section, we estimate the impacts on various measures of food security and health. For food security indicators, we show the impacts on the hunger scale [developed by the FANTA project, (Ballard, Coates, Swindale, & Deitchler, 2011)] and estimate the impacts on child and maternal health following the Feed the Future indicator handbook [(Feed the Future, 2014)] and the World Health Organization child and young infant feeding practices [(WHO, et al., 2008)].

Food Security: Hunger Scale

This measure indicates households experiencing moderate or severe hunger. For this indicator, respondents are asked about the frequency with which three events were experienced by household members in the last four weeks: having no food at all in the house; going to bed hungry; and going all day and night without eating. For each question, four responses are possible (never, rarely, sometimes, or often), which are collapsed and summed for each household to produce a household hunger score ranging from 0 to 6. Households with scores of 2 or more are classified as suffering from moderate or severe hunger.

In the baseline survey, only 3% of households were identified as suffering from chronic or persistent hunger. The lean season in the region starts around April, when the stocks from the postrera harvest in December have started to dwindle, and lasts until August, when the primera harvest becomes

available¹¹. Taking into account that each survey was conducted between May and July, hunger levels, as measured by the scale, seem extremely low for the lean season. We estimate the impact for the household hunger scale and for the probability of being classified as moderately or severely hungry and find that treatment households have lower scores on the hunger scale and are comparatively less likely to be classified as moderately or severely hungry (Table 42). These differences are not statistically significant at conventional levels, however, and we emphasize that since the initial level of hunger is low, it would be difficult to estimate small impacts on these indicators.

Table 42 Impacts of Food Security

	(1)	(2)	(3)	(4)
Indicator of households with moderate or severe hunger				
Treatment x 2013- Midline	-0.0058 [0.011]	-0.0056 [0.011]		
Treatment x 2015- Endline	-0.019 [0.011]*	-0.018 [0.011]*		
Treatment x Post- 2013 & 2015			-0.012 [0.0098]	-0.012 [0.0097]
Mean of Comp. at Baseline	0.03	0.031	0.03	0.031
SD of Comp. at Baseline	0.17	0.17	0.17	0.17
Observations	8728	8728	8728	8728
Hunger Scale				
Treatment x 2013- Midline	-0.024 [0.030]	-0.023 [0.029]		
Treatment x 2015- Endline	-0.046 [0.028]	-0.045 [0.028]		
Treatment x Post- 2013 & 2015			-0.035 [0.027]	-0.035 [0.026]
Mean of Comp. at Baseline	0.1	0.11	0.1	0.11
SD of Comp. at Baseline	0.46	0.47	0.46	0.47
Number of Clusters	301	301	301	301
Number of Households	3304	3304	3304	3304
Observations	8728	8728	8728	8728

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

¹¹ Postrera occurs between December to January and Primera occurs between August-September.

Child Nutrition and Health

Feeding practices are key to achieving an adequate childhood nutritional status, which in turn determines children's potential for development and growth. For example, in the first six months of life, breastfeeding provides children with the nutrients needed for proper growth and development and with immunization against a large number of common childhood diseases.

The World Health Organization (WHO) recommends that infants be exclusively breastfed for the first six months of life; to meet their new nutritional needs after six months, children should receive adequate and safe complementary food and should continue to breastfeed for up to two years of age or more.

Early initiation of complementary feeding without proper hygiene measures or adequate food preparation limits the value of breastfeeding by putting a child in contact with contaminated substances in the environment. Thus, breastfeeding practices and the use of complementary foods are certainly closely related to the risk of illness and death and the level of malnutrition in this population.

We gathered information on breastfeeding practices for children under 24 months of age; the introduction of solid foods; feeding frequency; and the use of fortified complementary foods or vitamin-mineral supplements. In this section, we present the impact results for measures constructed with this data. These impact estimates are obtained using the DID approach described in the methodology section, in which the data consists of individual-level data for each year. Mothers or caretakers answer questions regarding the feeding practices for each of the children living in the household in the eligible age range.

Table 43 presents the estimates for the probability of children under six months of age being exclusively breastfed and for the probability of children aged 6-23 months being breastfed. We find no significant impacts on these variables from the program.

Table 43 Impact on Breastfeeding Practices

	(1)	(2)	(3)	(4)
Indicator for exclusive breastfeeding (< 6 months)				
Treatment x 2013- Midline	-0.42 [0.31]	-0.45 [0.29]		
Treatment x 2015- Endline	-0.13 [0.19]	-0.15 [0.18]		
Treatment x Post- 2013 & 2015			-0.19 [0.25]	-0.23 [0.24]
Mean of Comp. at Baseline	0.82	0.81	0.82	0.81
SD of Comp. at Baseline	0.39	0.39	0.39	0.39
Observations	423	423	423	423
Indicator for current breastfeeding (6-23 months)				
Treatment x 2013- Midline	0.13 [0.075]*	0.12 [0.074]		
Treatment x 2015- Endline	0.034 [0.10]	0.026 [0.10]		
Treatment x Post- 2013 & 2015			0.098 [0.073]	0.09 [0.072]
Mean of Comp. at Baseline	0.55	0.55	0.55	0.55
SD of Comp. at Baseline	0.5	0.5	0.5	0.5
Number of Clusters	238	238	238	238
Number of Households	1091	1091	1091	1091
Observations	1497	1497	1497	1497

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the individual level and all regressions include household fixed effects, and controls for the number of children under five, the number of observations per household and sex. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

For the children age 6-23 months, we construct three measures to measure adequate nutrition: number of food groups consumed in the previous day (dietary diversity score), number of times that the child was fed (meal frequency), and an indicator for having a minimum acceptable diet.

We measure dietary diversity using a seven-group score. The seven foods groups used for tabulation of this indicator are grains, roots and tubers; legumes and nuts; dairy products (milk, yogurt, cheese); flesh foods (meat, fish, poultry, and liver/organ meats); eggs; vitamin-A rich fruits and vegetables; and other fruits and vegetables. Consumption of any amount of food from each food group is sufficient to increase the score. Consumption of at least four of these food groups on the previous day would indicate that the child (in most populations) had a high likelihood of consuming at least one animal-source food and at least one fruit or vegetable that day, in addition to a staple food. Minimum meal frequency gives us an estimate of energy intake from foods other than breast milk and helps us better measure the nutritional status of children. Feeding frequency for breastfed children includes only non-liquid feeds; for non-

breastfed children, feeding frequency includes both milk feeds and solid/semi-solid feeds. The minimum is defined as two feeding times for breastfed infants ages 6–8 months, three times for breastfed children ages 9–23 months, and four times for non-breastfed children ages 6–23 months. These two measures are then aggregated to construct a minimum acceptable diet indicator. A child is considered to have a minimum acceptable diet if they have a score¹² above four and have received the minimum meal frequency as defined above. Table 44 shows the impact estimates for these indicators; the impact estimates are not statistically significant different from zero.

We also compute indicators for the consumption of iron-rich or iron-fortified foods, vitamins, and oral hydration solutions for children ages 6-23 months. Iron-rich or iron-fortified foods include flesh foods, commercially foods fortified with iron and specially designed for infants and young children, or foods fortified in the home. We find no significant impacts for these indicators in Table 45.

These results suggest that even if treatment households were able to produce nutrient-dense foods, children did not benefit from this increased dietary diversity. Given the low levels of children consuming minimally acceptable diets at baseline, and the low prevalence of vitamin-rich food in children’s diets, these results suggest that USAID-ACCESO was not successful in encouraging households to adopt improved feeding practices for children.

¹² The score used to calculate this indicator is similar to the seven-group discussed above. The dietary diversity score used for MAD excludes milk feedings for non-breast fed children and includes a minimum *number* of two milk feedings to be considered adequate.

Table 44 Impact on Children Dietary Diversity

	(1)	(2)	(3)	(4)
Minimum dietary Diversity (7 Group Score)				
Treatment x 2013- Midline	-0.22 [0.22]	-0.2 [0.21]		
Treatment x 2015- Endline	0.19 [0.32]	0.2 [0.30]		
Treatment x Post- 2013 & 2015			-0.11 [0.21]	-0.093 [0.20]
Mean of Comp. at Baseline	2.36	2.34	2.36	2.34
SD of Comp. at Baseline	1.14	1.15	1.14	1.15
Observations	1512	1512	1512	1512
Minimum Meal Frequency (General)				
Treatment x 2013- Midline	0.086 [0.086]	0.086 [0.086]		
Treatment x 2015- Endline	-0.099 [0.098]	-0.097 [0.098]		
Treatment x Post- 2013 & 2015			0.036 [0.073]	0.035 [0.073]
Mean of Comp. at Baseline	0.42	0.42	0.42	0.42
SD of Comp. at Baseline	0.49	0.49	0.49	0.49
Observations	1512	1512	1512	1512
Minimum Acceptable Diet (6-23 months)				
Treatment x 2013- Midline	-0.093 [0.059]	-0.094 [0.058]		
Treatment x 2015- Endline	-0.026 [0.098]	-0.026 [0.099]		
Treatment x Post- 2013 & 2015			-0.075 [0.058]	-0.076 [0.057]
Mean of Comp. at Baseline	0.038	0.037	0.038	0.037
SD of Comp. at Baseline	0.19	0.19	0.19	0.19
Number of Clusters	239	239	239	239
Number of Households	1099	1099	1099	1099
Observations	1512	1512	1512	1512

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the individual level and all regressions include household fixed effects, and controls for the number of children under five, the number of observations per household and sex. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 45 Impact on Consumption of iron-rich foods, vitamins and hydration solutions (6-23)

	(1)	(2)	(3)	(4)
Consumption of hydration solutions				
Treatment x 2013- Midline	-0.0068 [0.039]	-0.008 [0.039]		
Treatment x 2015- Endline	-0.04 [0.043]	-0.041 [0.043]		
Treatment x Post- 2013 & 2015			-0.015 [0.031]	-0.017 [0.031]
Mean of Comp. at Baseline	1.93	1.93	1.93	1.93
SD of Comp. at Baseline	0.25	0.26	0.25	0.26
Observations	2010	2010	2010	2010
Consumption of vitamins				
Treatment x 2013- Midline	0.085 [0.083]	0.079 [0.082]		
Treatment x 2015- Endline	0.023 [0.12]	0.0038 [0.12]		
Treatment x Post- 2013 & 2015			0.07 [0.067]	0.059 [0.067]
Mean of Comp. at Baseline	1.63	1.62	1.63	1.62
SD of Comp. at Baseline	0.48	0.48	0.48	0.48
Observations	2010	2010	2010	2010
Consumption of iron-rich foods (6-23)				
Treatment x 2013- Midline	-0.03 [0.066]	-0.029 [0.066]		
Treatment x 2015- Endline	-0.036 [0.066]	-0.04 [0.068]		
Treatment x Post- 2013 & 2015			-0.029 [0.059]	-0.029 [0.058]
Mean of Comp. at Baseline	0.084	0.082	0.084	0.082
SD of Comp. at Baseline	0.28	0.27	0.28	0.27
Observations	1512	1512	1512	1512

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the individual level and all regressions include household fixed effects, and controls for the number of children under five, the number of observations per household and sex. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Child Anthropometry: Stunting, Wasting, and Underweight

Anthropometric data are a useful, simple, and practical way to describe the overall nutritional status of population groups. Their usefulness stems from anthropometry's close correlation with the multiple dimensions of individual health and development and their socio-economic and environmental determinants. In addition to the dietary information discussed in the previous section, these data are a useful way to design appropriate dietary interventions for targeted groups.

The standard indices of physical growth for children are height-for-age, weight-for-height, and weight-for-age. These indices are related to different aspects of nutritional status. A child who is below minus two standard deviations (-2 SD) from the median of a reference population in terms of height-for-age is considered short for his/her age, or stunted. Stunting reflects the cumulative effect of chronic malnutrition or long-term insufficient nutrient intake. A child who is below minus two standard deviations (-2 SD) from the median of a reference population in terms of weight-for-height is considered too thin for his/her height, or wasted. Wasting is a condition reflecting acute or recent nutritional deficit or severe food shortages. Weight-for-age is a composite index of stunting and wasting and is a good indicator to monitor nutritional status over time. A child who is below minus two standard deviations (-2 SD) from the median of a reference population in terms of weight-for-age is considered too thin for his/her age, or underweight.

Table 46 shows the impact results for the continuous indicators, the z-scores, for the measures of wasting, stunting, and underweight. We estimate impacts for the continuous z-scores for the children ages 0-59 months. In Table 47, we construct an indicator for being below minus 2 standard deviations from the median z-score of the reference population. An advantage of using the continuous variables is that we have more variation in the outcome measure, which facilitates the estimation of the impact; in the case of the indicator constructed using the z-score, the advantage is the ease of interpretation as a proportion of the children living in treatment households that has changed between baseline and follow-up when contrasted to the concurrent change in the comparison households. The impact estimates for these indices and the indicators for stunting, wasting, and underweight are not significantly different from zero in any of the specifications.

The fact that these indicators see no significant effects is not surprising. As we observed in the previous section, we see no impact on feeding practices among treatment households. While it is possible to

have impacts on severe acute malnutrition (SAM)¹³ or severe wasting (weight-for-height) in kids within the first 1,000 days of life, or in just a few months, wasting has a very low prevalence, and reducing SAM would require intensive targeting of wasted children. Given the low prevalence of wasting in the zone of influence (and in Honduras in general), it is difficult to detect changes in the whole population since there are not many children who are wasted. The question becomes more nuanced in Central America¹⁴. Given that wasting rates are low, increases in weight-for-height from baseline levels could indicate that the program is making children in the treatment group overweight. However, the results from the BMI regressions do not show significant increases in BMI for age.

In the case of impacts on height-for-age (stunting), we would not expect to see effects in the short run. Any intervention aimed to reduce the prevalence of stunting would have to be very intensive in addressing food/nutrient needs, encouraging behaviors that improve feeding practices (including exclusive breastfeeding and appropriate complementary feeding starting at 6 months of age), and reducing infections. To observe changes in this indicator, an intervention should target children as early and as intensively as possible (e.g. during pregnancy, or as early as possible during the first year of the child's life) and ideally should expose both the mother and the child to the intervention until the child reaches 24 months of age. It is difficult to impact stunting, and the earlier and longer you intervene within the first 1,000 days, the more likely the impact.

Child Anemia

We conduct anemia testing using a HemoCue photometer (Hb201+). Anemia is defined as a reduction in the normal number of red blood cells or a decrease in the concentration of hemoglobin in the blood. Symptoms of anemia range from pallor, fatigue, and weakness to shortness of breath and heart problems. Children who have hemoglobin levels below 11 g/dl are classified as anemic. Table 48 shows the results for the estimation on the hemoglobin level, on the prevalence of anemia, and on the presence of edema in children ages 6-59 month in the sample. The impact estimates for these measures are not statistically significant from zero. The data does not suggest that children living in households that participated in USAID-ACCESO have a lower prevalence of anemia or edema than similar children living in comparison households.

¹³ Severe acute malnutrition is defined by a very low weight for height (below -3z scores of the median WHO growth standards), by visible severe wasting, or by the presence of nutritional edema.[WHO].

¹⁴ In unreported regressions, when we estimate the impact for boys and girls separately, we find that boys are less likely to be underweight (a 15 percentage point difference from an 18 percent baseline).

Table 46 Impact on anthropometric z-scores

	(1)	(2)	(3)	(4)
Length/height-for-age Z-score				
Treatment x 2013- Midline	0.086 [0.14]	0.093 [0.14]		
Treatment x 2015- Endline	0.62 [0.70]	0.56 [0.65]		
Treatment x Post- 2013 & 2015			0.3 [0.30]	0.28 [0.28]
Mean of Comp. at Baseline	-1.65	-1.66	-1.65	-1.66
SD of Comp. at Baseline	2.02	2.05	2.02	2.05
Observations	3432	3432	3432	3432
Weight-for-age Z-score				
Treatment x 2013- Midline	-0.033 [0.092]	-0.026 [0.092]		
Treatment x 2015- Endline	-0.14 [0.19]	-0.1 [0.17]		
Treatment x Post- 2013 & 2015			-0.061 [0.10]	-0.054 [0.10]
Mean of Comp. at Baseline	-0.75	-0.76	-0.75	-0.76
SD of Comp. at Baseline	1.43	1.45	1.43	1.45
Observations	3300	3300	3300	3300
Weight-for-length/height Z-score				
Treatment x 2013- Midline	-0.068 [0.13]	-0.055 [0.13]		
Treatment x 2015- Endline	0.19 [0.15]	0.23 [0.15]		
Treatment x Post- 2013 & 2015			0.026 [0.12]	0.042 [0.12]
Mean of Comp. at Baseline	0.17	0.17	0.17	0.17
SD of Comp. at Baseline	1.09	1.09	1.09	1.09
Observations	3247	3247	3247	3247
BMI-for-age Z-score				
Treatment x 2013- Midline	0.25 [0.56]	0.21 [0.52]		
Treatment x 2015- Endline	1.54 [1.90]	1.75 [1.97]		
Treatment x Post- 2013 & 2015			0.74 [1.01]	0.79 [1.03]
Mean of Comp. at Baseline	0.47	0.48	0.47	0.48
SD of Comp. at Baseline	1.51	1.52	1.51	1.52
Number of Clusters	260	260	260	260
Number of Households	1573	1573	1573	1573
Observations	3289	3289	3289	3289

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the individual level and all regressions include household fixed effects, and controls for the number of children under five, the number of observations per household, sex and age. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 47 Impact on children underweight, stunting, and wasting

	(1)	(2)	(3)	(4)
Child is underweight				
Treatment x 2013- Midline	-0.00054 [0.024]	-0.0075 [0.025]		
Treatment x 2015- Endline	-0.009 [0.038]	-0.02 [0.039]		
Treatment x Post- 2013 & 2015			-0.0081 [0.022]	-0.012 [0.023]
Mean of Comp. at Baseline	0.16	0.16	0.16	0.16
SD of Comp. at Baseline	0.37	0.37	0.37	0.37
Observations	3363	3363	3363	3363
Child is Stunted				
Treatment x 2013- Midline	-0.031 [0.062]	-0.037 [0.060]		
Treatment x 2015- Endline	0.0064 [0.058]	0.0063 [0.054]		
Treatment x Post- 2013 & 2015			-0.024 [0.055]	-0.019 [0.053]
Mean of Comp. at Baseline	0.4	0.41	0.4	0.41
SD of Comp. at Baseline	0.49	0.49	0.49	0.49
Observations	3487	3487	3487	3487
Child is wasted				
Treatment x 2013- Midline	0.015 [0.020]	0.016 [0.020]		
Treatment x 2015- Endline	-0.024 [0.027]	-0.024 [0.027]		
Treatment x Post- 2013 & 2015			0.00049 [0.018]	0.00098 [0.019]
Mean of Comp. at Baseline	0.034	0.034	0.034	0.034
SD of Comp. at Baseline	0.18	0.18	0.18	0.18
Number of Clusters	260	260	260	260
Number of Households	1583	1583	1583	1583
Observations	3354	3354	3354	3354

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the individual level and all regressions include household fixed effects, and controls for the number of children under five, the number of observations per household, sex and age. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 48 Impact on Child anemia and edema

	(1)	(2)	(3)	(4)
Hemoglobine G/DI				
Treatment x 2013- Midline	0.041 [0.13]	0.059 [0.13]		
Treatment x 2015- Endline	0.17 [0.18]	0.19 [0.18]		
Treatment x Post- 2013 & 2015			0.097 [0.13]	0.12 [0.13]
Mean of Comp. at Baseline	11.6	11.6	11.6	11.6
SD of Comp. at Baseline	1.17	1.17	1.17	1.17
Observations	3004	3004	3004	3004
Child has anemia (6-59 months)				
Treatment x 2013- Midline	0.014 [0.065]	0.018 [0.063]		
Treatment x 2015- Endline	0.07 [0.063]	0.068 [0.062]		
Treatment x Post- 2013 & 2015			0.041 [0.055]	0.035 [0.054]
Mean of Comp. at Baseline	0.24	0.24	0.24	0.24
SD of Comp. at Baseline	0.43	0.43	0.43	0.43
Observations	3005	3005	3005	3005
Child has edema				
Treatment x 2013- Midline	0.015 [0.018]	0.015 [0.018]		
Treatment x 2015- Endline	-0.037 [0.027]	-0.034 [0.025]		
Treatment x Post- 2013 & 2015			-0.0043 [0.016]	-0.0048 [0.016]
Mean of Comp. at Baseline	0.028	0.028	0.028	0.028
SD of Comp. at Baseline	0.17	0.17	0.17	0.17
Number of Clusters	263	263	263	263
Number of Households	1619	1619	1619	1619
Observations	3499	3499	3499	3499

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the individual level and all regressions include household fixed effects, and controls for the number of children under five, the number of observations per household, sex and age. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Maternal Nutrition and Anemia

To characterize the effects of USAID-ACCESO on the health and nutrition on women of reproductive age (12-49 years of age), we will use three indicators. The first is the number of food groups consumed from a nine-item food group list. This indicator is used to proxy for micronutrient intake in women and their children (especially during pregnancy and lactation periods). Using a 10-food version of the dietary diversity score¹⁵, we construct a minimum acceptable dietary diversity indicator. This indicator is equal to one in which the female consumed five or more of the food groups in the past 24 hours. The results for these indicators are presented in Table 49, where we find significant impacts at endline; on average, after the intervention, women in treatment households consume 0.17 (5 percent of baseline level) more food groups than women in similar comparison households. We do not find significant effects on the indicator for minimum dietary diversity; only 40% of women consume a minimally diverse diet at baseline, with an average dietary diversity score of 3.23. An increase of 0.17 is insufficient to bring the average above 4, which is the threshold for consuming a minimally diverse diet.

Among women of reproductive age, height and weight measurement are used to calculate body mass index (BMI) and to assess a woman's risk of having difficulty during childbirth due to short stature (height <145 cm). BMI values are used to determine the percentage of the adult population that is normal, thin, overweight, and obese. In this study, the second indicator we will use to measure the nutritional status is the proportion of non-pregnant women of reproductive age who are underweight, as defined by body mass index (BMI) < 18.5. Table 50 shows the impact estimates for these outcomes. We do not find evidence of changes in these outcomes. At the 10% confidence level, we detect small differences in weight and height in the midline, with treatment women being marginally taller and heavier.

Lastly, we look at the impact of ACCESO on the anemia status of pregnant and non-pregnant women. Pregnant women who have hemoglobin levels below 11 g/dl are classified as anemic, while non-pregnant women with hemoglobin levels below 12 g/dl are classified as anemic. The results for this indicator and for the level of hemoglobin in women's blood are presented in Table 51; we find no significant differences between treatment and comparison women across specifications. Our results are unsurprising, as we find limited evidence of improved nutrition among women in our sample and only a

¹⁵ Because some food groups were agglomerated in one category, we are not able to separate some scores. To resolve this, we constructed an eight-food group score and ten-food group score by doubling the score for agglomerated food groups. The results are qualitatively the same.

very small fraction of this improvement would have been due to increased consumption of iron-rich foods.

Table 49: Impact on Maternal Nutrition Outcomes

	(1)	(2)	(3)	(4)
Dietary Diversity (Mean # of food groups (9) consumed)				
Treatment x 2013- Midline	0.12 [0.088]	0.11 [0.088]		
Treatment x 2015- Endline	0.22 [0.098]**	0.22 [0.098]**		
Treatment x Post- 2013 & 2015			0.17 [0.076]**	0.17 [0.077]**
Mean of Comp. at Baseline	3.23	3.23	3.23	3.23
SD of Comp. at Baseline	1.08	1.07	1.08	1.07
Observations	8395	8395	8395	8395
Dietary Diversity (Mean # of food groups (10) consumed)				
Treatment x 2013- Midline	0.12 [0.13]	0.12 [0.13]		
Treatment x 2015- Endline	0.27 [0.13]**	0.27 [0.13]**		
Treatment x Post- 2013 & 2015			0.21 [0.11]*	0.21 [0.11]*
Mean of Comp. at Baseline	4.23	4.22	4.23	4.22
SD of Comp. at Baseline	1.32	1.31	1.32	1.31
Observations	8395	8395	8395	8395
Women Min. Dietary Diversity (10 Food)				
Treatment x 2013- Midline	0.039 [0.044]	0.04 [0.045]		
Treatment x 2015- Endline	0.049 [0.054]	0.052 [0.056]		
Treatment x Post- 2013 & 2015			0.045 [0.044]	0.047 [0.046]
Mean of Comp. at Baseline	0.39	0.39	0.39	0.39
SD of Comp. at Baseline	0.49	0.49	0.49	0.49
Number of Clusters	289	289	289	289
Number of Households	2847	2847	2847	2847
Observations	8395	8395	8395	8395

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003).

Observations are at the individual level and all regressions include household fixed effects, and controls for the number of women age 15 to 49 and the number of observations per household and age. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 50 Impact on Maternal Anthropometric Outcomes

	(1)	(2)	(3)	(4)
Weight (kgs.)				
Treatment x 2013- Midline	1.06 [0.57]*	1.07 [0.57]*		
Treatment x 2015- Endline	0.37 [0.77]	0.39 [0.75]		
Treatment x Post- 2013 & 2015			0.7 [0.63]	0.71 [0.62]
Mean of Comp. at Baseline	53.9	53.9	53.9	53.9
SD of Comp. at Baseline	12.3	12.3	12.3	12.3
Observations	7216	7216	7216	7216
Height (cms.)				
Treatment x 2013- Midline	1.12 [0.63]*	1.1 [0.63]*		
Treatment x 2015- Endline	1.34 [0.94]	1.21 [0.91]		
Treatment x Post- 2013 & 2015			1.23 [0.65]*	1.15 [0.64]*
Mean of Comp. at Baseline	151.9	151.9	151.9	151.9
SD of Comp. at Baseline	8.28	8.25	8.28	8.25
Observations	7243	7243	7243	7243
BMI of Woman				
Treatment x 2013- Midline	-0.022 [0.16]	-0.01 [0.16]		
Treatment x 2015- Endline	-0.088 [0.23]	-0.073 [0.23]		
Treatment x Post- 2013 & 2015			-0.056 [0.17]	-0.043 [0.17]
Mean of Comp. at Baseline	23.1	23.1	23.1	23.1
SD of Comp. at Baseline	3.86	3.84	3.86	3.84
Observations	7104	7104	7104	7104
Women with BMI < 18.5 - Underweight				
Treatment x 2013- Midline	-0.0062 [0.015]	-0.0068 [0.015]		
Treatment x 2015- Endline	-0.029 [0.019]	-0.03 [0.019]		
Treatment x Post- 2013 & 2015			-0.018 [0.014]	-0.019 [0.014]
Mean of Comp. at Baseline	0.074	0.072	0.074	0.072
SD of Comp. at Baseline	0.26	0.26	0.26	0.26
Number of Clusters	284	284	284	284
Number of Households	2641	2641	2641	2641
Observations	6725	6725	6725	6725

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the individual level and all regressions include household fixed effects, and controls for the number of women age 15 to 49 and the number of observations per household and age. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 51 Impact on anemia in women of reproductive age

	(1)	(2)	(3)	(4)
Hemoglobin G/DI				
Treatment x 2013- Midline	0.066 [0.100]	0.062 [0.10]		
Treatment x 2015- Endline	0.087 [0.11]	0.088 [0.11]		
Treatment x Post- 2013 & 2015			0.075 [0.095]	0.072 [0.097]
Mean of Comp. at Baseline	13.4	13.3	13.4	13.3
SD of Comp. at Baseline	1.27	1.26	1.27	1.26
Observations	6648	6648	6648	6648
Woman has anemia				
Treatment x 2013- Midline	-0.0077 [0.024]	-0.0079 [0.024]		
Treatment x 2015- Endline	-0.028 [0.030]	-0.03 [0.030]		
Treatment x Post- 2013 & 2015			-0.018 [0.024]	-0.019 [0.024]
Mean of Comp. at Baseline	0.11	0.11	0.11	0.11
SD of Comp. at Baseline	0.31	0.31	0.31	0.31
Number of Clusters	284	284	284	284
Number of Households	2592	2592	2592	2592
Observations	6648	6648	6648	6648

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the individual level and all regressions include household fixed effects, and controls for the number of women age 15 to 49 and the number of observations per household and age. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

To summarize the impacts on hunger, child health and nutrition, and maternal health and nutrition:

Hunger

- 1.8 percentage point decrease in moderate to severe hunger, only at endline, imprecisely measured.

Child Health and Nutrition

- No significant robust program effects on anthropometrics, dietary diversity, or health.

Maternal Health and Nutrition

- No effects on maternal anemia;
- 1cm increase in height among treatment group , imprecisely estimated; and
- 0.17 increase in 9-group dietary diversity score.

4.4 WOMEN'S EMPOWERMENT AND TIME USE

Women play an essential role in efforts to promote agricultural growth in developing economies. However, some of the potential arenas in which women can participate in the agricultural sector are limited by social and economic constraints. The Women's Empowerment in Agriculture Index (WEAI) measures the empowerment, agency, and inclusion of women in the agricultural sector in an effort to identify ways to overcome those obstacles and constraints. Specifically, it measures the roles and extent of women's engagement in the agricultural sector in five domains: (1) decisions about agricultural production, (2) access to and decision-making power over productive resources, (3) control over use of income, (4) leadership in the community, and (5) time use. It also measures women's empowerment relative to men within their households. [(Alkire, et al., 2012)]

As the measurement of empowerment is at the center of the Feed the Future interventions, the questionnaire included the variables necessary to calculate this composite indicator. Table 52 shows the results from our calculations using the methodology in (Alkire, et al., 2012) for the baseline and follow-up survey.

The first element of the WEAI is the 5DE index, which measures whether women (or men) are empowered across the five dimensions mentioned within their households and their communities. It is composed by the disempowered headcount ratio and the average adequacy score of disempowered individuals, which measure disempowerment at the extensive and intensive margin. The disempowered headcount ratio measures the proportion of males and females in the population classified as disempowered. The average adequacy ratio measures the intensity of the disempowerment of each individual as a function of the areas in which each individual feels they can make decisions. The adequacy score ranges from 0, for individuals who do not make decisions in any of the five domains, and 1 for those who participate in all domains.

Figure 8 shows the percentage of women that are **empowered** (empowered headcount ratio) in each of the dimensions for each year of the survey. The main difference across years is the proportions for input in productive activities and for autonomy in production. The headcount for women being empowered in providing input in agricultural production decisions is decreasing; for autonomy in production decisions, the proportion is lower in the mid-line and rebounds in the endline. The proportion of women empowered with respect to the control of income derived from agriculture increases in both follow-up

surveys¹⁶. In addition, we decompose the empowered headcount ratio in the follow-up surveys by treatment status; the results are presented in Figure 9. The headcount ratios are very similar for the comparison and treatment group in each year, with the treatment group having a higher proportion of women empowered when it comes to make decisions about purchasing or sales of assets and access to credit.

After aggregating these measures for the zone of influence in each year, in Table 52, we see that 68.5% of the women in our sample are classified as disempowered in the baseline survey versus 65.1% in the endline survey. We make a similar comparison for males and find that 39.9% of males are classified as disempowered in the baseline versus 26.5% in the endline survey. The inadequacy scores are similar, but women experience inadequacy in more domains than men. The 65% of women who are not yet empowered have, on average, inadequate achievements in 38% of domains in the endline survey; this is very similar to what we found in the baseline survey. In addition to a lower proportion of males in the endline being classified as disempowered, the depth of disempowerment is also less, with males being disempowered in only 29.5% of the domains in the end line versus 31% found in the baseline. In Table 53, we decompose these figures by treatment status for the 2013 and 2015 follow-up data. By the endline survey, women in the treatment and comparison groups have similar levels of inadequacies (38.9 and 37.9% of domains in the endline survey, respectively). The largest difference between the groups is in the disempowered headcount ratio for both sexes, with the comparison group being proportionally more disempowered.

The 5DE index is one minus the product of these two measures ($5DE\ Index = 1 - H_{20p} * A_{20p}$). Our results in Table 52 imply that the women in the sample have adequate empowerment in 77% of the indicators in the endline compared to 75% in the baseline. The estimated level of empowerment using the 5DE index is below the 80% recommended cut-off in each year of the survey, above which a woman would be classified as empowered. In addition, the level is well below the empowerment of men, with 87.6% of the indicators being classified as adequate for men in the baseline and 92.3% in the endline.

¹⁶ The measure headcount ratio for empowerment in ownership of assets and asset decision-making for the follow-up survey uses the imputed values from the baseline in the cases where, due to non-response in the empowerment section, households are classified as having no assets. That is, we impute the lag value for the household sex pair if the household reports having the asset elsewhere in the questionnaire and has not answered the decision-making questions. Households that after this remained with no asset are classified as inadequate in the domain as described in (Alkire, et al., 2012).

Figure 10 shows the contribution of each domain to the empowerment index for both males and females in each year. The purpose of this is to compare men and women across the different dimensions and to see in which dimensions women are lagging. From the figure, we note that the largest differences in contribution come from control of income and access to resources (credit and ownership)¹⁷. The corresponding results for the decomposition by treatment status for the 5DE Index in 2013 and 2015 are presented in Figure 11. The empowerment of women across groups is very similar in both follow-up years. Within each group, the contrast between men and women is consistent with what we observed at the zone of influence level for each year; control of income and access to resources are the main drivers of the difference between men and women.

The second part of the WEAI is the Gender Parity Index (GPI), which measures the level of inequality within dual households. A household enjoys parity if the principal female is empowered or, if she is not empowered, if her adequacy score is greater than or equal to that of the primary male in her household. In a similar fashion to the 5DE index, the GPI is comprised of two margins of parity. On the extensive margin, we have the gender parity-inadequate headcount ratio, which measures the proportion of households in which women have not achieved adequate gender parity. On the intensive margin, we have the average empowerment gap (censored parity inadequacy scores), which measures the average percentage gap among the households classified as gender parity-inadequate, (i.e. the normalized difference between the female and males inadequacy scores for dual households that are gender parity-inadequate). The gender parity index is then obtained as one minus the product of the headcount and the average empowerment gap ($GPI = 1 - H_{GIP} * A_{GPI}$) in dual households. In the zone of influence in our follow-up sample, shown in Table 52, 14% of dual households are classified as gender parity-inadequate, a much lower estimate than at baseline. The intensity of the inequality within households is given by the average empowerment gap. Of the 14% of women who are less empowered than the men in their households, the empowerment gap between them and the males in their households is 46.6%, which is relatively large. Contrasting the parity numbers between baseline and endline and taking into consideration that the empowerment of men increased during the period, these results suggest that the empowerment of women in dual households is driving the increase in the 5DE index for women and that the women who remained disempowered saw an increase in the gap between themselves and the men in their household. The gender parity index in our sample is 0.874 in the baseline and 0.935 in the

¹⁷ The calculations for each contribution to empowerment can be found in “Appendix A – Decomposition of Women Empowerment Index in Agriculture”

endline. Table 53 shows the corresponding measures for the follow-up sample when we decompose the GPI by treatment status.

Finally, the WEAI is obtained by combining the GPI and the 5DE index. The index is:

$$WEAI = 0.9 * (5DE\ Index) + 0.1 * (GPI\ Index)$$

In the zone of influence, the WEAI in the follow-up is 0.75 and 0.77 in the endline. Both of these are below the 0.80 recommended cut-off to classify women as empowered. The decomposition by treatment status in Table 53 shows that women in the treatment group have higher levels of empowerment in both follow-ups and that both groups have similar levels of empowerment to the aggregate zone of influence figures discussed above, having experienced increases in this measure of women's empowerment. The differences across group are not significantly different. However, the Index and its components provide evidence on the domains of empowerment that are lacking for women in the ZOI.

Figure 8 Proportion of Women that are EMPOWERED by indicator in the ZOI of FTF in Honduras by Year

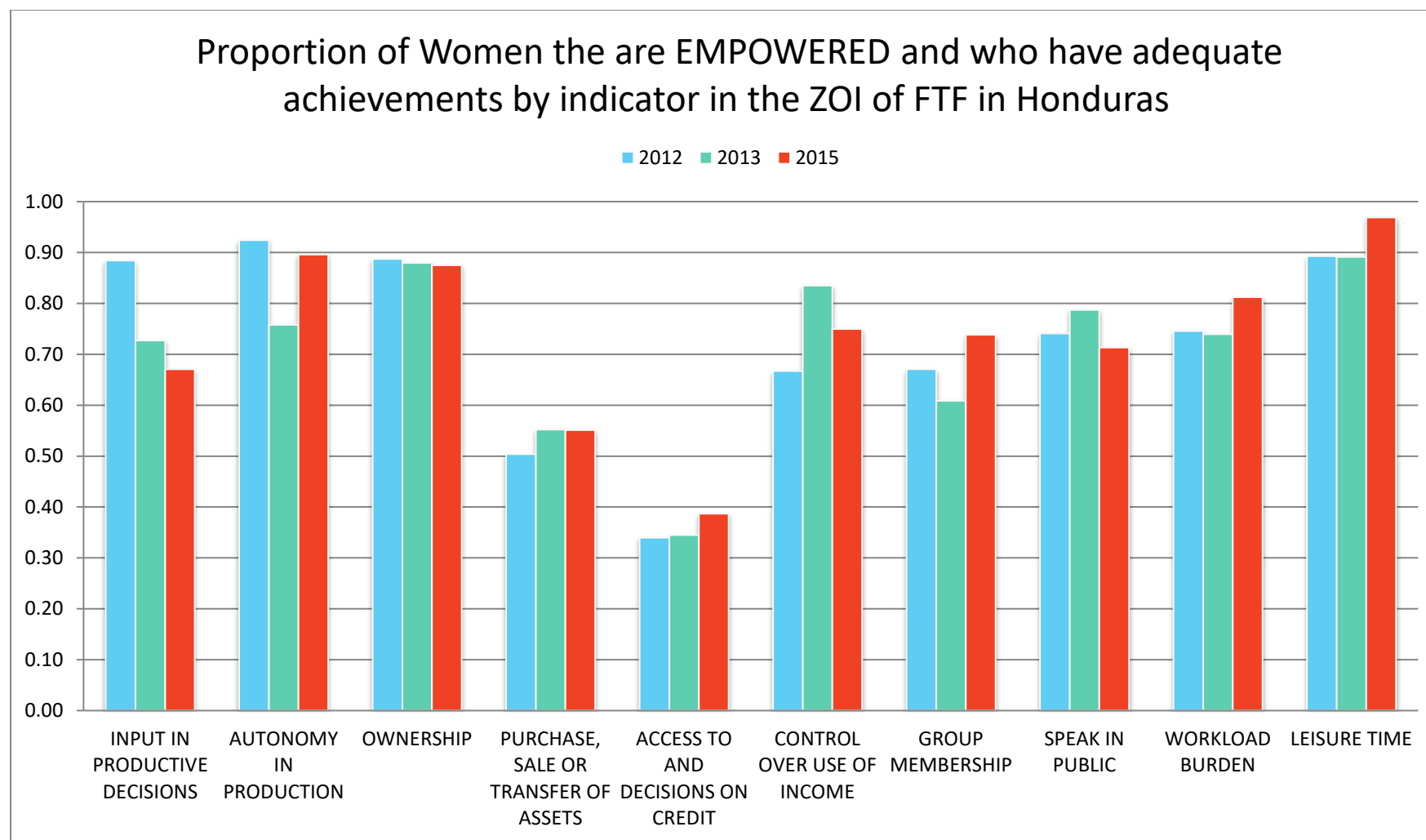


Figure 9 Proportion of Women that are EMPOWERED by indicator in the ZOI of FTF in Honduras by Treatment Status in 2013 and 2015

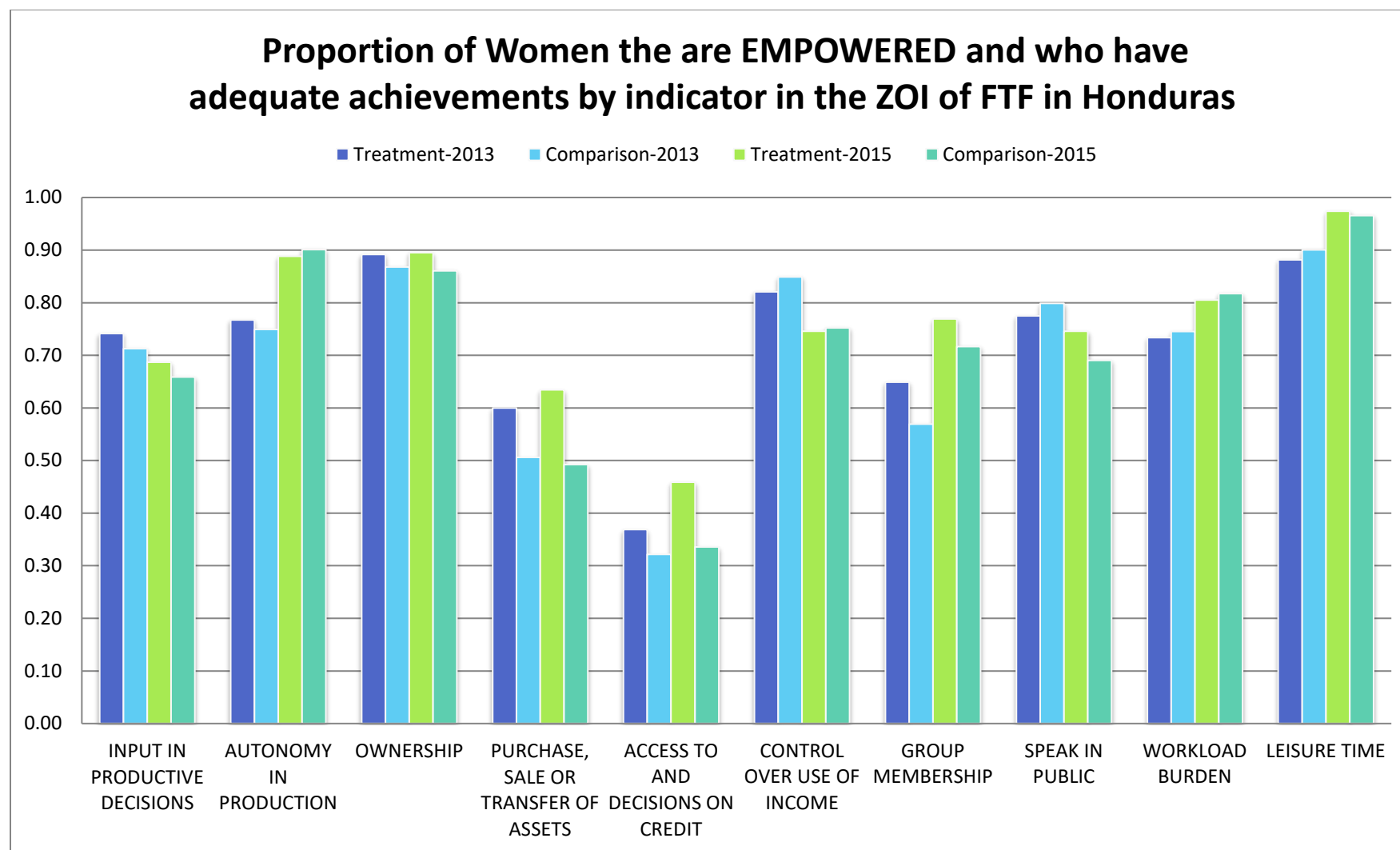


Figure 10 5DE Decomposed by Dimension and Indicator for each year

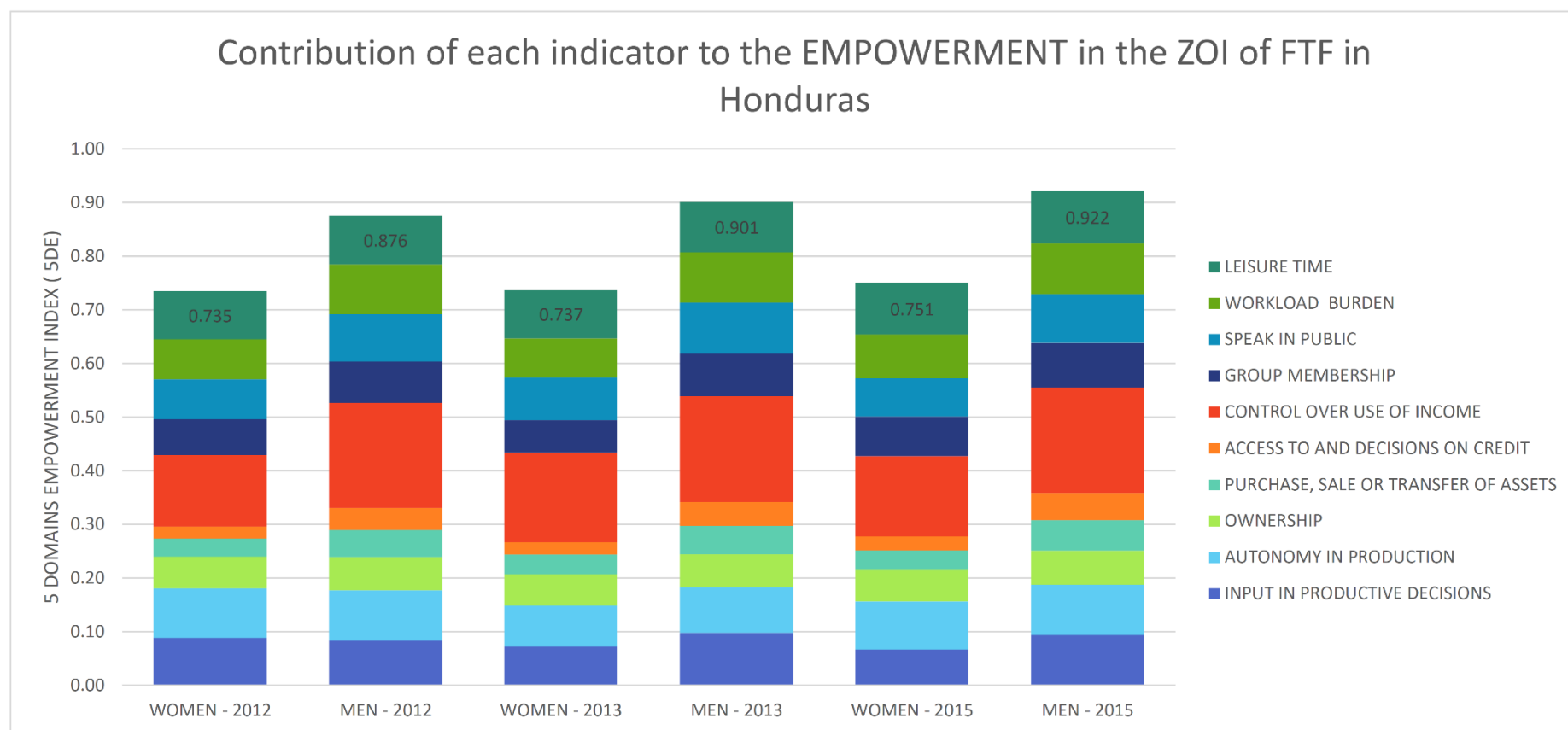


Figure 11 SDE Decomposed by Dimension and Indicator for each Treatment Status in 2013 and 2015

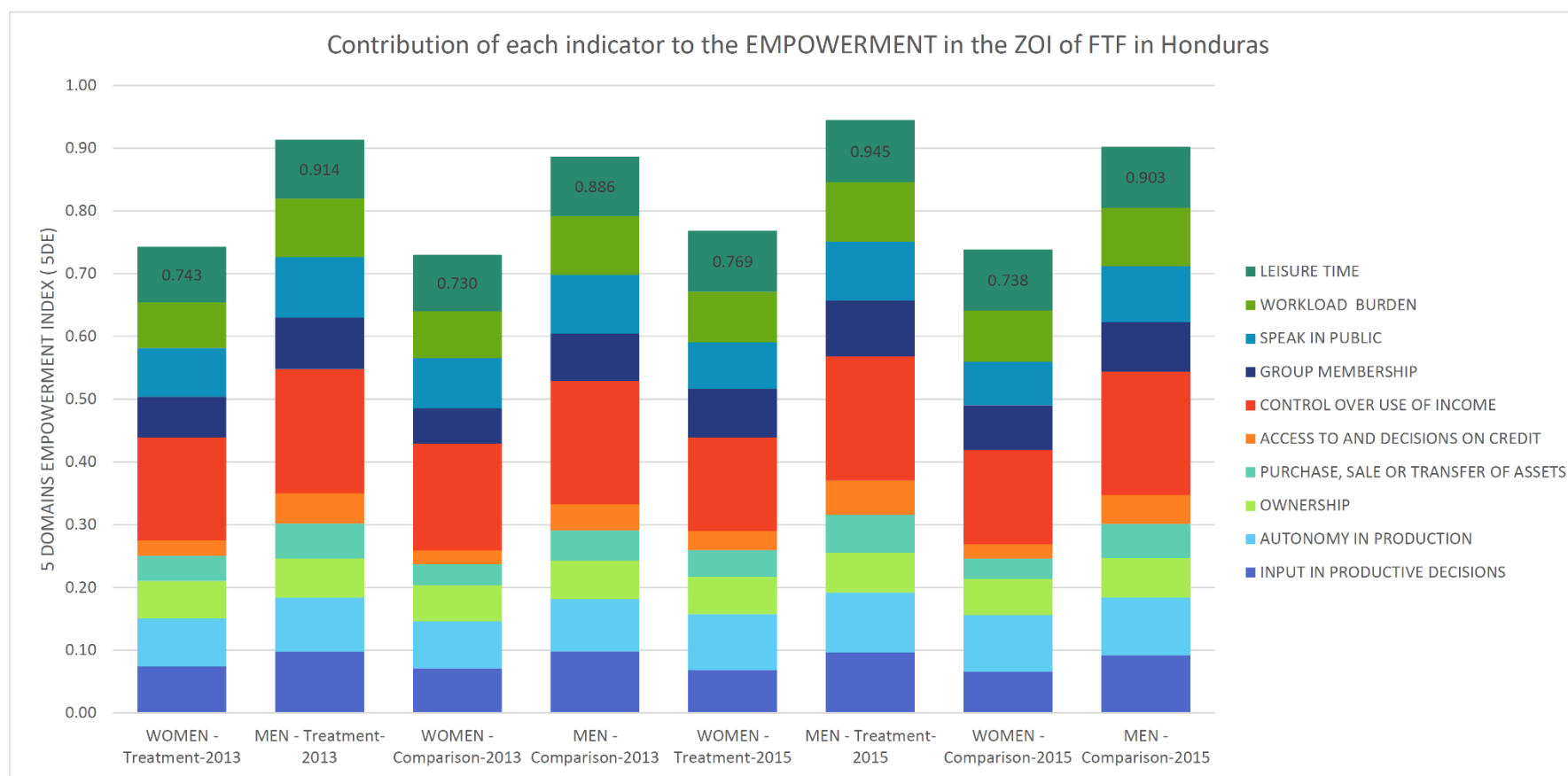


Table 52 Empowerment in Agriculture Index by Year

	2012		2013		2015	
	Female	Male	Female	Male	Female	Male
DISEMPOWERED HEADCOUNT (H_20p)	0.685	0.399	0.699	0.342	0.651	0.265
AVERAGE INADEQUACY SHARE (A_20p)	0.387	0.311	0.377	0.288	0.383	0.295
5 DOMAINS DISEMPOWERMENT INDEX (MO_20p)	0.265	0.124	0.263	0.099	0.249	0.078
5 DOMAINS EMPOWERMENT INDEX (EA_20P)	0.735	0.876	0.737	0.901	0.751	0.922
INADEQUACY HEAD COUNT (H_GPI)	0.581		0.591		0.140	
CENSORED INADEQUACY SCORES AVERAGE	0.218		0.224		0.466	
GENDER DISPARITY INDEX (PI)	0.126		0.132		0.065	
GENDER PARITY INDEX (GPI)	0.874		0.868		0.935	
EMPOWERMENT IN AGRICULTURE INDEX	0.749	0.876	0.750	0.898	0.769	0.923

Table 53 Empowerment in Agriculture Index by Treatment Status in 2013 and 2015

	2013				2015			
	Comparison		Treatment		Comparison		Treatment	
	Female	Male	Female	Male	Female	Male	Female	Male
DISEMPOWERED HEADCOUNT (H_20p)	0.728	0.397	0.669	0.297	0.690	0.326	0.595	0.192
AVERAGE INADEQUACY SHARE (A_20p)	0.371	0.286	0.384	0.290	0.379	0.299	0.389	0.287
5 DOMAINS DISEMPOWERMENT INDEX (MO_20p)	0.270	0.114	0.257	0.086	0.262	0.097	0.231	0.055
5 DOMAINS EMPOWERMENT INDEX (EA_20P)	0.730	0.886	0.743	0.914	0.738	0.903	0.769	0.945
INADEQUACY HEAD COUNT (H_GPI)	0.592		0.590		0.176		0.115	
CENSORED INADEQUACY SCORES AVERAGE	0.209		0.236		0.442		0.491	
GENDER DISPARITY INDEX (PI)	0.124		0.139		0.078		0.057	
GENDER PARITY INDEX (GPI)	0.876		0.861		0.922		0.943	
EMPOWERMENT IN AGRICULTURE INDEX	0.745	0.885	0.755	0.908	0.756	0.904	0.786	0.945

5 CONCLUSIONS

We have performed an in-depth analysis of the baseline and two follow-up surveys designed to evaluate the impact of the USAID-ACCESO intervention in the Occidental departments of Honduras. The goal was to estimate the impact across the main indicators and explore ways in which the impact evaluation design could have been compromised by changes that occurred between its conception and project finalization.

We estimate the impact of USAID-ACCESO across all the main outcomes that Feed the Future monitors and add other outcomes of interest to provide a narrative for the impact pathway when impacts are observed or to provide context when the expected impacts are not found. We find some significant effects across the main outcome measures, but most are not robust across model specifications. The evidence regarding the impact of USAID-ACCESO presented can be summarized as follows.

Agriculture

- A 10 to 15% increase in the coffee and corn area planted by endline
- Some evidence of increases in diversification through vegetable and tuber production, but with a small sample size, so few households affected
- No increases in the level of agricultural production or sales
- Decreases in the value of production and sales that depend on the valuing price, with no effects when using median prices and persistent negative effects when adjusted for dispersion (logs). Evidence that farmers might understate their prices or, less likely, that a large heterogeneity in prices exists in the area.
- Negative impacts on agricultural production and productivity in the midline and a reversion to the baseline levels in the endline. This suggests mean reversion, or returning to the long-run mean. The treatment households might have been able to recover after a negative shock at midline.
- Improvement in knowledge of coffee rust in both follow-up surveys, with increases in assistance for the treatment group

Income and Expenditure

- Increase in wage labor income and income investment in independent business

- Increase in labor supply in the agricultural sector and wage labor
- No impact on total household expenditures or across expenditure categories
- No effects on prevalence of poverty using international or local poverty lines

Women Health and Nutrition

- No effects on weight, height, or prevalence of anemia in women of reproductive age
- Increase in women's dietary diversity with women in treatment households consuming more food groups (5%) at endline

Child Health and Nutrition

- No effects on weight, height, or prevalence of anemia in children under five years old
- Disaggregating effects by sex shows a decrease in the proportion of boys underweight, but no effects for girls.
- No effect on feeding practices for children under 24 months of age

Empowerment

- Women experience a small increase in empowerment across time.
- Women in treatment households are slightly more empowered than women in comparison households.

Discussion of Impacts and Lessons for Future Interventions

We detect no effects on some measures that a priori should be affected by USAID-ACCESO, including agricultural productivity, agricultural sales, and nutritional outcomes. We expected that by the endline survey, households would have changed their agricultural and nutritional practices as they participated in USAID-ACCESO for longer periods of time, and that these changes in practices would translate into improvements in final outcomes. However, we observe no changes in intermediate outcomes or in the main final outcomes.

The fact that there were no significant effects on nutritional indicators is not surprising. Given the low prevalence of low weight-for-height (wasting), it is difficult to detect changes in the whole population since there are not many children who are wasted. To detect impacts on these indicators, we need to target the intervention to malnourished children and specifically follow them in the survey. Similarly, impacts on height-for-age (stunting) are a long-term goal; to observe changes in stunting rates due to an

intervention, that intervention needs to be very intensive in addressing food/nutrient needs, improving feeding practices (including exclusive breastfeeding and appropriate complementary feeding starting at 6 months of age), and reducing infections. Thus, the intervention should target children as early as possible (e.g. during pregnancy, or as early as possible during the first year of the child's life) and ideally should expose both the mother and the child to the intervention until the child reaches 24 months of age. It is difficult to impact stunting, and the earlier and longer you intervene within the first 1,000 days, the more likely the impact.

In the case of impacts on agricultural productivity and market access indicators, we expect that longer exposures to treatment, through new interventions, will move these indicators as farmers learn better techniques and become accustomed to new technologies. One possible explanation as to why we did not observe significant impacts on these indicators are the existence of large aggregate shocks. Although aggregate shocks are not likely to differentially affect the treatment and comparison groups, such a shock (like coffee rust) could decrease the resources available to all households and, as a consequence, the take-up of new practices and introduction of more nutritious foods. In addition, the initial conditions necessary to take advantage of new market opportunities might be lower in general, and even when the treatment households are receiving information and assistance, they might not be able to take advantage of these opportunities.

The results of this impact evaluation place us in a better position to understand the mechanisms through which interventions that couple agricultural and nutritional assistance affect the well-being of beneficiaries.

Lessons for the future

The results from the analysis show that the research design is appropriate, but changes in the treatment status in the final stage should be taken into account in order to improve the precision of the estimates by increasing the sample of treatment households.

The attrition in the sample was larger than expected; coupled with the need to estimate the effect of the program on specific sectors of the population (e.g. children under five), this requires a larger sample size to better estimate the impact on subpopulations, not just the means across the different strata.

The main lesson for the future is the need to focus on specific activities that would be expected to have large impacts on vulnerable populations. It is recommended that future impact evaluations balance the evaluation of a program as diverse as USAID-ACCESO as a package and the evaluation of specific and

highly targeted activities within the program. When evaluating packaged programs, the impacts of the different parts of the package are difficult to separately identify. The importance of separating these effects depends much on the context being studied. In our case, the levels of malnutrition in Honduras are large enough that it would be interesting to estimate these separately; however, this would entail a nutrition-targeted intervention and impact evaluation specifically designed for such an intervention. A targeted evaluation would bring new evidence on which, and what size, differences in implementation could improve the effectiveness of nutrition and agriculture programs similar to USAID-ACCESO. In addition it will also be important to include a qualitative assessment to better understand the impact pathways of the highly targeted activities to be able to better understand the reasons and process of the success or failure of these activities.

Finally, one important lesson is the need to improve the monitoring of the implementation of the project activities to assure that all activities are implemented according to the design and timeline. This will help not only to implement corrective measures but also to better understand if the non-impact is because of an improper implementation of the program or because the interventions are not appropriate.

6 REFERENCES

- Alkire, S., Meinzen-Dick, R., Peterman, A., Quisumbing, A. R., Seymour, G., & Vaz, A. (2012). *The Women's Empowerment in Agriculture Index*. International Food Policy Research Institute (IFPRI).
- Ballard, T., Coates, J., Swindale, A., & Deitchler, M. (2011). *Household Hunger Scale (HHS): Indicator Definition and Measurement Guide*. FHI 360, Food and Nutrition Technical Assistance II Project, Washington, DC. Retrieved from <http://www.fantaproject.org/sites/default/files/resources/HHS-Indicator-Guide-Aug2011.pdf>
- Deaton, A., & Zaidi, S. (1999). *Guidelines for Constructing Consumption Aggregates For Welfare Analysis*. Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies. Princeton University.
- Dehejia, R., & Wahba, S. (1994). Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs. *Journal of the American Statistical Society*, 1053-1062.
- Donner, A. (1998). Some Aspects of the Design and Analysis of Cluster Randomization Trials. *Journal of the Royal Statistical Society*, 47(1).
- Donner, A., & Birkett N, B. C. (1981). Randomization by cluster: sample size requirements and analysis. *American Journal of Epidemiology*(114), 906-914.
- Duflo, E., Glennerster, R., & Kremer, M. (2008). Using Randomization in Development Economics Research: A Toolkit. (T. Schultz, & J. Strauss, Eds.) *Handbook of Development Economics*, 4, 3895-3962.
- Feed the Future. (2014). *Feed the Future Handbook of Indicator Definitions*. Retrieved from https://feedthefuture.gov/sites/default/files/resource/files/ftf_handbook_indicators_october2014.pdf
- FEWS NET. (February 2014). *SPECIAL REPORT Central America, Coffee producer and laborer income to decline for a second consecutive year*. Retrieved from http://www.fews.net/sites/default/files/documents/reports/FEWS%20NET%20coffee%20rust%20shock%20report%20update_2014_02_en_0.pdf

- Fitzgerald, J., Gottschalk, P., & Moffitt, R. (1998). An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics. *Journal of Human Resources*, 33(2), 251-299.
- Heckman, J., & Robb, R. (1984). Alternative Methods for Evaluating the Impact of Interventions. In J. Heckman, & B. Singer (Eds.), *Longitudinal Analysis of Labor Market Data* (pp. 156-245). Cambridge, U.K.: Cambridge University Press.
- Heckman, J., Ichimura, H., & Todd, P. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program. *Review of Economic Studies*, 605-654.
- Heckman, J., Ichimura, H., & Todd, P. (1998). Matching as an Econometric Evaluation Estimator. *Review of Economic Studies*, 261-294.
- Heckman, J., Ichimura, H., Smith, J., & Todd, P. (1998). Characterizing Selection Bias Using Experimental Data. *Econometrica*, 1017-1098.
- Hirano, K., & Imbens, G. W. (2002). Estimation of Causal Effects using Propensity Score Weighting: An Application to Data on Right Heart Catheterization. *Health Services & Outcomes Research Methodology*, 2, 259-278.
- Hirano, K., Imbens, G. W., & Ridder, G. (2003). Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. *Econometrica*, 71(4), 1161-1189.
- INEI. (2012). *Encuesta Permanente de Hogares de Propósitos Múltiples*. Retrieved from <http://www.inec.gov.hk/index.php/censos-y-encuestas/encuestas-todos-las-encuestas-de-honduras/encuesta-permanente-de-hogares>
- Robins, J., & Rotnitzky, A. (1995). Semiparametric efficiency in multivariate regression models with missing data. *Journal of the American Statistical Association*, 90(429), 122-129.
- Rosenbaum, P., & Rubin, D. (1983). The Central Role of Propensity Score in Observational Studies. *Biometrika*, 41-55.
- WHO, UNICEF, USAID, AED, UCDAVIS, & IFPRI. (2008). *Indicators for assessing infant and young child feeding practices, Part I: Definitions*. Retrieved from <http://www.who.int/nutrition/publications/infantfeeding/9789241596664/en/>

APPENDIX A – DECOMPOSITION OF WOMEN EMPOWERMENT INDEX IN AGRICULTURE

Table 54 Censored headcount from 5DE Index by Survey Year

Censored headcount from 5DE Index		Baseline Value	Follow-Up Value	Endline Value
Domain	Indicator	% of Females achieving indicator	% of Females achieving indicator	% of Females achieving indicator
PRODUCTION	Input in productive decisions	88.43%	72.69%	67.04%
	Autonomy in production	92.42%	75.80%	89.59%
RESOURCES	Ownership of assets	88.70%	87.95%	87.50%
	Purchase, sale, or transfer of assets	50.41%	55.20%	55.10%
	Access to and decisions on credit	33.96%	34.50%	38.67%
INCOME	Control over use of income	66.71%	83.48%	74.96%
LEADERSHIP	Group member	67.05%	60.84%	73.83%
	Speaking in public	74.09%	78.72%	71.34%
TIME	Workload	74.56%	73.96%	81.23%
	Leisure	89.31%	89.12%	96.91%

Table 55 SDE Decomposed by Dimension and Indicator by Survey Year

	Statistics	PRODUCTION		RESOURCES			INCOME	LEADERSHIP		TIME	
		INPUT IN PRODUCTIVE DECISIONS	AUTONOMY IN PRODUCTION	OWNERSHIP	PURCHASE, SALE OR TRANSFER OF ASSETS	ACCESS TO AND DECISIONS ON CREDIT	CONTROL OVER USE OF INCOME	GROUP MEMBERSHIP	SPEAK IN PUBLIC	WORKLOAD BURDEN	LEISURE TIME
2012	WOMEN - 2012										
	Censored headcount	88.4%	92.4%	88.7%	50.4%	34.0%	66.7%	67.1%	74.1%	74.6%	89.3%
	% Contribution	1.6%	1.0%	1.0%	4.5%	6.0%	9.1%	4.5%	3.5%	3.5%	1.5%
	Contribution	0.088	0.092	0.059	0.034	0.023	0.133	0.067	0.074	0.075	0.089
	% Contr. by dimension		2.6%		11.5%		9.1%	8.0%		4.9%	
	MEN - 2012										
	Censored headcount	83.5%	94.0%	93.3%	74.5%	62.1%	98.1%	77.1%	88.2%	92.4%	91.4%
	% Contribution	1.9%	0.7%	0.5%	1.9%	2.9%	0.4%	2.6%	1.3%	0.9%	1.0%
2013	Contribution	0.084	0.094	0.062	0.050	0.041	0.196	0.077	0.088	0.092	0.091
	% Contr. by dimension		2.6%		5.3%		0.4%	4.0%		1.9%	
	WOMEN - 2013										
	Censored headcount	72.7%	75.8%	87.9%	55.2%	34.5%	83.5%	60.8%	78.7%	74.0%	89.1%
	% Contribution	3.7%	3.3%	1.1%	4.1%	5.9%	4.5%	5.3%	2.9%	3.5%	1.5%
	Contribution	0.073	0.076	0.059	0.037	0.023	0.167	0.061	0.079	0.074	0.089
	% Contr. by dimension		7.0%		11.1%		4.5%	8.2%		5.0%	
	MEN - 2013										
2015	Censored headcount	98.1%	84.7%	92.5%	79.0%	67.5%	98.7%	79.0%	95.4%	93.1%	94.5%
	% Contribution	0.2%	1.7%	0.6%	1.6%	2.4%	0.3%	2.3%	0.5%	0.8%	0.6%
	Contribution	0.098	0.085	0.062	0.053	0.045	0.197	0.079	0.095	0.093	0.094
	% Contr. by dimension		1.9%		4.5%		0.3%	2.8%		1.4%	
	WOMEN - 2015										
	Censored headcount	67.0%	89.6%	87.5%	55.1%	38.7%	75.0%	73.8%	71.3%	81.2%	96.9%
	% Contribution	1.6%	1.0%	1.0%	4.5%	6.0%	9.1%	4.5%	3.5%	3.5%	1.5%
	Contribution	0.067	0.090	0.058	0.037	0.026	0.150	0.074	0.071	0.081	0.097
2015	% Contr. by dimension		2.6%		11.5%		9.1%	8.0%		4.9%	
	MEN - 2015										
	Censored headcount	94.2%	93.7%	94.7%	85.6%	74.9%	98.4%	83.7%	91.2%	94.0%	98.0%
	% Contribution	1.9%	0.7%	0.5%	1.9%	2.9%	0.4%	2.6%	1.3%	0.9%	1.0%
	Contribution	0.094	0.094	0.063	0.057	0.050	0.197	0.084	0.091	0.094	0.098
	% Contr. by dimension		2.6%		5.3%		0.4%	4.0%		1.9%	

Table 56 Censored headcount from 5DE Index by Treatment Status in 2013

		2013		2015	
Censored headcount from 5DE Index		Treatment	Comparison	Treatment	Comparison
Domain	Indicator	% of Females achieving indicator	% of Females achieving indicator	% of Females achieving indicator	% of Females achieving indicator
PRODUCTION	Input in productive decisions	74.16%	71.26%	68.67%	65.89%
	Autonomy in production	76.73%	74.90%	88.85%	90.12%
RESOURCES	Ownership of assets	89.15%	86.78%	89.52%	86.07%
	Purchase, sale, or transfer of assets	59.96%	50.57%	63.43%	49.21%
	Access to and decisions on credit	36.88%	32.18%	45.88%	33.56%
INCOME	Control over use of income	82.05%	84.87%	74.59%	75.22%
LEADERSHIP	Group member	64.89%	56.90%	76.92%	71.65%
	Speaking in public	77.51%	79.89%	74.59%	69.05%
TIME	Workload	73.37%	74.52%	80.50%	81.74%
	Leisure	88.17%	90.04%	97.38%	96.57%

Table 57 Decomposed by Dimension and Indicator by Treatment Status in 2013 and 2015

		Statistics	PRODUCTION		RESOURCES		INCOME	LEADERSHIP		TIME		
			INPUT IN PRODUCTIVE DECISIONS	AUTONOMY IN PRODUCTION	OWNERSHIP	PURCHASE, SALE OR TRANSFER OF ASSETS	ACCESS TO AND DECISIONS ON CREDIT	CONTROL OVER USE OF INCOME	GROUP MEMBERSHIP	SPEAK IN PUBLIC	WORKLOAD BURDEN	LEISURE TIME
2013	Treatment-2013	WOMEN - Treatment-2013										
		Censored headcount	74.2%	76.7%	89.2%	60.0%	36.9%	82.1%	64.9%	77.5%	73.4%	88.2%
		% Contribution	3.5%	3.1%	1.0%	3.6%	5.7%	4.8%	4.7%	3.0%	3.6%	1.6%
		Contribution	0.074	0.077	0.059	0.040	0.025	0.164	0.065	0.078	0.073	0.088
		% Contr. by dimension		6.6%		10.2%		4.8%	7.8%			5.2%
		MEN - Treatment-2013										
	Comparison-2013	Censored headcount	97.9%	86.0%	93.4%	83.9%	72.0%	99.0%	81.8%	96.5%	93.4%	94.1%
		% Contribution	0.2%	1.5%	0.5%	1.2%	2.0%	0.2%	2.0%	0.4%	0.7%	0.7%
		Contribution	0.098	0.086	0.062	0.056	0.048	0.198	0.082	0.097	0.093	0.094
		% Contr. by dimension		1.8%		3.7%		0.2%	2.4%			1.4%
		WOMEN - Comparison-2013										
		Censored headcount	71.3%	74.9%	86.8%	50.6%	32.2%	84.9%	56.9%	79.9%	74.5%	90.0%
2015	Treatment-2015	% Contribution	3.9%	3.4%	1.2%	4.5%	6.2%	4.1%	5.9%	2.8%	3.5%	1.4%
		Contribution	0.071	0.075	0.058	0.034	0.021	0.170	0.057	0.080	0.075	0.090
		% Contr. by dimension		7.4%		11.9%		4.1%	8.7%			4.9%
		MEN - Comparison-2013										
		Censored headcount	98.3%	83.1%	91.6%	73.0%	62.0%	98.3%	75.5%	94.1%	92.8%	94.9%
		% Contribution	0.2%	1.9%	0.6%	2.0%	2.9%	0.4%	2.8%	0.7%	0.8%	0.6%
	Comparison-2015	Contribution	0.098	0.083	0.061	0.049	0.041	0.197	0.076	0.094	0.093	0.095
		% Contr. by dimension		2.1%		5.5%		0.4%	3.4%			1.4%
		WOMEN - Treatment-2015										
		Censored headcount	68.7%	88.8%	89.5%	63.4%	45.9%	74.6%	76.9%	74.6%	80.5%	97.4%
		% Contribution	4.1%	1.5%	0.9%	3.2%	4.7%	6.6%	3.0%	3.3%	2.5%	0.3%
		Contribution	0.069	0.089	0.060	0.042	0.031	0.149	0.077	0.075	0.081	0.097
	Comparison-2015	% Contr. by dimension		5.5%		8.8%		6.6%	6.3%			2.9%
		MEN - Treatment-2015										
		Censored headcount	96.8%	95.3%	95.6%	90.2%	82.0%	98.8%	88.4%	94.3%	95.1%	98.7%
		% Contribution	0.3%	0.5%	0.3%	0.7%	1.3%	0.2%	1.2%	0.6%	0.5%	0.1%
		Contribution	0.097	0.095	0.064	0.060	0.055	0.198	0.088	0.094	0.095	0.099
		% Contr. by dimension		0.8%		2.3%		0.2%	1.8%			0.7%
		WOMEN - Comparison-2015										
		Censored headcount	65.9%	90.1%	86.1%	49.2%	33.6%	75.2%	71.7%	69.0%	81.7%	96.6%
		% Contribution	4.6%	1.3%	1.3%	4.6%	6.0%	6.7%	3.8%	4.2%	2.5%	0.5%
		Contribution	0.066	0.090	0.057	0.033	0.022	0.150	0.072	0.069	0.082	0.097
	% Contr. by dimension		6.0%		11.8%		6.7%	8.0%			2.9%	
	Comparison-2015	MEN - Comparison-2015										
		Censored headcount	92.0%	92.3%	93.9%	81.8%	69.0%	98.1%	79.8%	88.6%	93.2%	97.4%
		% Contribution	0.9%	0.8%	0.4%	1.3%	2.3%	0.4%	2.2%	1.3%	0.8%	0.3%
		Contribution	0.092	0.092	0.063	0.055	0.046	0.196	0.080	0.089	0.093	0.097
		% Contr. by dimension		1.7%		4.1%		0.4%	3.5%			1.0%

APPENDIX B – IMPACT ESTIMATES: OTHER OUTCOMES

Table 58 Impact on household expenditures

	(1)	(2)	(3)	(4)
Annual HH Food Expenditures in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-42 [64.3]	-41.6 [64.4]		
Treatment x 2015- Endline	-33.2 [111.4]	-36.4 [114.1]		
Treatment x Post- 2013 & 2015			-39.3 [81.9]	-41 [84.6]
Mean of Comp. at Baseline	1894.6	1875	1894.6	1875
SD of Comp. at Baseline	937.7	932.4	937.7	932.4
Observations	9055	9055	9055	9055
Annual HH Non-Food Expenditures in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-144.9 [108.4]	-152.8 [113.5]		
Treatment x 2015- Endline	-172 [233.8]	-185.3 [245.3]		
Treatment x Post- 2013 & 2015			-158.9 [168.6]	-169.9 [178.5]
Mean of Comp. at Baseline	718.2	708.7	718.2	708.7
SD of Comp. at Baseline	709.3	704.6	709.3	704.6
Observations	9055	9055	9055	9055
Annual Rent Expenditures in 2005 PPP 2010 USD				
Treatment x 2013- Midline	30.7 [48.9]	34.8 [49.8]		
Treatment x 2015- Endline	116.6 [76.9]	117.2 [80.6]		
Treatment x Post- 2013 & 2015			74.3 [59.1]	77.4 [62.2]
Mean of Comp. at Baseline	773.1	771.9	773.1	771.9
SD of Comp. at Baseline	748	751.2	748	751.2
Observations	9055	9055	9055	9055
Annual Utilities Expenditures in 2005 PPP 2010 USD				
Treatment x 2013- Midline	2.22 [10.4]	2.02 [10.3]		
Treatment x 2015- Endline	-4.73 [14.1]	-4.86 [14.6]		
Treatment x Post- 2013 & 2015			-1.05 [10.4]	-1.23 [10.7]
Mean of Comp. at Baseline	439.2	435.6	439.2	435.6
SD of Comp. at Baseline	229.2	227.7	229.2	227.7
Observations	9055	9055	9055	9055
Annual Housing Related Expenditure in 2005 PPP 2010 USD				
Treatment x 2013- Midline	32.9 [48.4]	36.8 [48.9]		
Treatment x 2015- Endline	111.9 [69.3]	112.4 [72.2]		
Treatment x Post- 2013 & 2015			73.2 [54.3]	76.1 [56.7]
Mean of Comp. at Baseline	1212.3	1207.5	1212.3	1207.5
SD of Comp. at Baseline	837.5	840.4	837.5	840.4
Number of Clusters	301	301	301	301
Number of Households	3322	3322	3322	3322
Observations	9055	9055	9055	9055

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact

* p<0.10, ** p<0.05, *** p<0.01

Table 59 Impact on diarrhea in children under 5

Diarrhea in last 15 days	[1]	[2]	[3]	[4]
Treatment x 2013- Midline	0.04 [0.044]	0.04 [0.045]		
Treatment x 2015- Endline	0.14 [0.089]	0.14 [0.092]		
Treatment x Post- 2013 & 2015			0.078 [0.060]	0.08 [0.063]
Mean of Comp. at Baseline	1.81	1.81	1.81	1.81
SD of Comp. at Baseline	0.42	0.42	0.42	0.42
Observations	4886	4886	4886	4886
Diarrhea in last 30 days				
Treatment x 2013- Midline	0.0073 [0.028]	0.0082 [0.028]		
Treatment x 2015- Endline	-0.027 [0.032]	-0.024 [0.033]		
Treatment x Post- 2013 & 2015			-0.0058 [0.028]	-0.0042 [0.028]
Mean of Comp. at Baseline	1.93	1.93	1.93	1.93
SD of Comp. at Baseline	0.32	0.33	0.32	0.33
Number of Clusters	269	269	269	269
Number of Households	1793	1793	1793	1793
Observations	4066	4066	4066	4066

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the individual level and all regressions include household fixed effects, and controls for the number of children under five, the number of observations per household and sex. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 60 Impact on feeding practices for children 6-23 months

Milk feedings for non-breastfed	[1]	[2]	[3]	[4]
Treatment x 2013- Midline	-1.72 [1.48]	-1.44 [1.38]		
Treatment x 2015- Endline	-0.87 [1.31]	-0.7 [1.25]		
Treatment x Post- 2013 & 2015			-1.35 [1.36]	-1.09 [1.27]
Mean of Comp. at Baseline	1.04	1.05	1.04	1.05
SD of Comp. at Baseline	1.35	1.37	1.35	1.37
Observations	480	480	480	480
Feeds = Milk+Solids for non-breastfed				
Treatment x 2013- Midline	-0.35 [1.20]	-0.2 [1.15]		
Treatment x 2015- Endline	1.43 [1.66]	1.47 [1.63]		
Treatment x Post- 2013 & 2015			0.067 [1.25]	0.21 [1.20]
Mean of Comp. at Baseline	3.45	3.46	3.45	3.46
SD of Comp. at Baseline	1.3	1.35	1.3	1.35
Number of Clusters	157	157	157	157
Number of Households	377	377	377	377
Observations	436	436	436	436

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the individual level and all regressions include household fixed effects, and controls for the number of children under five, the number of observations per household and sex. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 61 Impact on birth control use

Birth Control used	[1]	[2]	[3]	[4]
Treatment x 2013- Midline	0.036 [0.025]	0.036 [0.025]		
Treatment x 2015- Endline	0.03 [0.027]	0.032 [0.027]		
Treatment x Post- 2013 & 2015			0.033 [0.022]	0.034 [0.023]
Mean of Comp. at Baseline	0.29	0.3	0.29	0.3
SD of Comp. at Baseline	0.46	0.46	0.46	0.46
Number of Clusters	288	288	288	288
Number of Households	2783	2783	2783	2783
Observations	7495	7495	7495	7495

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the individual level and all regressions include household fixed effects, and controls for the number of women age 15 to 49 and the number of observations per household and age. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 62 Impact on female child anthropometric measures

Females: Length/height-for-age Z-score	[1]	[2]	[3]	[4]
Treatment x 2013- Midline	0.12 [0.22]	0.09 [0.22]		
Treatment x 2015- Endline	1.94 [1.78]	1.7 [1.63]		
Treatment x Post- 2013 & 2015			0.73 [0.69]	0.63 [0.64]
Mean of Comp. at Baseline	-1.3	-1.31	-1.3	-1.31
SD of Comp. at Baseline	1.89	1.93	1.89	1.93
Observations	1523	1523	1523	1523
Females: Weight-for-age Z-score				
Treatment x 2013- Midline	-0.15 [0.13]	-0.15 [0.13]		
Treatment x 2015- Endline	-0.26 [0.17]	-0.24 [0.17]		
Treatment x Post- 2013 & 2015			-0.19 [0.12]	-0.18 [0.12]
Mean of Comp. at Baseline	-0.58	-0.57	-0.58	-0.57
SD of Comp. at Baseline	1.28	1.29	1.28	1.29
Observations	1482	1482	1482	1482
Females: Weight-for-length/height Z-score				
Treatment x 2013- Midline	-0.28 [0.20]	-0.24 [0.19]		
Treatment x 2015- Endline	-0.15 [0.29]	-0.057 [0.28]		
Treatment x Post- 2013 & 2015			-0.23 [0.21]	-0.18 [0.21]
Mean of Comp. at Baseline	0.19	0.21	0.19	0.21
SD of Comp. at Baseline	1.03	1.02	1.03	1.02
Observations	1472	1472	1472	1472
Females: BMI-for-age Z-score				
Treatment x 2013- Midline	-0.14 [0.45]	-0.12 [0.44]		
Treatment x 2015- Endline	0.15 [0.84]	0.27 [0.86]		
Treatment x Post- 2013 & 2015			-0.041 [0.42]	0.016 [0.42]
Mean of Comp. at Baseline	0.39	0.41	0.39	0.41
SD of Comp. at Baseline	1.2	1.22	1.2	1.22
Number of Clusters	224	224	224	224
Number of Households	926	926	926	926
Observations	1485	1485	1485	1485

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. The outcomes variable represent the proportion/average for the children (girls) in the household. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 63 Impact on female child wasting, stunting and underweight indicators

Females: Indicator for child being underweight	[1]	[2]	[3]	[4]
Treatment x 2013- Midline	0.0094 [0.035]	0.0066 [0.035]		
Treatment x 2015- Endline	0.034 [0.066]	0.025 [0.069]		
Treatment x Post- 2013 & 2015			0.018 [0.032]	0.013 [0.034]
Mean of Comp. at Baseline	0.13	0.13	0.13	0.13
SD of Comp. at Baseline	0.33	0.33	0.33	0.33
Observations	1496	1496	1496	1496
Females: Child has Height-for-Age z-score < -2				
Treatment x 2013- Midline	-0.11 [0.12]	-0.1 [0.12]		
Treatment x 2015- Endline	-0.089 [0.11]	-0.074 [0.10]		
Treatment x Post- 2013 & 2015			-0.1 [0.11]	-0.092 [0.11]
Mean of Comp. at Baseline	0.32	0.33	0.32	0.33
SD of Comp. at Baseline	0.45	0.45	0.45	0.45
Observations	1535	1535	1535	1535
Females: Child has Weight-for-Height z-score < -2				
Treatment x 2013- Midline	0.016 [0.027]	0.014 [0.027]		
Treatment x 2015- Endline	0.042 [0.033]	0.032 [0.034]		
Treatment x Post- 2013 & 2015			0.025 [0.023]	0.02 [0.023]
Mean of Comp. at Baseline	0.026	0.024	0.026	0.024
SD of Comp. at Baseline	0.15	0.15	0.15	0.15
Number of Clusters	224	224	224	224
Number of Households	930	930	930	930
Observations	1495	1495	1495	1495

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. The outcomes variable represent the proportion/average for the children (girls) in the household. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 64 Impact on female child anemia

Females: HEMOGLOBIN G/DL	[1]	[2]	[3]	[4]
Treatment x 2013- Midline	-0.037 [0.23]	-0.0068 [0.23]		
Treatment x 2015- Endline	0.39 [0.34]	0.4 [0.34]		
Treatment x Post- 2013 & 2015			0.1 [0.25]	0.13 [0.25]
Mean of Comp. at Baseline	11.7	11.7	11.7	11.7
SD of Comp. at Baseline	1.14	1.15	1.14	1.15
Observations	1388	1388	1388	1388
Females: Indicator for child anemia (6-59 months)				
Treatment x 2013- Midline	0.14 [0.11]	0.13 [0.10]		
Treatment x 2015- Endline	0.069 [0.089]	0.061 [0.090]		
Treatment x Post- 2013 & 2015			0.12 [0.096]	0.11 [0.095]
Mean of Comp. at Baseline	0.22	0.22	0.22	0.22
SD of Comp. at Baseline	0.4	0.4	0.4	0.4
Observations	1388	1388	1388	1388
Females: Child has edema				
Treatment x 2013- Midline	0.0065 [0.029]	0.0057 [0.029]		
Treatment x 2015- Endline	-0.013 [0.042]	-0.012 [0.042]		
Treatment x Post- 2013 & 2015			-0.000039 [0.028]	-0.00026 [0.028]
Mean of Comp. at Baseline	0.028	0.029	0.028	0.029
SD of Comp. at Baseline	0.17	0.17	0.17	0.17
Number of Clusters	225	225	225	225
Number of Households	953	953	953	953
Observations	1536	1536	1536	1536

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. The outcomes variable represent the proportion/average for the children (girls) in the household. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 65 Impact on male child anthropometric measures

Males: Length/height-for-age Z-score	[1]	[2]	[3]	[4]
Treatment x 2013- Midline	0.066 [0.23]	0.084 [0.23]		
Treatment x 2015- Endline	-0.26 [0.40]	-0.23 [0.40]		
Treatment x Post- 2013 & 2015			-0.047 [0.24]	-0.025 [0.24]
Mean of Comp. at Baseline	-1.96	-1.98	-1.96	-1.98
SD of Comp. at Baseline	1.87	1.87	1.87	1.87
Observations	1560	1560	1560	1560
Males: Weight-for-age Z-score				
Treatment x 2013- Midline	0.18 [0.15]	0.19 [0.15]		
Treatment x 2015- Endline	-0.007 [0.36]	0.019 [0.36]		
Treatment x Post- 2013 & 2015			0.11 [0.19]	0.12 [0.18]
Mean of Comp. at Baseline	-0.89	-0.91	-0.89	-0.91
SD of Comp. at Baseline	1.36	1.37	1.36	1.37
Observations	1528	1528	1528	1528
Males: Weight-for-length/height Z-score				
Treatment x 2013- Midline	0.24 [0.17]	0.22 [0.17]		
Treatment x 2015- Endline	0.36 [0.22]	0.34 [0.22]		
Treatment x Post- 2013 & 2015			0.28 [0.15]*	0.27 [0.15]*
Mean of Comp. at Baseline	0.12	0.11	0.12	0.11
SD of Comp. at Baseline	1.13	1.13	1.13	1.13
Observations	1507	1507	1507	1507
Males: BMI-for-age Z-score				
Treatment x 2013- Midline	1.31 [1.38]	1.3 [1.39]		
Treatment x 2015- Endline	4.03 [4.69]	4.19 [4.79]		
Treatment x Post- 2013 & 2015			2.23 [2.51]	2.29 [2.56]
Mean of Comp. at Baseline	0.5	0.49	0.5	0.49
SD of Comp. at Baseline	1.63	1.62	1.63	1.62
Number of Clusters	222	222	222	222
Number of Households	946	946	946	946
Observations	1521	1521	1521	1521

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. The outcomes variable represent the proportion/average for the children (boys) in the household. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 66 Impact on male child wasting, stunting and underweight indicators

Males: Indicator for child being underweight	[1]	[2]	[3]	[4]
Treatment x 2013- Midline	-0.042 [0.037]	-0.041 [0.038]		
Treatment x 2015- Endline	-0.15 [0.066]**	-0.15 [0.067]**		
Treatment x Post- 2013 & 2015			-0.079 [0.038]**	-0.079 [0.040]**
Mean of Comp. at Baseline	0.17	0.18	0.17	0.18
SD of Comp. at Baseline	0.37	0.37	0.37	0.37
Observations	1542	1542	1542	1542
Males: Child has Height-for-Age z-score < -2				
Treatment x 2013- Midline	0.029 [0.080]	0.026 [0.080]		
Treatment x 2015- Endline	0.067 [0.091]	0.069 [0.089]		
Treatment x Post- 2013 & 2015			0.043 [0.077]	0.042 [0.076]
Mean of Comp. at Baseline	0.46	0.46	0.46	0.46
SD of Comp. at Baseline	0.48	0.48	0.48	0.48
Observations	1573	1573	1573	1573
Males: Child has Weight-for-Height z-score < -2				
Treatment x 2013- Midline	-0.018 [0.029]	-0.013 [0.030]		
Treatment x 2015- Endline	-0.073 [0.048]	-0.065 [0.050]		
Treatment x Post- 2013 & 2015			-0.038 [0.028]	-0.032 [0.030]
Mean of Comp. at Baseline	0.041	0.044	0.041	0.044
SD of Comp. at Baseline	0.19	0.2	0.19	0.2
Number of Clusters	222	222	222	222
Number of Households	959	959	959	959
Observations	1539	1539	1539	1539

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. The outcomes variable represent the proportion/average for the children (boys) in the household. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 67 Impact on male child anemia

Males: HEMOGLOBIN G/DL	[1]	[2]	[3]	[4]
Treatment x 2013- Midline	0.19 [0.15]	0.2 [0.15]		
Treatment x 2015- Endline	0.0092 [0.32]	0.036 [0.33]		
Treatment x Post- 2013 & 2015			0.13 [0.17]	0.15 [0.18]
Mean of Comp. at Baseline	11.6	11.6	11.6	11.6
SD of Comp. at Baseline	1.11	1.12	1.11	1.12
Observations	1409	1409	1409	1409
Males: Indicator for child anemia (6-59 months)				
Treatment x 2013- Midline	-0.1 [0.057]*	-0.11 [0.057]*		
Treatment x 2015- Endline	0.051 [0.10]	0.044 [0.10]		
Treatment x Post- 2013 & 2015			-0.05 [0.061]	-0.054 [0.060]
Mean of Comp. at Baseline	0.26	0.26	0.26	0.26
SD of Comp. at Baseline	0.42	0.43	0.42	0.43
Observations	1409	1409	1409	1409
Males: Child has edema				
Treatment x 2013- Midline	0.017 [0.016]	0.018 [0.015]		
Treatment x 2015- Endline	-0.06 [0.040]	-0.061 [0.041]		
Treatment x Post- 2013 & 2015			-0.011 [0.020]	-0.011 [0.020]
Mean of Comp. at Baseline	0.027	0.026	0.027	0.026
SD of Comp. at Baseline	0.15	0.15	0.15	0.15
Number of Clusters	226	226	226	226
Number of Households	987	987	987	987
Observations	1577	1577	1577	1577

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. The outcomes variable represent the proportion/average for the children (boys) in the household. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights.

* p<0.10, ** p<0.05, *** p<0.01

Table 68 Impact on log income and cost by type of work: principal activity

	[1]	[2]	[3]	[4]
aSinH-Log - Annual income for -Wage-Labor - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	0.31 [0.36]	0.32 [0.36]		
Treatment x 2015- Endline	0.45 [0.35]	0.45 [0.35]		
Treatment x Post- 2013 & 2015			0.38 [0.29]	0.38 [0.29]
Mean of Comp. at Baseline	4.96	5.03	4.96	5.03
SD of Comp. at Baseline	4.28	4.27	4.28	4.27
Observations	5686	5686	5686	5686
aSinH-Log - Annual income for -Merchant - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	0.15 [0.62]	0.19 [0.64]		
Treatment x 2015- Endline	0.45 [0.62]	0.48 [0.63]		
Treatment x Post- 2013 & 2015			0.3 [0.54]	0.33 [0.55]
Mean of Comp. at Baseline	2.98	2.99	2.98	2.99
SD of Comp. at Baseline	4.12	4.13	4.12	4.13
Observations	1531	1531	1531	1531
aSinH-Log - Annual income for -Agriculture - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-0.25 [0.35]	-0.23 [0.35]		
Treatment x 2015- Endline	-0.13 [0.34]	-0.11 [0.34]		
Treatment x Post- 2013 & 2015			-0.19 [0.30]	-0.17 [0.30]
Mean of Comp. at Baseline	5.25	5.33	5.25	5.33
SD of Comp. at Baseline	4.49	4.49	4.49	4.49
Observations	6345	6345	6345	6345
aSinH-Log - Annual income for -Production-Processing - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-1.06 [0.69]	-0.96 [0.70]		
Treatment x 2015- Endline	-0.4 [0.82]	-0.44 [0.85]		
Treatment x Post- 2013 & 2015			-0.71 [0.66]	-0.68 [0.67]
Mean of Comp. at Baseline	2.38	2.42	2.38	2.42
SD of Comp. at Baseline	3.68	3.7	3.68	3.7
Observations	735	735	735	735
aSinH-Log - Annual income for -Professional - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-0.65 [0.57]	-0.66 [0.57]		
Treatment x 2015- Endline	-0.36 [0.75]	-0.32 [0.75]		
Treatment x Post- 2013 & 2015			-0.52 [0.58]	-0.51 [0.59]
Mean of Comp. at Baseline	3.9	3.99	3.9	3.99
SD of Comp. at Baseline	4.53	4.54	4.53	4.54
Number of Clusters	165	165	165	165
Number of Households	636	636	636	636
Observations	1720	1720	1720	1720

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using

* p<0.10, ** p<0.05, *** p<0.01

APPENDIX C – SAMPLE METHODOLOGY

SAMPLE DESIGN

The sample designs to be used for these surveys are called “analytical survey designs,” in contrast to the “descriptive survey designs” used for most sample surveys. The purpose of a descriptive survey is to describe the characteristics of a population of interest and various subpopulations. The purpose of an analytical survey is to develop detailed analytical models that describe the relationship of dependent variables to a variety of explanatory (independent) variables (including treatment variables). The principal objective of these surveys is to provide high precision for estimates of the double-difference estimate of program impact; i.e. the difference between the treatment and comparison groups, such as the difference in income between the beginning and end of the evaluation period.

We have balanced the purpose of these two types of surveys in our calculation of sample size, given that representativeness at the department level is desired and that under these constraints, we need to increase the precision of our impact estimates as much as possible. There are two major analysis objectives for the survey data. The first, and primary, objective is to obtain precise estimates of program impact, by means of double-difference estimates. The second is to investigate the relationship of program impact to a variety of explanatory variables, or “covariates.” The main evaluation research design used is a “pretest-posttest control-group design” with matching. In order to achieve the research objectives, we construct sample survey designs that (1) have substantial variation in the explanatory variables; (2) have a high degree of orthogonality among explanatory variables (that do not have a similar relationship to the dependent variable); and (3) have a high degree of correlation between members of the treatment and comparison groups (preferably by matching on a one-to-one basis and secondarily by matching of probability distributions). Note that we are concerned here with selection of the sample of PSUs (caseríos or aldeas), not households within PSUs.

For the proposed designs, the probabilities of selection will be determined so that the expected number of sample items (PSUs) in various categories will be close to desired levels. Those levels will be determined to achieve desired levels of variation and orthogonality among explanatory variables, subject to the requirement that all probabilities of selection are known and non-zero (so that all population-of-interest units are subject to sampling) and the total sample size is as desired.

This procedure will produce a representative sample at the department level for the baseline survey using the most complete sample frame available, the 2001 census. The survey will be representative of prevalence of poverty and/or incidence of undernourishment in rural areas for each department where ACCESO is being implemented. A random over-sample of the beneficiaries will be added to the sample to improve the power of the sampling design in detecting the expected changes in the target population.

Given that the list of beneficiaries is incomplete, we will likely increase the oversampling of beneficiaries for the endline survey (and potentially the controls), effectively increasing the size of the treatment group and allowing us to explore possible heterogeneous effects that depend on the time that beneficiaries are exposed to the program.

The multistage stratified sampling design with an over-sample will be as follows:

- 1) Stratify the sample frame by department.
- 2) Select the number of primary sampling units (PSU) necessary for the sample to be representative of the department¹⁸ by distributing the PSUs across urban-rural strata, and use demographic variables available in the census to balance the sample across gender, age groups, etc.
- 3) Randomly over-sample ACCESO beneficiary households. Using the list of beneficiaries, we will randomly select beneficiary PSUs and, within PSUs, select the beneficiary households that will form the treatment group. These households will be selected from the validated list provided by Fintrac (i.e. list of beneficiaries that comply with the poverty criteria).
- 4) Randomly select control PSUs around but far enough away from beneficiary PSUs to ensure a pool of controls that will not be treated by the end date of ACCESO.
- 5) Randomly select households within each PSU in the control group.
- 6) Calculate appropriate weights to adjust for the different probability of selection between treatment and controls.

¹⁸ Potentially we could further stratify the sample between treated PSUs and non-treated PSUs which could improve the estimates across these groups. This will depend on the number of PSU that remain as non-treated, given that Acceso is spread out geographically (see

Figure 12), there is the possibility that few PSU's have no beneficiaries.

The final sample size will also account for attrition among surveyed households, given that we will be implementing a three-round panel survey. In addition, this will allow us to follow a random sub-sample of the treatment households to ensure that implementation efforts are not affected by the evaluation, in the sense that more effort could be placed by the implementers among the households that are known to be in the evaluation sample to the detriment of other beneficiary households.

To arrive to a sample size that will comply with the above-mentioned requirements, we first simulate the effect of 30,000 households¹⁹ exiting poverty under various scenarios to be able to bound the expected impact of the program and to ensure that the sample is large enough to detect these impacts.

Scenario 1

Uses the published poverty rates in each department²⁰ and calculates the change that would be observed after the exit of these households under the assumptions of:

- Number of households that exit poverty is proportional to the poverty share of the department

Scenario 2

Uses the published poverty rates in each department and calculates the change that would be observed after the exit of these households under the assumptions of:

- Number of households that exit poverty is proportional to the number of household participating in ACCESO to date. See

¹⁹ This is the target number of households specified in Fintrac's proposal.

²⁰ All population numbers, poverty rates, and malnutrition rates were obtained from the National Statistical Institute of Honduras (INE-Honduras) and can be found at <http://www.ine.gob.hn/drupal/node/122>

- Figure 12 for the geographic distribution of beneficiary households

Scenario 3

Assumes a 50% poverty rate for all department and calculates the change that would be observed after the exit of these households under the assumptions of:

- Number of households that exit poverty is proportional to the poverty share of the department

Scenario 4

Assumes a 50% poverty rate for all department and calculates the change that would be observed after the exit of these households under the assumptions of:

- Number of households that exit poverty is proportional to the number of household participating in ACCESO to date

These four scenarios are calculated under three more situations:

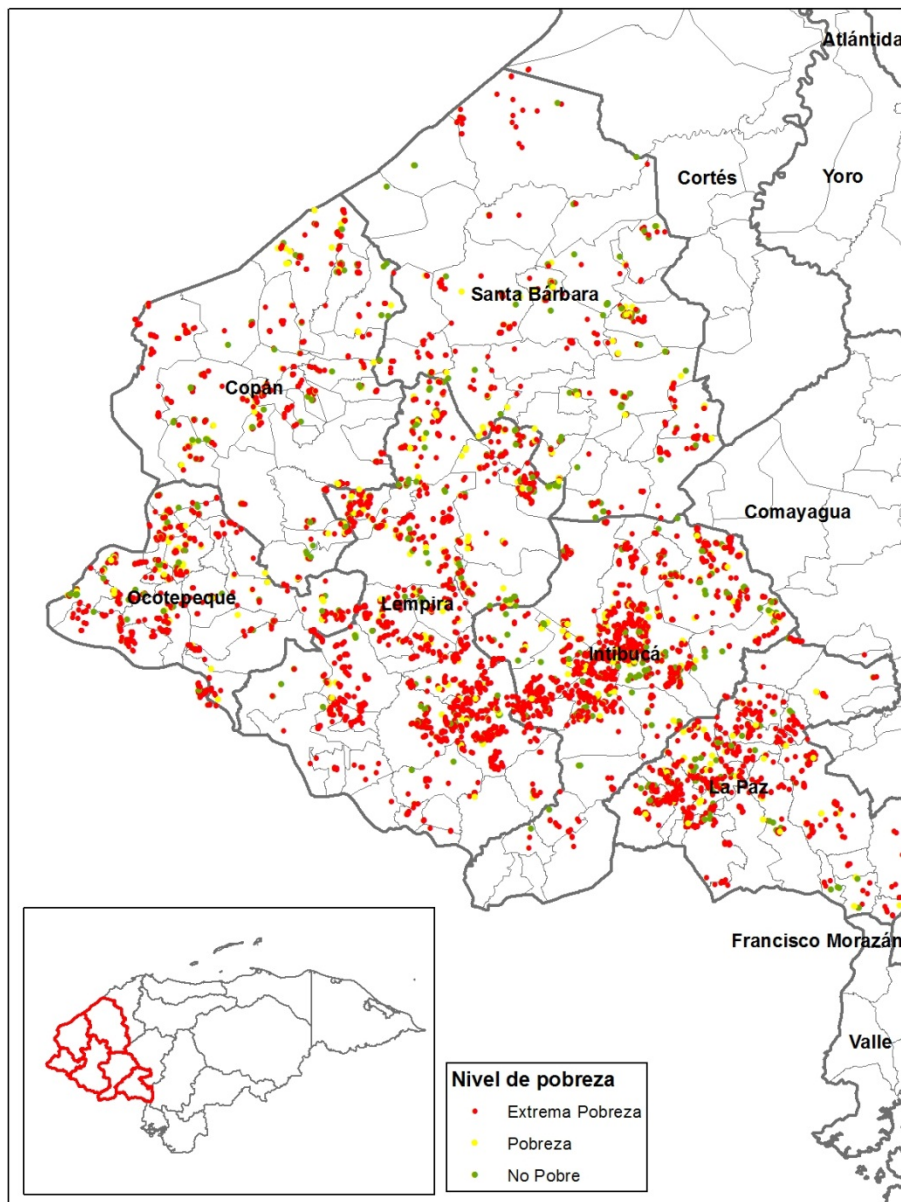
- No growth in population or poverty
- 3% increase in population and poverty
- 5% increase in population and poverty

The scenarios assumptions are summarized in Table 69 , and results are presented in Table 70.

Table 69 Simulation Scenarios

	Poverty	Distribution of impact	No population growth	3% growth	5% growth
Scenario 1	Observed	Proportional to Poverty	✓	✓	✓
Scenario 2	Observed	Proportional to ACCESO	✓	✓	✓
Scenario 3	50%	Proportional to Poverty	✓	✓	✓
Scenario 4	50%	Proportional to ACCESO	✓	✓	✓

Figure 12 Geographical Distribution of Beneficiary Households by Poverty Classification



Across the scenarios, the aggregate effect (for all six departments) of taking 30,000 households out of poverty is a decrease of between 4 and 12.4 percentage points in the poverty prevalence rate. Scenarios 3 and 4 illustrate that under lower levels of poverty prevalence, the effects appear larger; this is because the target of the program is set on the number of households that are taken out of poverty. Thus for lower initial levels, a decrease of 30,000 households represents a higher percentage decrease in the poverty prevalence.

We will discuss Scenarios 1 and 2 in more detail, given that these scenarios are better informed by the data. In scenario 1, the principal assumption is that the beneficiary households are allocated to the program as a function of the share of poverty in each department. For example, the six departments have 199,525 household below the poverty line; 41,566 of these live in Copán, so $41,566/199,525$ or 22% of the beneficiaries are allocated to the department of Copán. This serves us to allocate more beneficiaries in departments where the *number* of poor is greater, rather than where the poverty incidence (%) is greater. First, under an optimistic situation in which there is no increase in population, the effect of the program is observed as a decrease of over 11 percentage points in the prevalence of poverty in each department, with Lempira experiencing the highest change at 13 percentage points. Under a likely²¹ scenario, in which population and poverty increase 3 percent, the aggregate effect is similar to the optimistic situation, with a 11.5 percentage point decrease in the aggregate and over an 11 percentage point at the department level. Under a conservative situation, with 5% growth in population and poverty, the effect decreases to just over 4 percentage points in each department.

In Scenario 2, the principal assumption is that the beneficiary households are allocated to the program in a way that is consistent with the beneficiary list that Fintrac has provided us to date. This serves us to allocate beneficiaries in departments where the *Fintrac* has been more active, rather than where the number of poor is greater. The previous aggregate results do not change, given that the total number of poor that exit poverty remains unchanged; what changes is the distribution of the effects across the departments. First, under an optimistic situation in which there is no increase in poverty, the effect of the program is observed as a decrease in the prevalence of poverty in the range of 6.5 percentage points in Santa Bárbara to 22 percentage points in Ocotepeque. Under a likely scenario, in which population and poverty increase 3 percent, the effects are around 0.25 percentage points lower relative to the optimistic situation. Under a conservative situation, with 5% growth in population and poverty the effects decrease considerably; the range now is between a 15 percentage point *decrease*, in Ocotepeque, and a 2 percentage point *increase*, in Santa Bárbara. This scenario shows us the importance of aggregate factors that might confound what we can observe in the aggregate data. Similarly, a decrease in poverty due to other programs and/or improvements in aggregate/macroeconomic conditions would cause an attribution problem and/or bias the effect toward zero. These issues are addressed in our impact evaluation by the use of household-level data and

²¹ We call it “likely” given that from 2001 to 2010, rural poverty grew at a rate of 3% in Honduras (INE-Honduras)

repeated measures (panel) to estimate the impact, which then can be used to extrapolate the effects to the aggregate.

We present the results of the sample size calculations below. The information from the simulation informs the power analysis that follows, in that it allows us to gauge at what levels the sample will be able to detect the expected changes.

The formulae are the following,

$$n = [1 + \rho(H - 1)] \frac{4Np(1 - p)}{(N - 1)\alpha^2 + 4p(1 - p)}$$

where:

n is the sample size

N is the total population in the department

ρ is de intra-cluster correlation; we use 0.03²²

H is the number of households in a cluster; we use 8 observations per cluster.

p is the proportion of interest, in this case the poverty rate in the department. In addition, note that the function attains a maximum when p is 0.5, thus we also do the exercise in this case.

α is the type 1 error, which is set at 5%

Furthermore, we calculate the minimum detectable effect (MDE), see (Donner & Birkett N, 1981) and (Donner A. , 1998), with the following,

$$MDE = \frac{2\sigma \left(Z_{\frac{\alpha}{2}} + Z_{\beta} \right) \sqrt{[1 + \rho(H - 1)]}}{\sqrt{n}}$$

β is the type 2 error, which is set at 10%, to get a 90% power and Z is the critical value of the normal distribution at the subscript specified.

The results of the sample design are presented in Table 71.. As mentioned previously, we stratify the sample by department and select a representative sample for each department. The results of these calculations are in the columns titled “Calculated”. We then adjust the sample to have the same number

²² Using this value this value the design effect is 1.21. This value is common when calculating sample sizes for income measures in developing countries.

of primary sample units (PSU), by selecting the greatest number of PSUs calculated. Then we increase the sample by 10% to account for attrition and over-sample the treatment group in two ways. First, we select a sample of 10% of the available beneficiary list (this is around 1,200 households), and we add these households to the calculated sample. This gives us a total sample of 570 households per department, with an overall sample size of 3,417 households in 427 PSUs. In the other case, in which we maximize the sample size by setting p to 0.5, we take the adjusted sample and instead of adding our over-sample of treatment households, we split the total number by allocating 200 households in each department to be selected from the beneficiary list and the difference be selected from the sample frame as controls²³. In this case, we obtain a total sample of 3,220 households in 403 PSUs. In both cases, weights need to be used to account for the different probability of selection between treatment and controls when calculating aggregate measures.

The advantage of the first sample is that it can detect smaller differences across the treatment and control groups and is informed by the observed poverty rates. The benefit of the second strategy is that the sample size is maximized, thus allowing for precise estimates across different variables; however, we need to split the sample among the treatment and control to keep the sample size within the budget. In the end, both sample size calculations give us similar results, so that our calculations are not very sensitive to our assumptions. From these sample size calculations, we will use the first results and aim to have a sample of 3,417 households in the baseline survey²⁴. Finally, we present a graphical representation of the sample design in Figure 13.

²³ In the previous case, i.e. using the observed poverty rates, we do not split the sample because the number of control units that remains is too small for a matching procedure.

²⁴ The sample size exercise was similarly done using undernourishment prevalence in each department. These rates are all around 50%, which increase the sample as mentioned below. The results of these calculations are presented in annex 4. In summary, for the split sample we obtain 3,181 households and for the 'additive' sample, we obtain 4,348 households. We note that these nutrition based sample are more precise than needed, since the expected effect for undernourishment is much larger (20% in selected areas expected change in Fintrac's proposal). The sample to detect a 20% decrease in malnutrition is 3,298 (unreported calculations), which is below our preferred size of 3,417, thus we can detect such a change if it were realized.

In addition, we note that the simulations for undernourishment indicate that under optimistic assumptions, 70,000 children under 5 years need to exit undernourishment for such a change to be observed. This could be very challenging.

Figure 13 Graphical Representation of the Sample Design

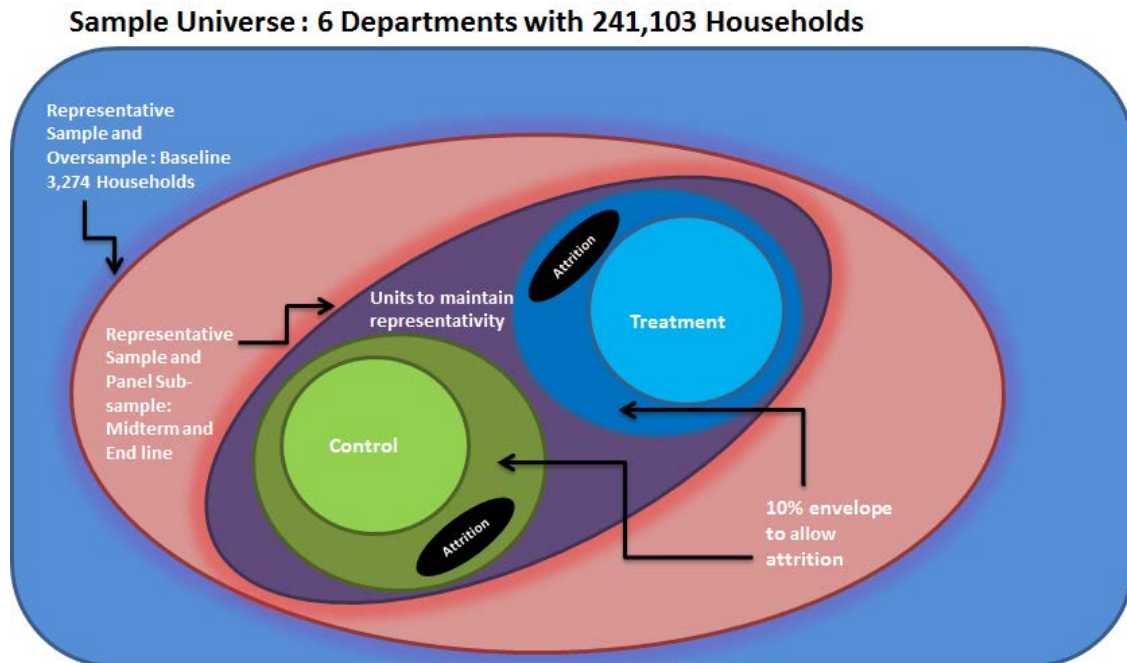


Table 70 Simulation of a decrease of 30,000 households in ACCESO's area of influence

							Optimistic			Likely			Conservative						
							Census				No population change			Assuming 3% poor population growth			Assuming 5% poor population growth		
							Households	% Poor	No. Poor	Shares	Effect	No. Poor	% Poor	% Change	No. Poor	% Poor	% Change	No. Poor	% Poor
Scenario 1																			
Copán	53,168	78%	41,566	21%	-	6,250	35,316	66.4%	-11.8%	36,563	66.8%	-11.4%	37,395	73.8%	-4.3%				
Intibucá	30,503	86%	26,120	13%	-	3,927	22,193	72.8%	-12.9%	22,977	73.1%	-12.5%	23,499	80.9%	-4.7%				
La Paz	27,510	78%	21,526	11%	-	3,237	18,290	66.5%	-11.8%	18,935	66.8%	-11.4%	19,366	73.9%	-4.3%				
Lempira	43,734	86%	37,710	19%	-	5,670	32,040	73.3%	-13.0%	33,171	73.6%	-12.6%	33,925	81.5%	-4.8%				
Ocatepeque	20,161	79%	15,955	8%	-	2,399	13,556	67.2%	-11.9%	14,035	67.6%	-11.6%	14,354	74.8%	-4.4%				
Santa Bárbara	66,027	86%	56,648	28%	-	8,517	48,130	72.9%	-12.9%	49,830	73.3%	-12.5%	50,963	81.0%	-4.8%				
Total	241,103	82%	199,525			-30,000	169,525	70.3%	-11.9%	175,511	70.7%	-11.5%	179,502	78.2%	-4.0%				
Scenario 2																			
Copán	53,168	78%	41,566	16%	-	4,854	36,712	69.0%	-9.1%	37,959	69.3%	-8.9%	38,790	76.6%	-1.6%				
Intibucá	30,503	86%	26,120	17%	-	5,230	20,890	68.5%	-17.1%	21,674	69.0%	-16.6%	22,196	76.4%	-9.2%				
La Paz	27,510	78%	21,526	14%	-	4,179	17,347	63.1%	-15.2%	17,993	63.5%	-14.7%	18,423	70.3%	-7.9%				
Lempira	43,734	86%	37,710	23%	-	7,009	30,701	70.2%	-16.0%	31,832	70.7%	-15.6%	32,586	78.2%	-8.0%				
Ocatepeque	20,161	79%	15,955	15%	-	4,453	11,502	57.1%	-22.1%	11,981	57.7%	-21.4%	12,300	64.1%	-15.1%				
Santa Bárbara	66,027	86%	56,648	14%	-	4,274	52,374	79.3%	-6.5%	54,073	79.5%	-6.3%	55,206	87.8%	2.0%				
Total	241,103	82%	199,525			-30,000	169,525	70.3%	-11.9%	175,511	70.7%	-11.5%	179,502	78.2%	-4.0%				
Scenario 3																			
Copán	53,168	50%	26,584	22%	-	6,616	19,968	37.6%	-12.4%	20,766	37.9%	-12.1%	21,298	42.1%	-7.9%				
Intibucá	30,503	50%	15,252	13%	-	3,795	11,456	37.6%	-12.4%	11,914	37.9%	-12.1%	12,219	42.1%	-7.9%				
La Paz	27,510	50%	13,755	11%	-	3,423	10,332	37.6%	-12.4%	10,745	37.9%	-12.1%	11,020	42.1%	-7.9%				
Lempira	43,734	50%	21,867	18%	-	5,442	16,425	37.6%	-12.4%	17,081	37.9%	-12.1%	17,519	42.1%	-7.9%				
Ocatepeque	20,161	50%	10,081	8%	-	2,509	7,572	37.6%	-12.4%	7,874	37.9%	-12.1%	8,076	42.1%	-7.9%				
Santa Bárbara	66,027	50%	33,014	27%	-	8,216	24,798	37.6%	-12.4%	25,788	37.9%	-12.1%	26,449	42.1%	-7.9%				
Total	241,103	50%	120,552			-30,000	90,552	37.6%	-12.4%	94,168	37.9%	-12.1%	96,579	42.1%	-7.9%				
Scenario 4																			
Copán	53,168	50%	26,584	16%	-	4,854	21,730	40.9%	-9.1%	22,527	41.1%	-8.9%	23,059	45.5%	-4.5%				
Intibucá	30,503	50%	15,252	17%	-	5,230	10,021	32.9%	-17.1%	10,479	33.4%	-16.6%	10,784	37.1%	-12.9%				
La Paz	27,510	50%	13,755	14%	-	4,179	9,576	34.8%	-15.2%	9,988	35.3%	-14.7%	10,263	39.2%	-10.8%				
Lempira	43,734	50%	21,867	23%	-	7,009	14,858	34.0%	-16.0%	15,514	34.4%	-15.6%	15,952	38.3%	-11.7%				
Ocatepeque	20,161	50%	10,081	15%	-	4,453	5,627	27.9%	-22.1%	5,930	28.6%	-21.4%	6,131	31.9%	-18.1%				
Santa Bárbara	66,027	50%	33,014	14%	-	4,274	28,740	43.5%	-6.5%	29,730	43.7%	-6.3%	30,390	48.3%	-1.7%				
Total	241,103	50%	120,552			-30,000	90,552	37.6%	-12.4%	94,168	37.9%	-12.1%	96,579	42.1%	-7.9%				

Table 71 Sample size calculations

			Calculated			Adjusted			Oversample			Final		
Department	Households	Poverty	Households	PSU	MDD %	Households	PSU	MDD %	10% Attrition	Treatment	Controls	Households	PSU	MDD %
<u>Poverty at actual and adding oversample of T</u>														
-														
Copán	53,168	78.4	329	42	16.2%	336	42	16.1%	34	200	336	570	71	12.3%
Intibucá	30,503	83.0	237	30	16.2%	336	42	13.6%	34	200	336	570	71	10.5%
La Paz	27,510	78.2	327	41	16.3%	336	42	16.0%	34	200	336	570	71	12.3%
Lempira	43,734	84.3	229	29	16.2%	336	42	13.4%	34	200	336	570	71	10.3%
Ocatepeque	20,161	78.2	316	40	16.3%	336	42	15.8%	34	200	336	570	71	12.1%
Santa Bárbara	66,027	85.6	236	30	16.2%	336	42	13.6%	34	200	336	570	71	10.4%
Total	241,103	81	1,674	212	6.7%	2,016	252	6.1%	201	1,200	2,016	3,417	427	4.7%
<u>Poverty at 50% and splitting sample T/C</u>														
-														
Copán	53,168	50.0	481	61	16.3%	488	61	16.1%	49	200	288	537	67	15.4%
Intibucá	30,503	50.0	478	60	16.3%	488	61	16.1%	49	200	288	537	67	15.4%
La Paz	27,510	50.0	478	60	16.3%	488	61	16.1%	49	200	288	537	67	15.4%
Lempira	43,734	50.0	480	60	16.3%	488	61	16.1%	49	200	288	537	67	15.4%
Ocatepeque	20,161	50.0	475	60	16.4%	488	61	16.1%	49	200	288	537	67	15.4%
Santa Barbara	66,027	50.0	482	61	16.2%	488	61	16.1%	49	200	288	537	67	15.4%
Total	241,103	50	2,874	362	6.6%	2,928	366	6.6%	292	1,200	1,728	3,220	403	6.3%

POWER ANALYSIS

Given this sample, we explore the power of the design under different types of indicators and minimum effect sizes. We present results using the minimum calculated clusters (212) and our adjusted sample (427). Figure 14 shows the minimum detectable effect for each power level under the assumptions of random assignment of clusters and three observations in time (one baseline and two follow-ups). This figure addresses the impact on continuous outcomes, such as expenditures, that is detectable given our sample. For our preferred sample, we can detect a 12% difference²⁵ between the treatment and control group with 80% power, and 14% difference between the groups with 90% power. For the smaller sample, the detectable difference 18% at 80% power and 20% at 90% power.

In Figure 15,²⁶ we present similar results for the differences detectable at baseline. For our preferred sample, we can detect a 10% difference between the treatment and control group with 80% power, and 12% difference between the groups with 90% power²⁷.

Figure 16 and Figure 17 show the power analysis for binary outcomes, such as the prevalence of poverty in each group. Figure 16 shows the number of clusters (of eight households) necessary to attain a level of power, under the assumptions that the prevalence of poverty in the control group is 82% and that the expected prevalence of poverty in the treatment group is 74%, for a difference of 8%. The figures shows that 450 clusters are needed to attain a 90% power under these assumptions and that 340 clusters are needed to attain an 80% power. Our sample will only have 427 clusters; thus we further explore the level of power we have for binary outcomes with this sample.

Finally we present the power of our sample to detect difference in the poverty prevalence across groups, fixing the poverty rate in the control group at 82% and allowing the expected poverty rate in the treatment group to vary. For our adjusted sample, 73.5% prevalence in the treatment group, for an 8.5%

²⁵ In this section all differences are standardized, i.e. the observed differences are divided by their standard deviations. For example, a 4.7% detectable difference presented in the sample size calculation is approximately equivalent to a 12.2% standard difference.

²⁶ Figures are obtained using software developed by **Invalid source specified**.

²⁷ Power is higher at baseline (given a MDE) because variability of the unconstrained estimate is lower. For example, given one cross section we have that the variance of the difference between the groups is the sum of the variances, call this A; assume that in the follow up the variance is the same, then the variance of the difference in difference is 2A, which is larger than A. This does not imply that more data or panel data decreases the reliability of an estimate; it means that measurement error needs to be addressed in panel studies, and that the benefits of using within household variation and covariates increase the precision of the estimates enough to offset the measurement error problem.

difference, can be detected at 90 % level; at 80% power we can detect a 7% difference in the prevalence rates.

In summary, the results from the power analysis show that the sample design is well powered to detected economically significant differences across beneficiaries and non-beneficiaries.

Figure 14 Standardized effect size (MDE) vs. Power under 3 survey waves for continuous outcomes

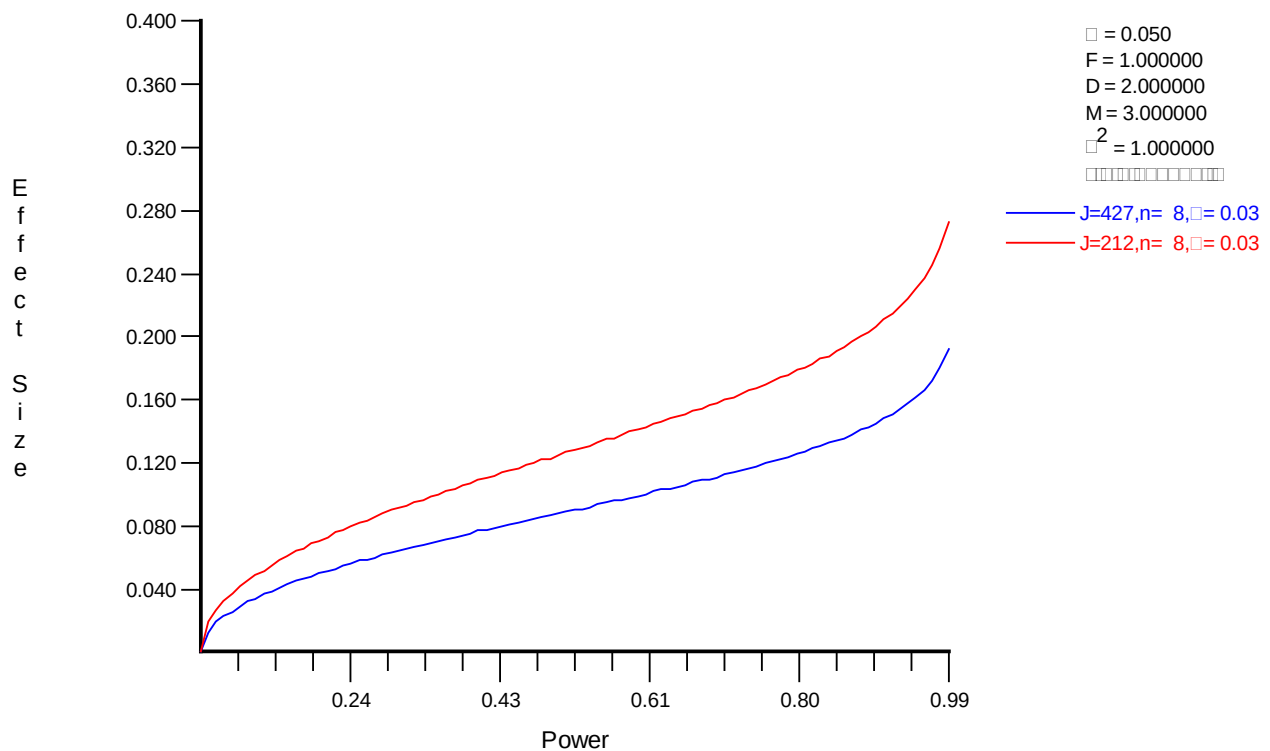


Figure 15 Standardized effect size (MDE) vs. Power at baseline for continuous outcomes

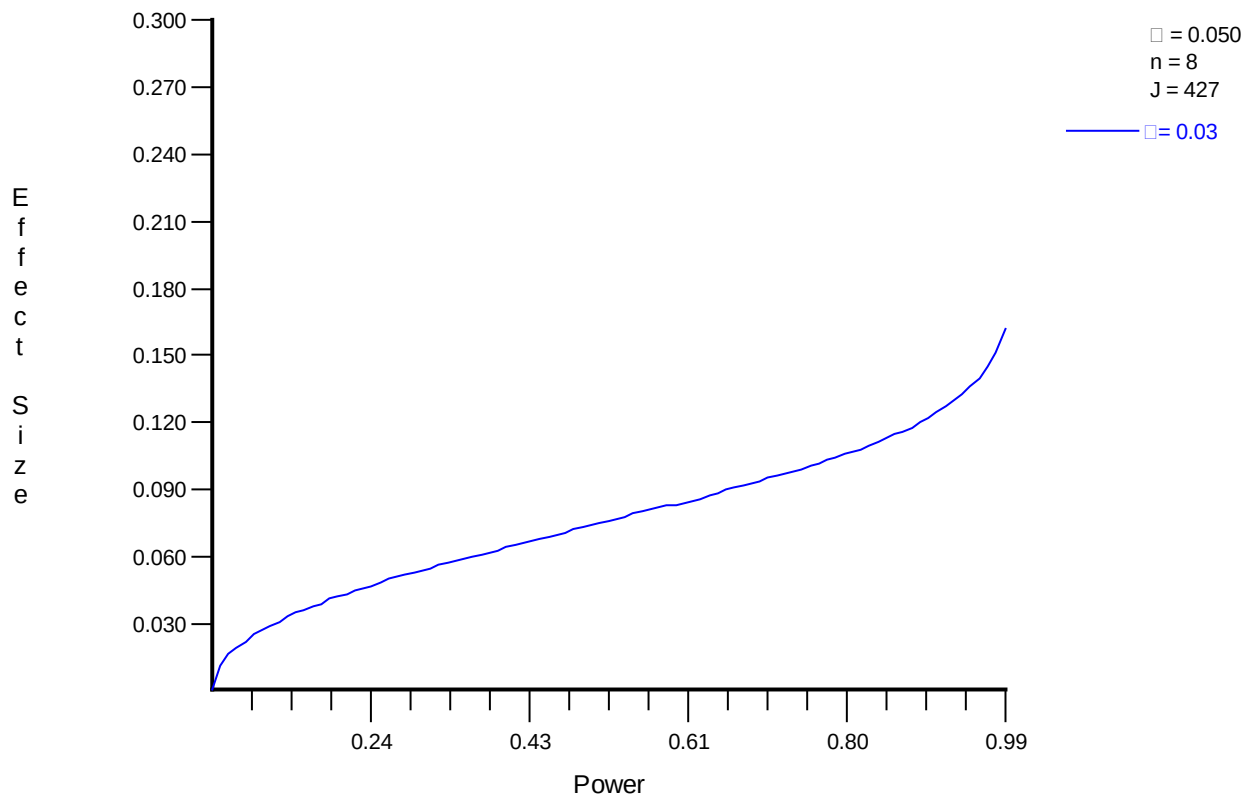


Figure 16 Power vs. Number of cluster vs. power for binary outcome

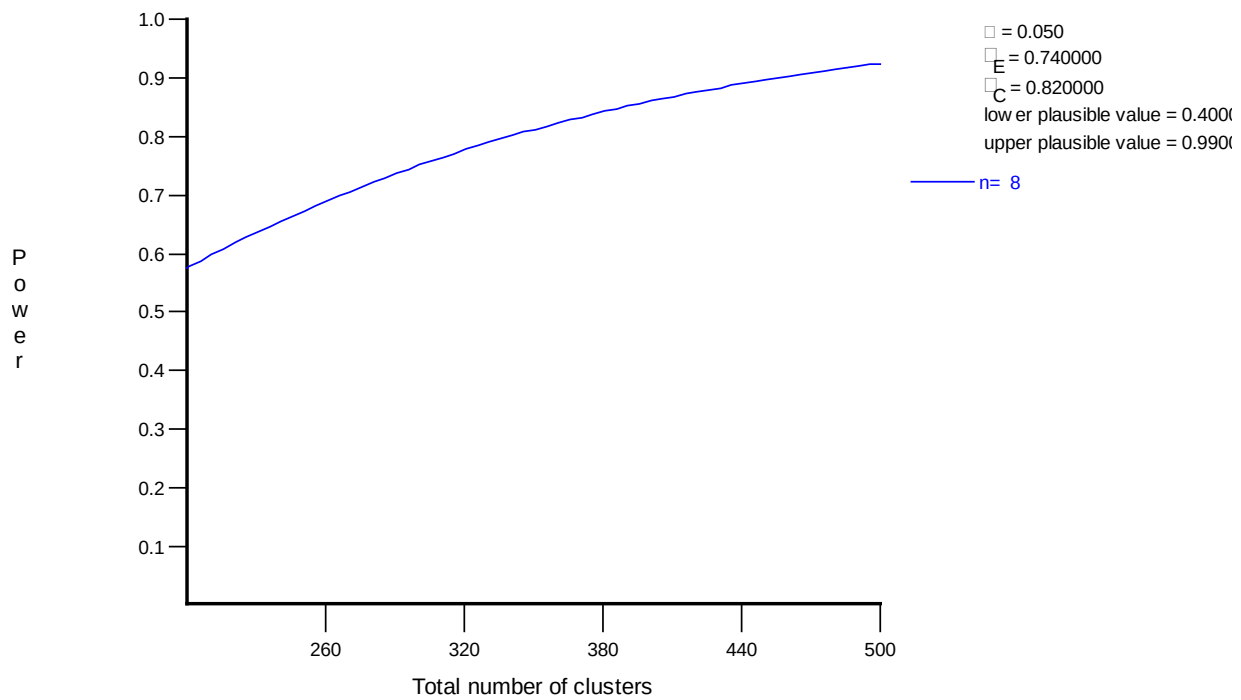
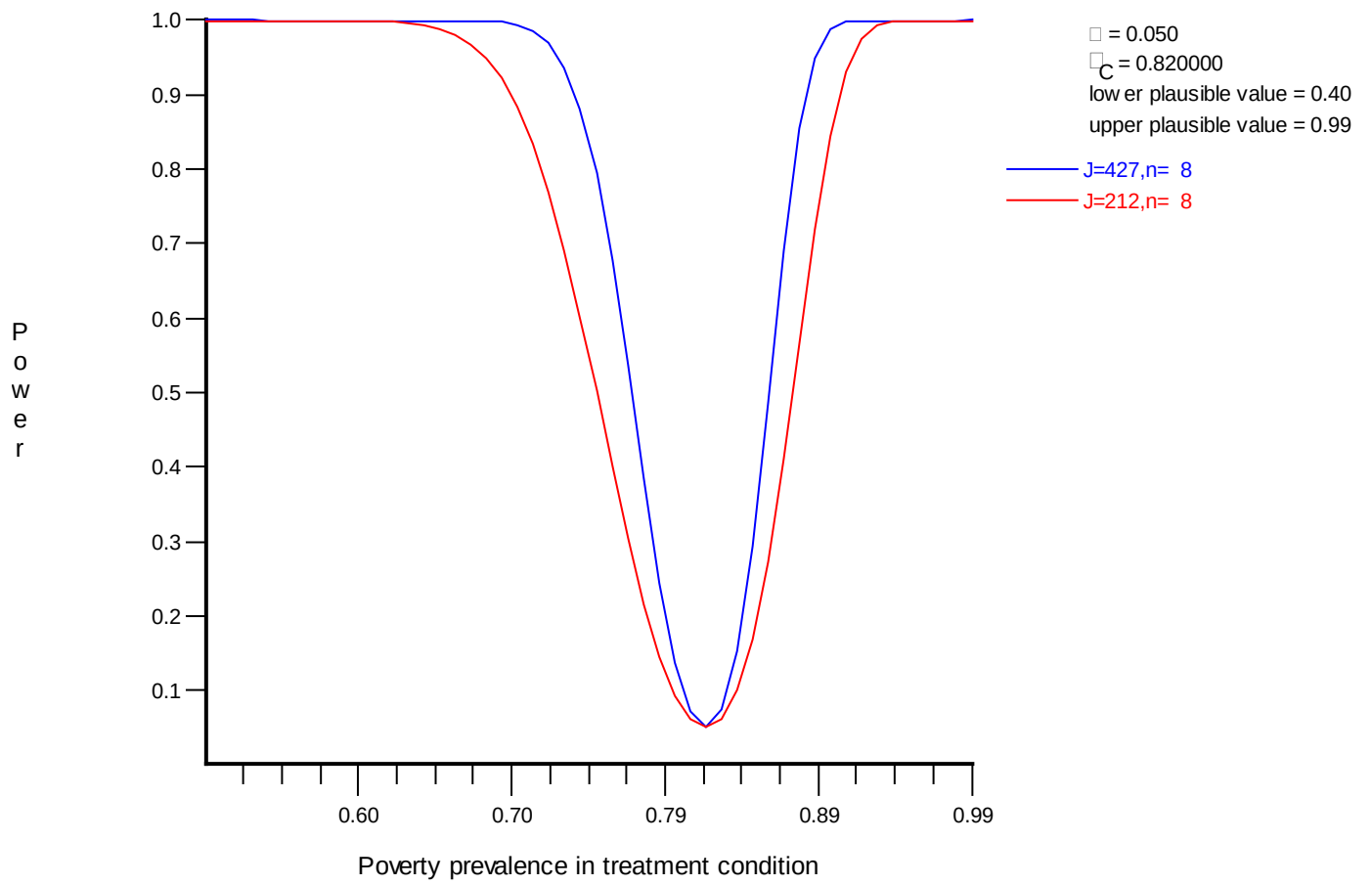


Figure 17 Power vs. Prevalence of poverty in treatment group, binary outcome



APPENDIX D - AVERAGE TREATMENT ON THE TREATED IMPACT ESTIMATES

THE SUB-SAMPLE OF HOUSEHOLDS REPORTING PARTICIPATION IN USAID-ACCESO

To evaluate USAID-ACCESO, we designed a survey that sampled 3,326 households in the activity's area of influence across six departments of Honduras: Copán, Intibucá, La Paz, Lempira, Ocotepeque, and Santa Bárbara. The baseline survey included 1,256 households in the treatment group, which were randomly selected from the roster of about 12,000 beneficiaries provided by the implementer, Fintrac, in April 2012. Our data showed that over fifty percent of households in the treatment group included in the survey do not self-report that they participate in the program in either the baseline or the two follow-up rounds of the survey. In the endline impact evaluation report the treatment group was comprised by all the households that were interviewed from the beneficiary list and other households that reported participation in USAID-ACCESO. In this discussion, we estimate the impact across a set of variables to for the group households that report participating and construct alternative propensity score weights using the same pool of comparison households and excluding the treatment households that do not report participating. The idea is to compare and contrast the effects among this specific group; in a sense providing the average treatment effect on the treated (ATT) to compare to the intention to treat (ITT) effects in the impact evaluation report.

Before discussing the results of this estimation, we note two points. First, the intention to treat effects in the impact evaluation report are the impact of the program. They include the effect of retaining households that were offered the program at some point and did not continue. This is one of the outcomes in any program offering a service. Second, the effects that can be detected among this selected group would need to be larger to be detected with the reduced sample. If the previous results are being diluted because of misclassified households, then the effects among this subsample should be larger.

The main conclusion we derive from this additional analysis is that the impact estimates among the households that report participating in the program are similar to those estimated with the complete treatment group. Below we present some estimates of the effects for a selection of indicators.

Expenditure and Poverty. We find no significant effects in the expenditure measure. The size of the estimated effects in the ATT sample suggest a difference of 2 percent in expenditure between the treatment and comparison group in the years after the baseline, but this is not significantly different

from zero. The effects on poverty mirror the results on poverty. We find a non-significant increase in extreme poverty, with the ATT groups being 4.7 percent more likely to be classified as extremely poor; in comparison, the ITT estimate was 3.3 percent.

Child nutrition and anthropometrics. The ATT estimates on feeding practices are similar to the ITT presented in the report. The impact estimates on breastfeeding practices are significant. Children living in treatment households that report participating in the program are less likely to be exclusively breastfed for the under 6 month old sample. The estimates suggest that children age 6 to 23 months in treatment households are more likely to be breastfed in the first follow-up survey, and no significant difference in the endline survey (2015). The estimates that use data on the weight and height of children are similar to the ITT estimates. The main difference we find is on the weight for age measure. We find that children in treatment households have higher weight for age z-scores in the first follow-up. This difference is not reflected in the wasting, underweight or stunting prevalence. The results on the probability of having anemia show no significant differences between the treatment and the comparison group.

Women dietary diversity and anthropometrics. The impact estimates for the number of food groups consumed are significant, as were the ITT estimates. Women in treatment households consume 0.2 more food groups (from a 3.21 baseline and 9 food group score), in comparison, the ITT estimate was 0.17 food groups more. The underweight and anemia status indicators are not different between women that live in treatment household and those that do not.

Agriculture. The results for the agriculture variables are consistent with the estimates in the report. The main difference we find is that this group of farmers in the treatment group have has lower yields in beans production. While the ITT estimates were different from zero only for the midline impact estimate, the ATT estimates are larger and suggest that the group of farmers that report participating has a lower land productivity in beans production. We find no systematic differences between treatment and comparison group in the production or sales of vegetables, fruits or tubers. We find that among this subsample, treatment farmers are less likely to produce fruits and vegetables. As noted in the report, this is likely due to farmers in the treatment farmer being more likely to have fruits and vegetables, and the proportion decreasing with respect to the comparison group.

Labor and Income. We estimate the impact for different types of income (wage, independent business) and by different types of occupation (merchants, agriculture, etc.). We find no significant effects, and

the estimates are of similar sign and size as those in the main report. We further estimate the effects on the probability of participating in waged labor, agriculture, etc. We find that this sub-group of the treatment households is more likely to participate in wage labor. These households are 18 percentage points more likely to participate in wage labor than a similar comparison group. The ITT effects were lower, with treatment households being 12 percentage points more likely to participate in wage labor. The impact estimates on the number of days worked per year show no significant differences between the groups.

Coffee Rust. Given the effects of coffee rust in Central America, we introduced specific questions on coffee management and prevention practices. As in the main report, we found some evidence that farmers in the treatment have more knowledge of preventive practices and had less area affected by the rust. We also find that treatment farmers are 18 percentage points more likely to have received assistance.

Table 72 ATT Impact on total expenditure

	(1)	(2)	(3)	(4)
Expenditure PC per day in 2005PPP to 2010 Prices				
Treatment x 2013- Midline	0.0089 [0.10]	0.013 [0.10]		
Treatment x 2015- Endline	-0.075 [0.096]	-0.07 [0.099]		
Treatment x Post- 2013 & 2015			-0.034 [0.086]	-0.03 [0.088]
Mean of Comp. at Baseline	1.97	1.98	1.97	1.98
SD of Comp. at Baseline	1.27	1.28	1.27	1.28
Observations	6531	6531	6531	6531
Expenditure PC per day in 2005 PPP USD				
Treatment x 2013- Midline	0.0079 [0.093]	0.012 [0.093]		
Treatment x 2015- Endline	-0.067 [0.086]	-0.063 [0.089]		
Treatment x Post- 2013 & 2015			-0.03 [0.077]	-0.026 [0.078]
Mean of Comp. at Baseline	1.77	1.77	1.77	1.77
SD of Comp. at Baseline	1.14	1.15	1.14	1.15
Observations	6531	6531	6531	6531
Log-Expenditure PC in 2005 PPP 2010 USD				
Treatment x 2013- Midline	0.011 [0.045]	0.014 [0.045]		
Treatment x 2015- Endline	0.012 [0.044]	0.013 [0.045]		
Treatment x Post- 2013 & 2015			0.011 [0.038]	0.014 [0.039]
Mean of Comp. at Baseline	0.51	0.51	0.51	0.51
SD of Comp. at Baseline	0.59	0.59	0.59	0.59
Observations	6531	6531	6531	6531
Log-Expenditure PC in 2005 PPP USD				
Treatment x 2013- Midline	0.011 [0.045]	0.014 [0.045]		
Treatment x 2015- Endline	0.012 [0.044]	0.013 [0.045]		
Treatment x Post- 2013 & 2015			0.011 [0.038]	0.014 [0.039]
Mean of Comp. at Baseline	0.4	0.4	0.4	0.4
SD of Comp. at Baseline	0.59	0.59	0.59	0.59
Number of Clusters	276	276	276	276
Number of Households	2404	2404	2404	2404
Observations	6531	6531	6531	6531

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Log variables represent percentage changes. Corresponds to table 40 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 73 ATT Impact on Poverty

	(1)	(2)	(3)	(4)
Poverty <= 1.25 2005 PPP USD PC per Day				
Treatment x 2013- Midline	0.002 [0.048]	0.0035 [0.047]		
Treatment x 2015- Endline	-0.013 [0.039]	-0.011 [0.038]		
Treatment x Post- 2013 & 2015			-0.0056 [0.039]	-0.0039 [0.038]
Mean of Comp. at Baseline	0.38	0.38	0.38	0.38
SD of Comp. at Baseline	0.49	0.49	0.49	0.49
Observations	6531	6531	6531	6531
Extreme Poverty - <= 2.84 2005 PPP 2010 USD				
Treatment x 2013- Midline	0.035 [0.033]	0.035 [0.034]		
Treatment x 2015- Endline	0.058 [0.042]	0.058 [0.043]		
Treatment x Post- 2013 & 2015			0.046 [0.034]	0.047 [0.035]
Mean of Comp. at Baseline	0.83	0.83	0.83	0.83
SD of Comp. at Baseline	0.38	0.38	0.38	0.38
Observations	6531	6531	6531	6531
Local Poverty <= 3.79 2005 PPP 2010 USD				
Treatment x 2013- Midline	0.013 [0.025]	0.013 [0.026]		
Treatment x 2015- Endline	0.045 [0.031]	0.048 [0.033]		
Treatment x Post- 2013 & 2015			0.029 [0.025]	0.031 [0.026]
Mean of Comp. at Baseline	0.92	0.92	0.92	0.92
SD of Comp. at Baseline	0.27	0.27	0.27	0.27
Observations	6531	6531	6531	6531
Relative Poverty - (2.84,3.79] 2005 PPP 2010 USD				
Treatment x 2013- Midline	-0.022 [0.030]	-0.021 [0.030]		
Treatment x 2015- Endline	-0.013 [0.026]	-0.011 [0.026]		
Treatment x Post- 2013 & 2015			-0.018 [0.024]	-0.016 [0.024]
Mean of Comp. at Baseline	0.096	0.095	0.096	0.095
SD of Comp. at Baseline	0.3	0.29	0.3	0.29
Number of Clusters	276	276	276	276
Number of Households	2404	2404	2404	2404
Observations	6531	6531	6531	6531

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 41 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 74 ATT Impact on Breastfeeding Practices

	(1)	(2)	(3)	(4)
Indicator for exclusive breastfeeding (< 6 months)				
Treatment x 2013- Midline	-0.41 [0.11]***	-0.42 [0.13]***		
Treatment x 2015- Endline	0.27 [0.35]	0.28 [0.34]		
Treatment x Post- 2013 & 2015			-0.41 [0.11]***	-0.42 [0.13]***
Mean of Comp. at Baseline	0.8	0.78	0.8	0.78
SD of Comp. at Baseline	0.41	0.42	0.41	0.42
Observations	296	296	296	296
Indicator for current breastfeeding (6-23 months)				
Treatment x 2013- Midline	0.23 [0.11]**	0.22 [0.10]**		
Treatment x 2015- Endline	-0.012 [0.17]	-0.013 [0.17]		
Treatment x Post- 2013 & 2015			0.16 [0.11]	0.15 [0.11]
Mean of Comp. at Baseline	0.57	0.56	0.57	0.56
SD of Comp. at Baseline	0.5	0.5	0.5	0.5
Number of Clusters	214	214	214	214
Number of Households	776	776	776	776
Observations	1068	1068	1068	1068

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 43 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 75 ATT Impact on anthropometric z-scores

	(1)	(2)	(3)	(4)
Length/height-for-age Z-score				
Treatment x 2013- Midline	-0.022 [0.19]	-0.013 [0.19]		
Treatment x 2015- Endline	0.25 [0.22]	0.28 [0.22]		
Treatment x Post- 2013 & 2015			0.099 [0.18]	0.11 [0.18]
Mean of Comp. at Baseline	-1.62	-1.64	-1.62	-1.64
SD of Comp. at Baseline	1.64	1.66	1.64	1.66
Observations	2675	2675	2675	2675
Weight-for-age Z-score				
Treatment x 2013- Midline	0.35 [0.12]***	0.35 [0.13]***		
Treatment x 2015- Endline	0.28 [0.24]	0.3 [0.24]		
Treatment x Post- 2013 & 2015			0.32 [0.14]**	0.33 [0.14]**
Mean of Comp. at Baseline	-0.74	-0.75	-0.74	-0.75
SD of Comp. at Baseline	1.38	1.39	1.38	1.39
Observations	2627	2627	2627	2627
Weight-for-length/height Z-score				
Treatment x 2013- Midline	0.33 [0.21]	0.3 [0.22]		
Treatment x 2015- Endline	0.23 [0.25]	0.26 [0.25]		
Treatment x Post- 2013 & 2015			0.26 [0.20]	0.27 [0.20]
Mean of Comp. at Baseline	0.15	0.14	0.15	0.14
SD of Comp. at Baseline	1.11	1.1	1.11	1.1
Observations	2580	2580	2580	2580
BMI-for-age Z-score				
Treatment x 2013- Midline	0.14 [0.24]	0.12 [0.24]		
Treatment x 2015- Endline	0.14 [0.28]	0.16 [0.28]		
Treatment x Post- 2013 & 2015			0.11 [0.22]	0.12 [0.23]
Mean of Comp. at Baseline	0.36	0.35	0.36	0.35
SD of Comp. at Baseline	1.21	1.21	1.21	1.21
Number of Clusters	240	240	240	240
Number of Households	1194	1194	1194	1194
Observations	2594	2594	2594	2594

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 46 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 76 ATT Impact on children underweight, stunting, and wasting

	(1)	(2)	(3)	(4)
Probability of child being underweight				
Treatment x 2013- Midline	-0.029 [0.041]	-0.026 [0.042]		
Treatment x 2015- Endline	-0.019 [0.056]	-0.033 [0.059]		
Treatment x Post- 2013 & 2015			-0.022 [0.039]	-0.029 [0.041]
Mean of Comp. at Baseline	0.14	0.14	0.14	0.14
SD of Comp. at Baseline	0.35	0.35	0.35	0.35
Observations	2594	2594	2594	2594
Probability of child being stunted				
Treatment x 2013- Midline	-0.06 [0.087]	-0.063 [0.087]		
Treatment x 2015- Endline	-0.071 [0.079]	-0.07 [0.080]		
Treatment x Post- 2013 & 2015			-0.065 [0.079]	-0.066 [0.079]
Mean of Comp. at Baseline	0.38	0.39	0.38	0.39
SD of Comp. at Baseline	0.49	0.49	0.49	0.49
Observations	2675	2675	2675	2675
Probability of child being wasted				
Treatment x 2013- Midline	-0.023 [0.041]	-0.016 [0.042]		
Treatment x 2015- Endline	-0.028 [0.029]	-0.027 [0.031]		
Treatment x Post- 2013 & 2015			-0.023 [0.029]	-0.02 [0.030]
Mean of Comp. at Baseline	0.037	0.04	0.037	0.04
SD of Comp. at Baseline	0.19	0.2	0.19	0.2
Number of Clusters	240	240	240	240
Number of Households	1206	1206	1206	1206
Observations	2674	2674	2674	2674

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 47 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 77 ATT Impact on Child anemia

	(1)	(2)	(3)	(4)
Hemoglobin level G/DL				
Treatment x 2013- Midline	-0.2 [0.19]	-0.2 [0.20]		
Treatment x 2015- Endline	-0.049 [0.23]	-0.019 [0.23]		
Treatment x Post- 2013 & 2015			-0.16 [0.20]	-0.13 [0.20]
Mean of Comp. at Baseline	11.6	11.6	11.6	11.6
SD of Comp. at Baseline	1.18	1.18	1.18	1.18
Observations	2201	2201	2201	2201
Indicator for child anemia (6-59 months)				
Treatment x 2013- Midline	0.1 [0.085]	0.096 [0.088]		
Treatment x 2015- Endline	0.12 [0.087]	0.12 [0.086]		
Treatment x Post- 2013 & 2015			0.11 [0.082]	0.1 [0.081]
Mean of Comp. at Baseline	0.24	0.24	0.24	0.24
SD of Comp. at Baseline	0.43	0.43	0.43	0.43
Observations	2187	2187	2187	2187
Child has edema				
Treatment x 2013- Midline	0.016 [0.023]	0.013 [0.024]		
Treatment x 2015- Endline	-0.0055 [0.032]	-0.0023 [0.032]		
Treatment x Post- 2013 & 2015			0.0063 [0.017]	0.0069 [0.017]
Mean of Comp. at Baseline	0.03	0.03	0.03	0.03
SD of Comp. at Baseline	0.17	0.17	0.17	0.17
Number of Clusters	242	242	242	242
Number of Households	1228	1228	1228	1228
Observations	2787	2787	2787	2787

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 48 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 78 ATT Impact on Women's dietary diversity

	(1)	(2)	(3)	(4)
Dietary Diversity (Mean x of food groups (9) consumed)				
Treatment x 2013- Midline	0.15 [0.11]	0.15 [0.11]		
Treatment x 2015- Endline	0.25 [0.15]*	0.24 [0.14]*		
Treatment x Post- 2013 & 2015			0.2 [0.098]**	0.2 [0.098]**
Mean of Comp. at Baseline	3.21	3.21	3.21	3.21
SD of Comp. at Baseline	1.06	1.05	1.06	1.05
Observations	5948	5948	5948	5948
Dietary Diversity (Mean x of food groups (10) consumed)				
Treatment x 2013- Midline	0.2 [0.14]	0.2 [0.14]		
Treatment x 2015- Endline	0.3 [0.17]*	0.3 [0.17]*		
Treatment x Post- 2013 & 2015			0.26 [0.11]**	0.25 [0.11]**
Mean of Comp. at Baseline	4.2	4.2	4.2	4.2
SD of Comp. at Baseline	1.31	1.3	1.31	1.3
Observations	5948	5948	5948	5948
Women Min. Dietary Diversity (10 Food)				
Treatment x 2013- Midline	0.059 [0.048]	0.062 [0.049]		
Treatment x 2015- Endline	0.059 [0.053]	0.063 [0.054]		
Treatment x Post- 2013 & 2015			0.058 [0.041]	0.062 [0.042]
Mean of Comp. at Baseline	0.38	0.38	0.38	0.38
SD of Comp. at Baseline	0.49	0.49	0.49	0.49
Number of Clusters	266	266	266	266
Number of Households	2041	2041	2041	2041
Observations	5948	5948	5948	5948

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 49 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 79 ATT Land productivity: Yield (Kg/Ha) for Corn

	(1)	(2)	(3)	(4)
Yield (Kg/Ha)				
Treatment x 2013- Midline	-394.7 [230.0]*	-423.6 [235.7]*		
Treatment x 2015- Endline	-345.4 [229.6]	-388.1 [255.5]		
Treatment x Post- 2013 & 2015			-377.1 [208.3]*	-412.8 [223.2]*
Mean of Comp. at Baseline	1603.1	1641.3	1603.1	1641.3
SD of Comp. at Baseline	2125.8	2291.8	2125.8	2291.8
Observations	2714	2714	2714	2714
Yield Kg/Ha Med. Impute				
Treatment x 2013- Midline	-189.2 [194.9]	-196.1 [196.4]		
Treatment x 2015- Endline	-107.9 [185.9]	-107.3 [187.2]		
Treatment x Post- 2013 & 2015			-156.1 [171.6]	-159.3 [173.1]
Mean of Comp. at Baseline	1456	1460.9	1456	1460.9
SD of Comp. at Baseline	1326.6	1337	1326.6	1337
Observations	2714	2714	2714	2714
Yield Kg/Ha trimmed				
Treatment x 2013- Midline	-185.8 [194.6]	-192.6 [195.9]		
Treatment x 2015- Endline	-103.4 [185.8]	-102.7 [187.0]		
Treatment x Post- 2013 & 2015			-151.8 [171.4]	-154.8 [172.8]
Mean of Comp. at Baseline	1457.3	1462.5	1457.3	1462.5
SD of Comp. at Baseline	1332.3	1343.9	1332.3	1343.9
Number of Clusters	245	245	245	245
Number of Households	1430	1430	1430	1430
Observations	2686	2686	2686	2686

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level (Corn) and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 31 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 80 ATT Ag. Sales revenue household aggregate impact with alternative prices for Corn

	(1)	(2)	(3)	(4)
Sales revenue with reported price - in 2005 PPP to 2010 USD				
Treatment x 2013- Midline	2.85 [110.5]	-2.65 [107.5]		
Treatment x 2015- Endline	60.4 [121.6]	53.6 [118.0]		
Treatment x Post- 2013 & 2015			32.8 [107.9]	26.9 [104.8]
Mean of Comp. at Baseline	355.4	352.7	355.4	352.7
SD of Comp. at Baseline	504.3	501.5	504.3	501.5
Observations	858	858	858	858
Sales revenue with median price - in 2005 PPP to 2010 USD				
Treatment x 2013- Midline	20.9 [112.4]	21.7 [108.8]		
Treatment x 2015- Endline	91.9 [187.2]	90 [179.8]		
Treatment x Post- 2013 & 2015			67 [124.6]	66.3 [120.9]
Mean of Comp. at Baseline	336.4	331.3	336.4	331.3
SD of Comp. at Baseline	459.5	458.5	459.5	458.5
Observations	944	944	944	944
Sales revenue with department median price - in 2005 PPP to 2010 USD				
Treatment x 2013- Midline	53.2 [98.2]	53.3 [95.1]		
Treatment x 2015- Endline	100.8 [163.9]	98 [157.3]		
Treatment x Post- 2013 & 2015			83.7 [106.2]	82.2 [103.1]
Mean of Comp. at Baseline	311.1	306.9	311.1	306.9
SD of Comp. at Baseline	417.8	417.2	417.8	417.2
Number of Clusters	138	138	138	138
Number of Households	418	418	418	418
Observations	944	944	944	944

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level (Corn) and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 29 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 81 ATT Land productivity: Yield (Kg/Ha) for Beans

	(1)	(2)	(3)	(4)
Yield (Kg/Ha)				
Treatment=1 x 2013- Midline	-637.6 [332.2]*	-627.9 [315.2]**		
Treatment=1 x 2015- Endline	-175.7 [188.0]	-182.3 [187.9]		
Treatment=1 x Post- 2013 & 2015			-438 [216.5]**	-435.4 [212.8]**
Mean of Comp. at Baseline	722.2	723.5	722.2	723.5
SD of Comp. at Baseline	2042.5	1988	2042.5	1988
Observations	1218	1218	1218	1218
Yield Kg/Ha Med. Impute				
Treatment=1 x 2013- Midline	-610.4 [241.5]**	-613.7 [243.0]**		
Treatment=1 x 2015- Endline	-415.4 [155.9]***	-412.1 [157.4]***		
Treatment=1 x Post- 2013 & 2015			-520.9 [190.8]***	-519.5 [192.8]***
Mean of Comp. at Baseline	607.5	613.7	607.5	613.7
SD of Comp. at Baseline	591	593.3	591	593.3
Observations	1218	1218	1218	1218
Yield Kg/Ha trimmed				
Treatment=1 x 2013- Midline	-587.1 [236.1]**	-593.1 [237.6]**		
Treatment=1 x 2015- Endline	-410.6 [154.1]***	-409.4 [155.5]***		
Treatment=1 x Post- 2013 & 2015			-510.3 [185.7]***	-511 [187.5]***
Mean of Comp. at Baseline	608.5	614.8	608.5	614.8
SD of Comp. at Baseline	594.5	596.6	594.5	596.6
Number of Clusters	197	197	197	197
Number of Households	826	826	826	826
Observations	1186	1186	1186	1186

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level (Beans) and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 31 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 82 ATT Ag. Sales revenue household aggregate impact with alternative prices for beans

	(1)	(2)	(3)	(4)
Sales revenue with reported price - in 2005 PPP to 2010 USD				
Treatment x 2013- Midline	-103.1 [107.8]	-109.9 [108.3]		
Treatment x 2015- Endline	10.1 [155.7]	25.8 [158.2]		
Treatment x Post- 2013 & 2015			-14.3 [97.9]	-10 [96.1]
Mean of Comp. at Baseline	306.3	317	306.3	317
SD of Comp. at Baseline	371.9	376.7	371.9	376.7
Observations	410	410	410	410
Sales revenue with median price - in 2005 PPP to 2010 USD				
Treatment x 2013- Midline	-112.4 [116.5]	-121.8 [115.3]		
Treatment x 2015- Endline	-47.1 [137.8]	-42.3 [135.7]		
Treatment x Post- 2013 & 2015			-55.4 [99.8]	-57.7 [96.0]
Mean of Comp. at Baseline	267.1	276.1	267.1	276.1
SD of Comp. at Baseline	323.6	328.3	323.6	328.3
Observations	429	429	429	429
Sales revenue with department median price - in 2005 PPP to 2010 USD				
Treatment x 2013- Midline	-123.2 [106.4]	-133.9 [106.7]		
Treatment x 2015- Endline	-70.1 [128.1]	-67 [125.9]		
Treatment x Post- 2013 & 2015			-71.1 [89.2]	-75 [85.8]
Mean of Comp. at Baseline	277.8	285.4	277.8	285.4
SD of Comp. at Baseline	333.5	336.5	333.5	336.5
Number of Clusters	95	95	95	95
Number of Households	236	236	236	236
Observations	429	429	429	429

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level (Beans) and impact estimation strategy consists of fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 29 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 83 ATT Probability Household Grows Corn, Beans or Coffee

	(1)	(2)	(3)	(4)
Probability of growing corn				
Treatment x 2013- Midline	0.037 [0.040]	0.035 [0.040]		
Treatment x 2015- Endline	0.1 [0.052]**	0.1 [0.054]*		
Treatment x Post- 2013 & 2015			0.073 [0.040]*	0.072 [0.041]*
Mean of Comp. at Baseline	0.8	0.8	0.8	0.8
SD of Comp. at Baseline	0.4	0.4	0.4	0.4
Observations	4078	4078	4078	4078
Probability of growing beans				
Treatment x 2013- Midline	0.0054 [0.043]	0.0057 [0.042]		
Treatment x 2015- Endline	0.054 [0.049]	0.058 [0.048]		
Treatment x Post- 2013 & 2015			0.031 [0.038]	0.034 [0.037]
Mean of Comp. at Baseline	0.38	0.38	0.38	0.38
SD of Comp. at Baseline	0.49	0.48	0.49	0.48
Observations	4078	4078	4078	4078
Probability of growing coffee				
Treatment x 2013- Midline	-0.032 [0.027]	-0.032 [0.027]		
Treatment x 2015- Endline	-0.042 [0.028]	-0.038 [0.028]		
Treatment x Post- 2013 & 2015			-0.037 [0.024]	-0.035 [0.024]
Mean of Comp. at Baseline	0.49	0.49	0.49	0.49
SD of Comp. at Baseline	0.5	0.5	0.5	0.5
Number of Clusters	264	264	264	264
Number of Households	1800	1800	1800	1800
Observations	4078	4078	4078	4078

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 11 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 84 ATT Prob. Household Grows vegetables, fruits, or tubers

	(1)	(2)	(3)	(4)
Probability of growing vegetables				
Treatment x 2013- Midline	-0.075 [0.031]**	-0.076 [0.031]**		
Treatment x 2015- Endline	-0.043 [0.030]	-0.046 [0.030]		
Treatment x Post- 2013 & 2015			-0.058 [0.023]**	-0.06 [0.023]**
Mean of Comp. at Baseline	0.027	0.026	0.027	0.026
SD of Comp. at Baseline	0.16	0.16	0.16	0.16
Observations	4078	4078	4078	4078
Probability of growing fruits				
Treatment x 2013- Midline	-0.033 [0.024]	-0.032 [0.023]		
Treatment x 2015- Endline	-0.072 [0.033]**	-0.067 [0.032]**		
Treatment x Post- 2013 & 2015			-0.054 [0.025]**	-0.051 [0.024]**
Mean of Comp. at Baseline	0.021	0.02	0.021	0.02
SD of Comp. at Baseline	0.14	0.14	0.14	0.14
Observations	4078	4078	4078	4078
Probability of growing tubers				
Treatment x 2013- Midline	-0.0028 [0.019]	-0.0022 [0.020]		
Treatment x 2015- Endline	-0.01 [0.020]	-0.0092 [0.021]		
Treatment x Post- 2013 & 2015			-0.0065 [0.019]	-0.0059 [0.019]
Mean of Comp. at Baseline	0.067	0.065	0.067	0.065
SD of Comp. at Baseline	0.25	0.25	0.25	0.25
Number of Clusters	264	264	264	264
Number of Households	1800	1800	1800	1800
Observations	4078	4078	4078	4078

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 12 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 85 ATT Impact of Labor Income from three principal activities

	(1)	(2)	(3)	(4)
Annual Income from dependent activity - Total - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	11 [313.2]	47.4 [305.3]		
Treatment x 2015- Endline	150.4 [341.9]	174.5 [341.3]		
Treatment x Post- 2013 & 2015			76.8 [295.9]	106.7 [292.6]
Mean of Comp. at Baseline	2634.9	2678.4	2634.9	2678.4
SD of Comp. at Baseline	3385.4	3415.9	3385.4	3415.9
Observations	4711	4711	4711	4711
Annual Income from independent activity - Total - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-1148.8 [650.9]*	-1115.4 [650.5]*		
Treatment x 2015- Endline	-1085.3 [695.8]	-1069.5 [704.9]		
Treatment x Post- 2013 & 2015			-1105.9 [646.8]*	-1080.1 [650.1]*
Mean of Comp. at Baseline	3044.8	3146.9	3044.8	3146.9
SD of Comp. at Baseline	6507.5	6629.8	6507.5	6629.8
Observations	4298	4298	4298	4298
Annual Invest./Cost for independent activity - Total - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-63.3 [198.9]	-49 [199.0]		
Treatment x 2015- Endline	303.4 [219.5]	344.6 [223.5]		
Treatment x Post- 2013 & 2015			119.8 [192.3]	149.6 [193.9]
Mean of Comp. at Baseline	1281.3	1339.2	1281.3	1339.2
SD of Comp. at Baseline	3239.8	3340.1	3239.8	3340.1
Number of Clusters	250	250	250	250
Number of Households	1476	1476	1476	1476
Observations	4006	4006	4006	4006

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 38 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 86 ATT Impacts on Income by Source

	(1)	(2)	(3)	(4)
Annual income for -Wage-Labor - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	118.5 [306.4]	157.6 [301.5]		
Treatment x 2015- Endline	296 [275.6]	303.6 [276.6]		
Treatment x Post- 2013 & 2015			205.6 [228.6]	230.1 [229.8]
Mean of Comp. at Baseline	2146.7	2171.9	2146.7	2171.9
SD of Comp. at Baseline	2948.9	2950	2948.9	2950
Observations	4146	4146	4146	4146
Annual income for -Merchant - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	659.6 [616.2]	712 [617.6]		
Treatment x 2015- Endline	23.3 [731.0]	48.8 [724.1]		
Treatment x Post- 2013 & 2015			371.2 [587.8]	400.5 [581.0]
Mean of Comp. at Baseline	1681.7	1711.7	1681.7	1711.7
SD of Comp. at Baseline	4788.9	4846.1	4788.9	4846.1
Observations	1101	1101	1101	1101
Annual income for -Agriculture - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	-891.2 [523.9]*	-840.8 [524.9]		
Treatment x 2015- Endline	-612.5 [664.3]	-553.9 [668.9]		
Treatment x Post- 2013 & 2015			-748.4 [573.0]	-691.9 [575.7]
Mean of Comp. at Baseline	2675	2766.9	2675	2766.9
SD of Comp. at Baseline	5043.2	5187.6	5043.2	5187.6
Observations	4211	4211	4211	4211
Annual income for -Professional - in 2005 PPP 2010 USD				
Treatment x 2013- Midline	144.4 [701.2]	184.3 [685.1]		
Treatment x 2015- Endline	480.1 [893.1]	555.6 [872.2]		
Treatment x Post- 2013 & 2015			273.4 [705.3]	325.1 [687.5]
Mean of Comp. at Baseline	2591.8	2720.4	2591.8	2720.4
SD of Comp. at Baseline	4612.7	4617.3	4612.7	4617.3
Number of Clusters	135	135	135	135
Number of Households	488	488	488	488
Observations	1299	1299	1299	1299

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 36 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 87 ATT Labor Market Participation

	(1)	(2)	(3)	(4)
Participation 1 - Wage-Labor				
Treatment x 2013- Midline	0.093 [0.096]	0.093 [0.096]		
Treatment x 2015- Endline	0.27 [0.10]***	0.25 [0.097]***		
Treatment x Post- 2013 & 2015			0.18 [0.078]**	0.17 [0.077]**
Mean of Comp. at Baseline	0.84	0.84	0.84	0.84
SD of Comp. at Baseline	0.95	0.94	0.95	0.94
Observations	3785	3785	3785	3785
Participation 1 - Merchant				
Treatment x 2013- Midline	-0.15 [0.13]	-0.15 [0.13]		
Treatment x 2015- Endline	0.074 [0.10]	0.078 [0.10]		
Treatment x Post- 2013 & 2015			-0.034 [0.093]	-0.031 [0.095]
Mean of Comp. at Baseline	0.51	0.51	0.51	0.51
SD of Comp. at Baseline	0.57	0.57	0.57	0.57
Observations	1034	1034	1034	1034
Participation 1 - Agriculture				
Treatment x 2013- Midline	0.058 [0.061]	0.057 [0.060]		
Treatment x 2015- Endline	0.051 [0.072]	0.054 [0.070]		
Treatment x Post- 2013 & 2015			0.055 [0.055]	0.056 [0.054]
Mean of Comp. at Baseline	1.22	1.22	1.22	1.22
SD of Comp. at Baseline	0.99	0.99	0.99	0.99
Observations	4630	4630	4630	4630
Participation 1 - Professional				
Treatment x 2013- Midline	-0.14 [0.11]	-0.15 [0.11]		
Treatment x 2015- Endline	0.053 [0.18]	0.052 [0.18]		
Treatment x Post- 2013 & 2015			-0.051 [0.14]	-0.054 [0.13]
Mean of Comp. at Baseline	0.62	0.63	0.62	0.63
SD of Comp. at Baseline	0.67	0.66	0.67	0.66
Number of Clusters	129	129	129	129
Number of Households	457	457	457	457
Observations	1227	1227	1227	1227

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and all regressions include household fixed effects. Column (1) and (3) presents the PSM weighted impact estimates and columns (2) and (4) uses attrition adjusted weights. Corresponds to table 35 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 88 ATT Coffee Rust: Impact on Assistance and Practices

	(1)	(2)
Farmer believes coffee rust can be prevented		
Treatment	0.03 [0.038]	0.034 [0.039]
2015- Endline	0.19 [0.047]***	0.19 [0.047]***
Mean of Comp. at Baseline	0.57	0.57
SD of Comp. at Baseline	0.5	0.5
Observations	1118	1118
Number of Ways Farmer knows to prevent Coffee Rust		
Treatment	0.14 [0.12]	0.14 [0.13]
2015- Endline	0.23 [0.13]*	0.23 [0.14]*
Mean of Comp. at Baseline	1.44	1.44
SD of Comp. at Baseline	0.78	0.78
Observations	730	730
Farmer Affected by Coffee Rust in Season of Survey		
Treatment	-0.049 [0.048]	-0.051 [0.049]
2015- Endline	-0.18 [0.034]***	-0.18 [0.035]***
Mean of Comp. at Baseline	0.81	0.81
SD of Comp. at Baseline	0.4	0.4
Number of Clusters	178	178
Number of Households	795	795
Observations	1118	1118

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and impact estimation strategy consists of single difference controlling for survey year. Column (1) presents the PSM weighted impact estimates and columns (2) uses attrition adjusted weights. Corresponds to table 36 in main report. Corresponds to table 33 in main report.

* p<0.10, ** p<0.05, *** p<0.01

Table 89 Coffee Rust: Impact on assistance and area affected

	(1)	(2)
Farm received any assistance with coffee rust		
Treatment	0.18 [0.033]***	0.18 [0.033]***
2015- Endline	-0.012 [0.035]	-0.018 [0.036]
Mean of Comp. at Baseline	0.078	0.078
SD of Comp. at Baseline	0.27	0.27
Observations	754	754
Number of Ways farmer knows to control Coffee Rust		
Treatment	-0.13 [0.077]	-0.13 [0.081]
2015- Endline	0.28 [0.067]***	0.29 [0.070]***
Mean of Comp. at Baseline	0.84	0.83
SD of Comp. at Baseline	0.65	0.65
Observations	859	859
Area affected by Coffee Rust (Ha.)		
Treatment	-5.21 [5.32]	-4.91 [5.03]
2015- Endline	-4.02 [4.48]	-3.85 [4.29]
Mean of Comp. at Baseline	11.4	10.9
SD of Comp. at Baseline	58.9	57.6
Observations	744	744
% of coffee-growing land affected by coffee rust		
Treatment	-0.038 [0.033]	-0.039 [0.033]
2015- Endline	-0.13 [0.028]***	-0.13 [0.029]***
Mean of Comp. at Baseline	0.47	0.47
SD of Comp. at Baseline	0.38	0.38
Number of Clusters	176	176
Number of Households	748	748
Observations	1032	1032

Standard errors in brackets

Std. errors are clustered at the village level. Treatment status and year indicator interaction coefficients are the impact estimates using weighted propensity score matching on the common support region and in Hirano et. al (2003). Observations are at the household level and impact estimation strategy consists of single difference controlling for survey year. Column (1) presents the PSM weighted impact estimates and columns (2) uses attrition adjusted weights.

Corresponds to table 36 in main report. Corresponds to table 33 and 34 in main report.

* p<0.10, ** p<0.05, *** p<0.01

APPENDIX E – STATEMENT OF DIFFERENCES

According to USAID policy the implementing partner(s), funder(s), and evaluation team members are to be given the opportunity to write a statement of differences regarding any significant unresolved differences of opinion. Below are the statements of difference from USAID/Honduras based on the April 2016 version of the evaluation.

	Issue	USAID/Honduras response	Evaluation team response
1.	Selection of treatment and control groups	Issues with the methodology used to select the treatment group and the control group create uncertainty about the validity of the report’s findings. With regard to the treatment group, IFPRI reports that “51% of households in the beneficiary roster do not self-report that they participated in the program in either the baseline of the two follow-up rounds of the survey”²⁸. IFPRI posits a “plausible explanation” for this: “assistance might not be sufficiently salient and thus households may not effectively recall access to the program”²⁹. This explanation, however, is based on conjecture. The information presented does not make it possible to determine the true cause. An equally plausible explanation is that the wrong individual may have been interviewed. A question arises as the extent to which survey respondents in the sample of treatment households were correctly selected; that is, the extent to which they were the individuals in the treatment household with the most reliable information on the household’s participation in the ACCESO activity. Self-reporting on participation in ACCESO was obtained by asking respondents “Participa algún miembro	The Treatment Group The saliency of the participation in ACCESO is one of the outputs of the evaluation. If the implementer is visiting and providing technical assistance, training, etc. to one or multiple members of the households, they will know who is providing this assistance. In addition, the branding of the ACCESO should make this even more salient for beneficiary households; the trucks, uniforms, training sheets, etc., all bore the ACCESO logo. The explanation advanced is not equally plausible. The household interviewed in the treatment group came from the list provided by the implementers at baseline. Households were selected by name from that list and identified in the communities by name, even in cases when the list had beneficiaries located in the wrong communities, we followed then and visit them in the community where they lived. The unit of treatment is the household, not the individual that responds to the survey. In any survey there are key

²⁸ IFPRI End-line Report, (section 3, p.8)

²⁹ IBID, p. 8

	Issue	USAID/Honduras response	Evaluation team response
		del hogar en el programa de capacitación/extensión USAID-ACCESO”³⁰. The IFPRI report does indicate whether the person or persons interviewed were the individuals who themselves had received ACCESO training and extension services and it is noted that the survey questionnaire did not include instructions or a protocol at the beginning of the survey to ensure that the respondent or respondents included the individual(s) in the household who participated in training and extension support provided by ACCESO. A third plausible explanation is that there were was incorrect information with the beneficiary list that was provided to IFPRI by ACCESO. With regard to the control group, the probability of treatment methodology that IFPRI applied to estimate the likelihood of receiving the treatment, employed variables that were different from the actual criteria used by ACCESO to select clients. For example, according to IFPRI’s probability model, an additional productive member in the household increases by eight percentage points the probability that the household is treated³¹. However, the number of productive household members was not a factor used by ACCESO to select clients. The detailed characteristics of participating households were only revealed once the household had become a client and ACCESO applied a questionnaire to that household. Household income (below the poverty line) and willingness to participate was the primary determinant of participation in ACCESO. IFPRI did attempt to match control households based on their income level, but the poverty rate at baseline of the control group was at the village level, rather than the household level.	respondents and in this case those were defined as those that made decisions about agriculture and/or expenditures in the household. The questions to verify participation in the treatment refer to anyone in the household and then identify which person participated. That the respondent to the extension modules is not the person receiving extension by ACCESO does not explain that no one knows about the program in the household. The third explanation advanced of wrong information on the initial list of the implementer is not supported by the data. When we matched to the updated list of beneficiaries in table 2 of the report, most households were marked as active households and only 50% mentioned ACCESO in the survey. The only argument that the data support is that active households are not well aware of the benefits or activities of the program. This points to a lack of effect in the first link of the impact change. The outputs (trainings, stoves, floor treatments, etc.) of the program are not evident or salient to the beneficiaries. The Control Group: The probability model does not aim to replicate the implementer’s decision-making process. It just needs to be related to the variables that will enter the decision process. While the number of productive members in the household might not be an explicit variable in the implementer’s decision, it would clearly affect the participation in the program. For example, in deciding between two identical

³⁰ Section 7, Question 2 of IFPRI’s Encuesta de Seguimiento/Línea Base-Alimentar el Futuro 2015

³¹ IFPRI End-line Report, section 3.1 p.10

	Issue	USAID/Honduras response	Evaluation team response
		The selection of the control group is also affected by the issue of the correct selection of the treatment group. If the treatment sample was not correctly identified, then the matching control sample would be prone to selection error as it would not be representative of ACCESO's actual client universe.	households one that has no one of prime age and the other that has one person in prime age, the one with the prime age person would be more likely to be selected. The model is consistent with this. Note that proxies of income are included (poverty rates of the village, household composition, etc.) and variables that would affect the "willingness to participate" are included (market access, access and use of land, etc.). The probability model shows that it can explain the variation in treatment assignment and that it can balance the variables of interest. That is, it can create a group that is similar in means to the treatment group at baseline. We doubt that the interviewed household are not the ones that were in the list provided by the implementers. If that were the case, we would not be able to find the names of the beneficiaries from the implementer's list in the survey household list. We find them. If the list were wrong at baseline, or if they were populated by people that were not followed up after by the implementer, then both point to shortcomings of the implementation process which are accurately reflected in the results of the impact evaluation.
2.	Methodology of analysis	IFPRI's analysis is based on the methodology of "intention to treat" (ITT) that includes every household in the sample drawn from the beneficiary list, regardless of whether the beneficiary household applied ACCESO-promoted agricultural practices. While methodologically sound, this approach does not provide information on the effectiveness of the set of agricultural practices that ACCESO promoted, by the fact that it includes non-applying households in the analysis. Using a "Per Protocol"	As mentioned above, the first link in the impact chain are the outputs of the program. This is why the design is based on the "intention to treat". If the implementer provides a list of names which they say they will do X to, the first step in what does X do to the beneficiaries is to check that the implementer did X. Then we can do a "treatment on the treated" analysis that takes into account that not all intended beneficiaries end up participating. We presented such analysis

	Issue	USAID/Honduras response	Evaluation team response
		(PP) analysis in addition to the ITT method could have provided more actionable information for decision-making by USAID on the effectiveness of ACCESO-promoted practices in the case of households that correctly applied them.	(see appendix C) and the results among these are qualitatively the same. Using a per protocol analysis, i.e. describing what happen to the ones that participated, learned, and adopted a practice is not part of the impact evaluation design. We note that to determine the effectiveness of a practice, one would need a research design aimed at evaluating that practice. To provide more actionable information, USAID should consider a performance evaluation of aspects of the program promoted; being careful to note that the subsample of people that correctly applied it is a very selected sample of the targeted population.
3.	Treatment Effects	Although IFPRI accurately describes ACCESO's Theory of Change (TOC) related to producers of basic crops, the evaluation only analyzes a few elements of the ACCESO's intervention. IFPRI's analysis of treatment effects only considered a few elements of the ACCESO's treatment package, and analyzed those elements separately, in isolation, rather than part of an integrated package. For example, for each crop, respondents were asked how much seeds, chemical and non-chemical fertilizers, and herbicides were applied. But ACCESO's treatment package at the individual farmer level included 24 technologies and systems relating to overall farm and business operations and management ³² , including practices such as using less amounts of fertilizer and herbicide applied at the right time, planting densities, weed control, post-harvest handling and record-keeping. The report does not analyze these aspects, limiting itself to a comparison of whether the treatment group used	The impact evaluation evaluates the package that was provided by ACCESO. Therefore, the treatment variable is at the household level and all the households included in the implementer's beneficiary roster are included as treated households. It was discussed from the initial design stages that the impact evaluation was not going to be able to detect differences for individualize packages of practices, benefits, etc. To do this each package (say maternal nutrition, child nutrition, horticulture practices, coffee practices, etc.) would need a design. The analysis indicators across the different paths of the theory of change to identify if a mechanisms and impact path could be identified.

³² ACCESO Final Report, p. 9

	Issue	USAID/Honduras response	Evaluation team response
		more inputs than the control group. In order to analyze the effectiveness of ACCESO's approach, it would have been useful for IFPRI to develop a variable or index that measured the extent to which the whole package was correctly applied. Moreover, the IFPRI evaluation does not analyze the impact of the irrigation systems introduced by ACCESO. For example, the section dealing with productivity does not differentiate between productivity under irrigation and non-irrigated crops. Similarly, the support provided by ACCESO to help facilitate access to financing is absent from IFPRI's analysis. It is important to note ACCESO clients did not receive one standard intervention. The binary treatment model (participated/did not participate) used in IFPRI's analysis does not adequately reflect the mix of interventions provided by ACCESO. On a related note, IFPRI's analysis focused on aspects that were not part of ACCESO's theory of change. For example, IFPRI analyzed the probability of growing corn and beans³³, but ACCESO's intervention was not intended to increase the number of clients growing corn and beans; rather, it was geared toward changing the way they grew them; for them to increase density through a combined package of techniques that would allow them to grow these subsistence crops on a smaller area of land, thereby liberating some land for income-generating crops. IFPRI found that "...treatment households are nine percentage points more likely to have their own agriculture production activities³⁴". The evaluation report does not analyze the impact	The creation of the index would show if any of the practices were applied and could be useful general information. In a future impact evaluation, a specific index and set of practices should be included at the design stage to show any effect that might be detected in the index is the product of changes in practices among the treatment group and not the product of ex-post selection of variables. We reiterate that the research design that was agreed on at the design stage made clear that individual interventions were not going to be evaluated separately. The argument that the analysis includes aspects that were not included in the theory of change fails to link the substitution effects in the target population. Changes the crop composition, land allocation, etc. would be reflected in these outcome indicators. That the analysis shows that "...treatment households are nine percentage points more likely to have their own agriculture production activities" shows that the treatment group selected by the implementer is more likely to be engaged in agriculture. The we can follow and see what income effects we have and try to link these two facts; always using the treatment assignment at the household level. The last paragraph in the statement by USAID wants to estimate the impact on income among households that have their own agricultural production. The evaluation design is not powered or meant to do that analysis. The program is not likely to be changing the composition of households that have

³³ IFPRI End-line Report, section 4.1, part "Distribution of crops", p. 30 and Table 11 on p. 33

³⁴ IFPRI Endline Report, section 4.2, part "Income and Labor Supply", p.73

	Issue	USAID/Honduras response	Evaluation team response
		on household income/consumption of households with their own agricultural production activities.	their own agricultural production activities. Then estimating the impact among those households could be biased by the selection of households that are included in the estimation (e.g. the households that are in the treatment group and have their own land are likely different than the other households in the treatment group that have less or no land).
4.	Child health and nutrition	IFPRI states that “...even if treatment households were able to produce nutrient-dense foods, children did not benefit from this increased dietary diversity. Given the low levels of children consuming minimally acceptable diets at baseline, and the low prevalence of vitamin-rich food in children’s diets, these results suggest that USAID-ACCESO was not successful in encouraging households to adopt improved feeding practices for children³⁵.” USAID/Honduras believes that one of the weaknesses in the design of the impact evaluation was that the evaluation team assumed that every treatment household received both the agricultural and nutritional components of the intervention. However, because USAID/Honduras does not receive any nutrition funds, far fewer households received nutritional assistance as compared to agricultural assistance. To truly ascertain if the nutritional component of the intervention had an impact on the health and nutrition indicators, the evaluation team would have had to form a treatment group using households that received nutritional assistance and compared them to those that did not. This wasn’t done. Therefore, the impact evaluation isn’t able to accurately measure the impact of the nutrition component on the health and nutrition indicators.	The evaluation design was not predicated on the assumption that every household would receive the same intervention. It is explained in the design that the evaluation is done as a package and that that package would depended on the conditions and needs the implementers found in the households. To get at the paths through which the program might affect the beneficiaries, the design proposed a set of outcomes which would reflect impacts across different paths. As mentioned above, creating treatment and control groups within specific themes of the intervention falls outside of the research design originally agreed, and the strength of the evidence this would provide if done ex-post is questionable.

³⁵ IFPRI End-Line Report, section 4.3, part “Child Nutrition and Health”, p.87

	Issue	USAID/Honduras response	Evaluation team response
5.	Overall		In general, the differences stated by USAID in this document point to the lack of disaggregated analysis in the report. This type of analysis needs specific research designs aimed at answering the questions of interest. The analysis IFPRI presented follows the initial design agreed upon and the strength of the analysis rests on having presented how the analysis would be done before doing the analysis. Analysis of variables and groups that were not initially discussed would need to be presented as ex-post exploration of the data. These explorations, might be useful to the decision-making process for USAID and the implementers, but would be weak evidence of the effects of the program among the target population. The statement seems to miss some lessons of the evaluation. The data and results convincingly show that there was a lack of connection between the program beneficiaries that were presented at baseline and the beneficiaries that ended up receiving the program. This is a lesson for a future intervention to consider: namely, the differences in the time that a beneficiary starts participating. The other important lesson is that if USAID is interested interventions and their effects, they should be noted at the design stage and an evaluation design be done for those specific interventions. During the design stage IFPRI discussed that the evaluation design was going to be in a package and that thematical outcomes would be analyzed across different impact paths. That proposal is different to the analysis among sub-groups that are brought up in this statement of difference.