



Rwanda's Improved Services for Vulnerable Populations Project

Impact Evaluation

Summary of End Line Findings

September 2019



Rwanda's Improved Services for Vulnerable Populations Project

Impact Evaluation

Summary of End Line Findings

September 2019

Jessica Fehringer, PhD, MHS, MEASURE Evaluation

Peter Lance, PhD, MEASURE Evaluation

Kenn Ndirangu, Incisive Africa

Aimee Benson, MA, MEASURE Evaluation

Gustavo Angeles, PhD, MEASURE Evaluation

Lisa Parker, PhD, MEASURE Evaluation

Shaylen Foley, MPH, MEASURE Evaluation

Stacie Gobin, MPH, MEASURE Evaluation

Veronica Caridad Varela, MPH, MEASURE Evaluation

Lauren Morris, PhD, MEASURE Evaluation

Abby Cannon, MPH, MSW, MEASURE Evaluation

Brittany Iskarpatyoti, MPH, MEASURE Evaluation

MEASURE Evaluation

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

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. TR-19-363

ISBN: 978-1-64232-171-5



ACKNOWLEDGMENTS

We thank the United States Agency for International Development (USAID) and the United States President's Emergency Plan for AIDS Relief (PEPFAR) for their support of this work.

We are thankful to the many participants who gave us their time and energy, and to the local leaders and community staff who helped us in the field.

We thank the Government of Rwanda—particularly the Ministry of Local Government, Ministry of Gender, National Institute of Statistics Rwanda, and the Rwanda National Ethics Committee—for supporting this evaluation.

The National University of Rwanda, School of Public Health, conducted initial data collection; we thank Sabine Musange, Ina Kalisa, Albert Ndagijimana, Vedaste Ndahindwa, and the many enumerators and supervisors for their hard work. Incisive Africa conducted end line data collection and contributed to data interpretation and end line report development; Yvette Mugeni, Denyse Nyinawishuli, Gahunga Sudi, and the enumerators and supervisors deserve thanks for their dedication and the quality of their contributions.

We are grateful to the many stakeholder participants in the meeting to share and validate results in Kigali on August 7, 2019 for their help.

Daniel Handel of USAID was instrumental in developing the evaluation design and launching the evaluation. We thank Mary de Boer, Esron Niyonsaba, and Célestin Mutumayi, of USAID/Rwanda, and Lindsay Little (USAID) and Tona Isibo, Absolom Muramira, and other staff of the Improved Services for Vulnerable Populations Project for the effort they put into communication, coordination, strategizing, and information sharing with the evaluation team. We thank Ana Djapovic Scholl of the USAID/Office of HIV and AIDS and Christine Fu, formerly of USAID, for their management assistance.

A large, talented team at the USAID- and PEPFAR-funded MEASURE Evaluation project made this work possible: Martha Priedman-Skiles, who led initial design of the study; Jessica Fehringer, who led the evaluation team and study; Peter Lance, who led sample design and selection and end line impact analyses; Aimee Benson, who contributed significantly to survey design, led tablet programming and data cleaning, carried out data analysis, and contributed to data interpretation; Gustavo Angeles, who contributed to evaluation and survey design, guided consumption analyses, provided senior expertise on analyses in general, and contributed to data interpretation; Lisa Parker, who contributed to evaluation and survey design, and interpretation of findings; Lisa Marie Albert, who contributed to survey development, managed initial data collection survey development and carried out local staff training; Judy Tatwangire (consultant), who provided data collection quality control at both time points; Brittany Iskarpatyoti, who helped with writing; Abby Cannon, who helped manage the qualitative data analysis and with writing; Madison Alexander, who helped with editing; and Becky Wilkes, for her work in mapping study sites. Stacie Gobin completed the cost-effectiveness analysis (CEA) assessment and contributed to data analysis. Shaylen Foley managed remote data collection, data analysis, and report writing for the CEA. Lisa Parker, Lauren Morris, and Scott Moreland supported the CEA report writing and interpretation. Gabriela Escudero and Heather Davis assisted with logistics and management.

We thank MEASURE Evaluation's knowledge management team, based at the University of North Carolina at Chapel Hill, for editorial, design, and production services.

Suggested citation

Fehringer, J., Lance, P., Ndirangu, K., Benson, A., Angeles, G., Parker, L., . . . Iskarpatyoti, B. (2019). Rwanda's Improved Services for Vulnerable Populations Project: Impact Evaluation: Summary of End Line Findings. Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina.

CONTENTS

Abbreviations.....	7
Introduction.....	8
Research Questions and Methods.....	9
Analysis.....	10
Results.....	11
Limitations	15
Discussion.....	16
Recommendations	17
Conclusions.....	18
References	23

TABLES

Table 1. Summary of key findings.....	11
Table 2. Summary of impact results, per DID estimation	19
Table 3. Summary of key outcomes at initial data collection and end line.....	21

ABBREVIATIONS

CEA	cost-effectiveness analysis
DID	difference-in-differences
ECD	early childhood development
GC	Global Communities
HES	household economic strengthening
ISLG	integrated savings and lending group
ISVP	Improved Services for Vulnerable Populations
KII	key informant interviews
NGO	nongovernmental organization
OVC	orphans and vulnerable children
USAID	United States Agency for International Development
PEPFAR	United States President's Emergency Plan for AIDS Relief

INTRODUCTION

The Improved Services for Vulnerable Populations (ISVP) project, known locally as Twiyubake, aims to improve the health, nutrition, and well-being of the populations on which the project focuses—orphans and vulnerable children (OVC), people living with HIV/AIDS, and economically vulnerable families—by strengthening their capacity in those domains. The ISVP is led by Global Communities (GC), along with international nongovernmental organizations (NGOs) and local civil society organizations. GC and its partners work within the Rwandan Ministries of Health (MOH) and Gender and Family Promotion framework. ISVP is supported by the United States Agency for International Development (USAID) Rwanda Mission.

The USAID- and United States President’s Emergency Plan for AIDS Relief-funded MEASURE Evaluation project—with support from USAID/Rwanda and in collaboration with Incisive Africa and the National University of Rwanda, College of Medicine and Health Sciences, School of Public Health—conducted an impact evaluation of the ISVP project. The evaluation sought to measure the impact of the interventions on the health, education, and economic well-being of vulnerable children and their families. This summary report shares end line impact, trend, and cost-effectiveness results using data from the 2017 and 2018 surveys, and costing data collected from ISVP. A full-length end line report on the evaluation is available, as well as a report from the initial data collection.

RESEARCH QUESTIONS AND METHODS

The ISVP evaluation was a prospective, cluster-randomized, controlled trial design, using a difference-in-differences (DID) estimation strategy with fixed-effects modeling to evaluate the impact on economic, health, and education outcomes. Administrative sectors were randomly assigned to study groups receiving different intervention packages: (1) a household economic strengthening (HES)-only group, which provides a core platform of financial interventions; (2) a full ISVP group, which supplements the core platform of HES interventions with health, education, and skill-building services; and (3) a control group, which received no ISVP interventions. The final number of assigned sectors ranged from 22 to 24: (1) full ISVP program, 23 sectors; (2) HES-only activities, 22 sectors; and (3) control, 24 sectors.

The evaluation sought to answer three primary research questions:

1. Does the full package of Twiyubake services (i.e., full ISVP) strengthen household economic status and provide additional support to motivate economically strengthened families to realize health and education benefits?
2. Can HES-only activities provide the economic stability for households to access health and education services, and improve individual health and educational well-being?
3. Which approach is more cost-effective, full ISVP or HES-only?

Initial household survey data were collected from April 2017 to July 2017 and September 2017. At that time, the program had been operating between 12 and 18 months in program areas. End line survey and qualitative data were collected from November 2018 to January 2019. This end line timing was selected because the program began implementation in early to mid-2016, and the program estimated that a two-year exposure period was needed to see changes in longer-term outcomes.

Program group households consisted of program beneficiaries sampled from beneficiary lists. Control group households consisted of vulnerable households sampled from the Government of Rwanda's Most Vulnerable Children list. Within each selected household, we selected the following members for the study: (1) all eligible primary caregivers of children ages 0 to 17 years old, (2) the primary member of an integrated savings and lending group (ISLG) (in the program groups), and (3) one randomly selected 10- to 17-year-old. The final household sample sizes at end line were 1,374 for full ISVP, 1,270 for HES-only, and 1,169 for control.

We conducted key informant interviews (KIIs) with 36 community leaders and 8 program staff at end line.

ANALYSIS

We used quantitative methods to compare data on the key outcome indicators in the full ISVP project intervention areas to those in the HES-only intervention and control areas, using STATA 15.1. Analyses included basic descriptive frequencies and some statistical testing of mean differences at end line. To examine changes in means within groups over time, we used cluster bootstrapping. We tested the difference in means between treatment and control groups using a regression model for selected indicators collected only at end line. We report indicators as either percentages or means, and weighted them using the sampling weights.

We determined the impact of the ISVP interventions on selected health, education, and economic outcomes using the DID model. This model identifies the impact of a program as the difference between a sample of participants and a control of nonparticipants regarding the trends each experienced in an outcome from baseline (before program implementation) to end line (after program implementation). In order to identify potential gender-related patterns in outcomes, DID estimation included subgroup analyses by sex (sex of household head, caregiver, and/or child, depending on outcome); only significant subgroup results are reported. Qualitative data provided context to the quantitative data so we could better understand what other programs were under way in the study areas and what major changes the program had made over time. Researchers coded the transcripts of the KIIs using deductive codes stemming from the interview guide topics. They used broad code categories, rather than detailed, line-by-line coding. We then reviewed content by code to assess emerging themes. For the community KIIs, we also compared codes across the three study groups to determine whether any differences existed among them. Last, researchers synthesized the themes into an initial draft report of qualitative findings. We then reread coded content and selected transcripts to confirm draft findings, and made adjustments in the text, as needed, to develop final content.

RESULTS

Key findings are summarized in Table 1 below, according to evaluation question. Table 2 also presents the detailed impact results, and Table 3 shows details on outcomes at initial data collection and end line.

Table 1. Summary of key findings

Primary Evaluation Questions	Key Findings
<p>1. Does the full package of Twiyubake services strengthen household economic status and provide the additional support to motivate economically strengthened families to realize health and education benefits?</p>	<p>Economic strengthening</p> <ul style="list-style-type: none"> • The percentage of households with moderate or severe household hunger decreased from 79.0 to 67.2 percent and was statistically significant for full ISVP ($p=0.000$). However, the DID model did not find that these changes were significant compared to control or HES-only. Households with male heads benefited more from the program (5 percentage points, $p=0.084$). • The percentage of households able to meet basic needs increased significantly, from 54.7 to 62.3 percent for full ISVP ($p=0.02$). However, the DID model did not find these changes to be significant. • The percentage of households owning livestock increased significantly from 58.0 to 66.9 percent ($p=0.000$). The DID model found that the change for full ISVP households was 3.7 percentage points higher than for control households ($p=0.038$). • Full ISVP households saw decreases in the percentage of household consumption for food (49.4% to 46.5%, $p=p=0.01$), health (6.3% to 4.6%, $p=0.001$), and education (4.2% to 2.3%, $p=0.001$). The DID model found that the full ISVP program offered a protective effect for food shares of about 3 percent, whereas the percentage shares going to food did not decrease as much as for the control group ($p=0.063$). DID results also showed the change for full ISVP health shares was 1 percentage point smaller compared to control and HES-only ($p=0.068$ and $p=0.041$, respectively). DID results showed the change for full ISVP education shares was 1.7 percentage points smaller compared to control and 1.5 percentage points smaller than HES-only ($p=0.000$ for both). • The percentage of households with savings accounts at end line was highest for full ISVP (75 percent), followed by HES-only (64 percent), and control (41 percent); the difference between full ISVP and both groups was statistically significant ($p=0.000$ for both). • The percentage of youth reporting new employment increased from 0.3 to 2.1 percent ($p=0.003$). The sample sizes were too low to include in DID analyses as originally planned. <p>Household decision making and gender-related attitudes and behaviors</p> <ul style="list-style-type: none"> • Support for harsh child punishment decreased significantly for full ISVP, from 46.3 percent to 32.4 percent ($p=0.000$). The DID model showed that the change over time was 16 points greater for full ISVP versus control ($p=0.01$) and 15 points greater compared to HES-only ($p=0.000$).

Primary Evaluation Questions	Key Findings
	<ul style="list-style-type: none"> • There were no other significant findings in this category. <p>Child health</p> <ul style="list-style-type: none"> • Birth registration increased significantly over time, from 85.2 to 90.8 percent (p=0.000). The DID model showed that the full ISVP change was 4.4 percentage points smaller than that for HES-only (p=0.025). • Caregiver's knowledge of child HIV status decreased from 30.0 percent to 26.8 percent (p=0.057); however because the control values decreased more, the DID model showed a protective effect for full ISVP, with the full ISVP change being 6.4 percentage points less than for control (p=0.000). <p>Youth health</p> <ul style="list-style-type: none"> • Youth (ages 10–17) reporting ever being testing for HIV increased from 26.7 percent to 33.4 percent; results were significant (p=0.029). The DID model showed that the change for full ISVP was 4.8 percentage points greater than control (p=0.042). • The percentage of youth knowing about HIV prevention strategies slightly decreased from 40.9 to 35.5 percent, but this change was not statistically significant. Because HES-only values increased, however, the DID model showed that full ISVP youth had a 10 percentage point decrease relative to HES-only (p=0.009). <p>Caregiver health</p> <ul style="list-style-type: none"> • Caregiver reports of having ever been tested for HIV increased from 63.1 percent to 68.1 percent, but this change was not statistically significant. <p>Education</p> <ul style="list-style-type: none"> • Early childhood development (ECD) attendance decreased for full ISVP, from 38.1 percent to 33 percent, but the change was not statistically significant. This change likely was because the end line survey did not capture in-home ECD, as the program had changed its ECD strategy before end line. At initial data collection, ECD attendance for full ISVP was much higher (likely because of program exposure before initial data collection) compared to both other groups and, over time, control ECD attendance increased slightly; because of these dynamics, the DID model showed a negative program impact for full ISVP compared to control (13.8 percentage points, p=0.009) and HES-only (9.4 percentage points, p=0.085). • Regular secondary school attendance among 13- to 17-year-olds stayed relatively stable, going from 17.7 percent to 18.9 percent. DID estimation showed no program impact on this outcome. • Full ISVP progression in school from the previous year decreased significantly, from 90.7 percent to 79.1 percent (p<0.001). The DID estimation showed no program impact on this outcome. • Youth graduating from primary and returning to secondary school decreased slightly, from 13.0 to 10.0

Primary Evaluation Questions	Key Findings
	<p>percent, and the change was statistically significant ($p < 0.05$). The DID estimation showed no program impact on this outcome; however, the change for females was 14 percentage points greater than for males ($p = 0.088$).</p>
<p>2. Can HES-only activities provide the economic stability for households to access health and education services, and improve individual health and educational well-being?</p>	<p>Economic strengthening</p> <ul style="list-style-type: none"> • Moderate or severe household hunger stayed roughly the same; the DID model, however, showed that HES-only households reported a 5.3 percentage point decrease in moderate or severe hunger, compared to control ($p = 0.043$). • The percentage of household consumption for food, health, and education decreased but the result was not significant for HES-only households. The DID model found that the HES-only program offered a protective effect for food shares of 3.1 percent, because the percentage of shares going to food did not decrease as much as for the control group ($p = 0.023$). • The percentage of households with savings accounts at end line was higher for HES-only (64 percent) than control households (41 percent); this difference was statistically significant ($p = 0.000$). • Ownership of livestock remained stable; the DID model did not find any significant differences between HES-only and control. • The percentage of youth reporting new employment went from 0.8 to 2.4 percent ($p = 0.078$); the sample size was too small to include in DID analyses as originally planned. <p>Household decision making and gender-related attitudes and behaviors</p> <ul style="list-style-type: none"> • There were no significant findings. <p>Child health</p> <ul style="list-style-type: none"> • HES-only saw increases in birth registration, from 79.3 percent to 89.1 percent ($p = 0.000$). The DID model found this increase was 3.9 percentage points greater than that in the control group ($p = 0.039$). • Caregiver nutrition knowledge increased significantly, from 15.9 percent to 32.1 percent ($p = 0.002$); the DID model showed HES-only had a borderline statistically significant 7.6 percentage point greater change than control ($p = 0.073$). • Caregivers' knowledge of a child's HIV status decreased slightly over time, from 18.8 percent to 14.8 percent, but was not statistically significant; the DID model showed a protective effect for HES-only, however, with those households being 3.9 percentage points higher relative to control ($p = 0.054$). <p>Youth health</p> <ul style="list-style-type: none"> • HES-only youth knowledge of HIV prevention strategies had a slight, not statistically significant increase over time; the DID model showed that HES-only youth had an 11 percentage point increase relative to control youth ($p = 0.002$). • HES-only youth reported increases in ever testing for HIV, from 18.4 percent to 29.1 percent ($p = 0.006$). The DID

Primary Evaluation Questions	Key Findings
	<p>model showed a 5.1 percentage point increase relative to control youth (p=0.037).</p> <p>Caregiver health</p> <ul style="list-style-type: none"> HES-only caregiver reports of ever testing for HIV increased from 42.4 percent to 60.4 percent (p=0.000). Because the control group experienced a greater increase, however, the DID model found that HES-only households had a 4.5 percentage point smaller increase compared to control. <p>Education</p> <ul style="list-style-type: none"> There was a borderline statistically significant decrease in child ECD attendance, from 13.6 percent to 8.6 percent (p=0.084). The DID model showed no program impact on ECD attendance. HES-only children progressing in school from the previous year decreased from 83.8 to 71.7 percent (p=0.000). The DID model showed no program impact on this outcome. Females fared better than males on the outcome of regular secondary school attendance (7 percentage points higher, p=0.003).
<p>3. Which of these two approaches is more cost-effective?</p>	<p>Economic strengthening</p> <ul style="list-style-type: none"> To increase the proportion of household consumption that is for food by 1 percent required \$12.59 in the full ISVP group under IR 2 and \$9.35 in the HES-only group. HES-only was more cost-effective. <p>Health</p> <ul style="list-style-type: none"> To increase caregiver's knowledge of a child's HIV status by 1 percent in the full ISVP intervention group compared to the control group requires \$7.84 per beneficiary, when considering the cost of the IR1 household interventions, or \$19.45 in the full package. HES-only did not show significant effects for this indicator. A 1 percent increase in youth reporting HIV testing cost \$25.93 per beneficiary in the full ISVP group, \$5.23 per beneficiary under only IR 3 programming in full ISVP compared to the control group, and \$5.68 per beneficiary in the HES-only group compared to the control group. HES-only was more cost-effective unless you use only the IR3 cost for full ISVP in comparison. <p>Schooled</p> <ul style="list-style-type: none"> No schooling outcomes showed significant effects so cost effectiveness calculations could not be carried out. <p>Safe</p> <ul style="list-style-type: none"> To see a 1 percent reduction in caregiver support for harsh punishment in school or home in the full ISVP group compared to the HES-only group cost an additional \$6.58 per beneficiary. A 1 percent reduction cost \$7.93 per beneficiary in the full ISVP package compared to the control, only slightly more than the cost to achieve the same impact when comparing the two intervention arms to each other. HES did not have program impact on this outcome, so it is unclear which approach is more cost-effective.

LIMITATIONS

The program components were rolled out 12 to 18 months before data collection, so program participants were already potentially experiencing changes because of the program at initial data collection; thus, the DID results for shorter-term outcomes may show less change for program arms than occurred in reality. In addition, the end line was timed to correspond to two to two-and-a-half years of program exposure and we expected many participants to have graduated by the time these data were collected. (Twiubake was envisioned as a two- to three-year program.) At end line, however, only a low percentage of participants reported having graduated from the program. Thus, the end line timing may have been too early in the outcome trajectory for the program to have benefitted participants meaningfully.

The basic assumption of DID analysis is that the program group would, in the absence of the program, have experienced a trend in outcomes parallel to that of the control group. It is not possible to test directly for violations of this assumption without pre-program data; as a solution, we controlled for differences in the trend experienced by the comparison group and what the program group would have experienced in the absence of the program. We did so with a regression version of the DID model.

Although we randomized the sectors into study groups, results from balance testing at the time of initial data collection suggest that the control group was different from the program groups, and the program groups were different from each other in some systematic ways. However, the DID model we used controlled for such differences.

Contamination is a concern. Control households reported some exposure to various services and information like that provided by the program, albeit at consistently lower levels than program arms; similarly, HES-only households reported exposure to such services and information, but at lower levels than that of full ISVP households. Contamination by government programs/donations was also found in all groups. Qualitative informants did not point to any programs that were systematically distributed in one arm versus another and randomization should mean that the effects of government interventions would have been equally distributed throughout the arms. However, this contamination may have made the group outcomes more similar than they would have been otherwise. It also negatively affects our confidence in attributing the observed changes to the programs.

All three groups had the same basic package of nutrition services offered by various projects under a nationwide government push, so nutrition-related results must be interpreted in light of this fact.

The sample size was powered to detect changes for two outcomes: progression from primary to secondary school and household hunger outcomes. The final study sample size was slightly smaller than planned, possibly limiting our ability to detect potential differences between groups for these outcomes if they were present.

Qualitative data from beneficiaries would have helped to elucidate some of the DID results.

Finally, there was a change in the local research implementing partner between initial data collection and end line. It is unlikely there were any systematic differences between fieldwork implementation, however, because there were no major changes in survey administration (tablet) or content between data collection points.

DISCUSSION

Despite the study limitations, the program was still able to affect significant change in multiple program areas. Each program arm saw positive impact for seven outcomes and negative impact for one; full ISVP also trended positive for an eighth outcome but just missed statistical significance. If the initial data collection had taken place prior to program start, full ISVP likely would have shown impact on at least two other outcomes, clearly outperforming HES-only.

HES-only provided some economic stability to households and contributed to improvements in the health and economic strengthening primary outcomes, and several secondary child and youth health outcomes. HES-only had a negative impact on caregiver report of HIV testing and no impact on educational outcomes. Full ISVP strengthened economic status and contributed to improvements in the health primary outcome, as well as several secondary child, youth, economic strengthening, gender-related attitudes, and caregiver health outcomes. Full ISVP households had better outcomes in savings and productive assets at end line compared to HES-only and control households. For several secondary outcomes, it appeared that full ISVP performed worse than HES-only and/or control. However, these were outcomes for which full ISVP started at a much higher level compared to the other groups (because of initial data collection taking place after program start) and remained higher at end line. The late initial data collection (after program start) meant that the full ISVP group had already started seeing changes in outcomes related to program services before the evaluation period. Had there been a “clean” baseline, the full ISVP group likely would have seen positive impacts for these outcomes. Full ISVP did not contribute to educational well-being and appeared to have negative impacts on youth HIV prevention strategies.

This study found moderate spending per beneficiary by the full ISVP program but, when combined with the resulting impact for specific outcome measures, the cost of achieving improvements constituted a relatively large proportion of national spending on health per person. Deciding whether it is worthwhile to spend this amount to achieve the corresponding outcomes is up to national governments, individual programs, or donors. There is evidence suggesting that some costly programs may be cost-effective in the long run by reducing more costly, negative outcomes later (Foster, Jones, & Conduct Problems Prevention Research Group, 2006).

When assessing the additive effects of the full ISVP program to the HES-only program, the impact was moderate, and the full ISVP program was only more cost-effective for achieving improvement in one outcome—youth reporting of HIV testing—and only if using IR3 costs alone. When interpreting cost-effectiveness analysis (CEA) ratios for the full ISVP, it is important to remember that the costs included encompass a large volume of activities that may be unrelated to the outcome being considered. In some ways, the full ISVP costs by IR area are more comparable to the HES-only group than the total package.

RECOMMENDATIONS

In light of these findings, we make the following preliminary recommendations:

- To improve education outcomes, future programming should consider different approaches or increased exposure to current approaches under the education result area; schooling outcomes were the only category of outcomes unaffected by the program in the current study.
- To better affect change in youth job readiness and improved employment, future programming should consider different approaches or increased exposure to current approaches under this pillar; related outcomes in the current study were very low.
- Future ISLG programs should consider how to address program participant concerns regarding the starting amounts for ISLG group participation; it was the primary reported reason for program drop-out.
- Future programming should address gender-related findings in the following ways:
 - Consider how to better support households with female heads to decrease household hunger; in this study we found full ISVP worked better for households with male heads.
 - Consider different approaches or increased exposure to current approaches to encouraging more equitable gender norms among boys; this evaluation did not find change in this area.
- USAID should consider conducting a qualitative study with beneficiaries six or more months after graduation to compare full ISVP and HES-only households; longer-term follow-up of beneficiaries would help to show whether full ISVP households achieve a better sustainability of outcomes, considering their greater savings and livestock ownership at end line; it would also help to clarify some of the results and capture more nuanced changes that the study was unable to capture. For example, education outcomes were unaffected in both arms, possibly because any impacts were just too small for our study to capture, but qualitative exploration could illuminate this issue.
- Considering the challenges this study encountered in the CEA, future complex OVC program CEA research should consider the following:
 - Using cost utility analysis to assess the benefits of OVC programs using standard and comparable measures such as disability-adjusted life years (DALYs) or quality-adjusted life years (QALYs) instead of natural units used in CEA, such as caregiver knowledge of HIV status (Husereau, et al., 2014). Use of nonnatural units would make comparison to other programs easier but does not necessarily solve the issue of outcome sensitivity in complex programs.
 - Combined outcome measures could also be considered when looking at the cost of complex, packaged service programs. Combined outcomes may better measure the wraparound effects that were difficult to identify or link to the costs in this study.
 - It might be prudent to consider studies set up in a similar way that have a narrower focus, randomizing beneficiary households to exposure to specific activities. For example, if the aim is to increase HIV testing, consider assessing the impact of risk assessments and referrals during home visiting compared to index patient testing. Costs in this context would have to be collected in an entirely bottom-up manner—an endeavor both more time consuming and expensive than the approach used in this study.

CONCLUSIONS

Contamination by government programs and timing limitations likely minimized the differences observed between groups. Despite this, the evaluation found positive impact for both program groups. HES-only drove change in the primary economic strengthening and health outcomes, whereas full ISVP shifted only the primary health outcome; the trend for the economic strengthening outcome in full ISVP was similar to that of HES-only, but just barely missed the statistical significance cutoff. The primary education outcome was not changed by either program. Both programs positively affected some secondary outcomes, including several child and youth HIV outcomes, and each program also uniquely affected a few outcomes. Several secondary outcomes appeared to be negatively impacted by full ISVP, likely because of the late initial data collection period (after program activities had begun). Last, full ISVP households had greater savings and productive assets at end line. Over time, these greater savings and productive assets may help full ISVP households better ensure sustainability of outcomes and achieve other outcome shifts.

When we assessed the additive effects of the full ISVP program compared to the HES-only program, the impact was moderate, and the full ISVP program was only more cost-effective for achieving improvement in one outcome—youth reporting of HIV testing in the past 12 months and knowing the result—and only if using IR3 costs alone. Because end line data collection took place before most households graduated from the programs; however, these results may not reflect the long term cost-effectiveness of the program. Follow-up with beneficiaries would help clarify the answer to this question.

Table 2. Summary of impact results, per DID estimation

Outcome	Program impact (regression coefficient)		
	Full ISVP vs. control	HES-only vs. control	Full ISVP vs. HES-only
Economic strengthening			
Households with moderate or severe household hunger	None (-0.050) p=0.122	Positive (-0.053)* p=0.043	None (-0.016) p=0.928
Households able to meet basic needs (school, food, major unexpected expenses)	None (0.020) p=0.658	None (0.045) p=0.314	None (-0.026) p=0.488
Food shares (% of household consumption for food)	Increase (0.026+) p=0.063	Increase (0.031*) p=0.023	None (-0.005) p=0.681
Health shares (% of household consumption for health)	Decrease (-0.009+) p=0.068	None (0.001) p=0.873	Decrease (-0.009*) p=0.041
Education shares (% of household consumption for education)	Decrease (-0.017***) p=0.000	None (-0.002) p=0.582	Decrease (-0.015***) p=0.000
Household owns livestock	Positive (0.037)* p=0.038	None (0.016) p=0.529	None (0.041) p=0.105
Household decision making and gender-related attitudes and behaviors			
Any physical or sexual violence against female caregiver	None (0.005) p=0.892	None (0.031) p=0.476	None (-0.025) p=0.502
Any physical violence against female caregiver	None (-0.021) p=0.584	None (-0.004) p=0.926	None (-0.017) p=0.629
Any sexual violence against female caregiver	None (0.023) p=0.456	None (0.024) p=0.477	None (-0.001) p=0.977
Joint or sole decision-making power on all six decisions	None (0.017) p=0.573	None (-0.003) p=0.911	None (0.020) p=0.453
Support for harsh punishment in school or at home	Decrease (-0.157**) p=0.001	None (-0.012) p=0.802	Decrease (-0.145***) p=0.000
Youth reporting high gender equitable beliefs, per Gender Equitable Men (GEM) Scale score	None (-0.013) p=0.797	None (-0.041) p=0.802	None (0.145) p=0.555
Health			
Child health			
Children ages 0–17 for whom caregivers report birth was registered at the sector level	None (-0.005) p=0.773	Increase (0.039*) p=0.039	Decrease (-0.044*) p=0.025
Nutrition knowledge	None (0.054) p=0.156	Increase (0.076*) p=0.073	None (-0.022) p=0.621
Children ages 0–59 months receiving growth	None (-0.028) p=0.529	None (0.009) p=0.830	None (-0.037) p=0.392

Outcome	Program impact (regression coefficient)		
	Full ISVP vs. control	HES-only vs. control	Full ISVP vs. HES-only
monitoring and nutrition services			
Percentage of children ages 6–59 months who meet minimum dietary diversity	None (0.026) p=0.536	None (0.047) p=0.285	None (-0.021) p=0.611
Caregiver's knowledge of child's HIV status	Increase (0.064^{***}) p=0.000	Increase (0.039[*]) p=0.054	None (0.026) p=0.152
Youth knowledge of HIV prevention behaviors	None (0.009) p=0.821	Increase (0.110^{**}) p=0.002	Decrease (-0.102^{**}) p=0.009
Youth (10- to 17-year-olds) ever tested for HIV (youth report)	Increase (0.048[*]) p=0.042	Increase (0.051[*]) p=0.037	None (0.001) p=0.890
Caregiver health			
Caregiver ever tested for HIV	None -0.019 (p=0.322)	Decrease -0.045 (p=0.037[*])	None 0.026 (p=0.111)
Education			
Early childhood development			
Children ages 36–59 months attending early childhood development program	Decrease (-0.138^{**}) p=0.009	None (-0.043) p=0.276	Decrease (-0.094⁺) p=0.085
Schooling			
Regular secondary school attendance among 13- to 17-year-olds	None (-0.027) p=-0.192	None (-0.027) p= 0.126	None (-0.001) p=0.977
School-age children who progressed in school from previous year	None (0.016) p=0.435	None (0.017) p=0.452	None (-0.002) p=0.940
Youth graduating from primary school and returning to secondary school	None (-0.086) p=0.244	None (-0.057) p=0.487	None (-0.028) p=0.682

+ Borderline statistically significant at p<0.10.

* Statistically significant at p<0.05.

** Statistically significant at p<0.01.

*** Statistically significant at p<0.01.

Table 3. Summary of key outcomes at initial data collection and end line

Outcome	Full ISVP		HES-only		Control	
	Initial	End line	Initial	End line	Initial	End line
Economic strengthening						
Households with moderate or severe household hunger	79	67.2***	72.4	69.3	66.4	62.6+
Households able to meet basic needs (school, food, major unexpected expenses)	54.7	62.3*	60.4	69.3+	60.5	68.5
Food shares (% of household consumption for food)	49.4	46.5**	51.5	49.7+	54	50.4+
Health shares (% of household consumption for health)	6.3	4.6***	5.1	4.5	4.7	4.2+
Education shares (% of household consumption for education)	4.2	2.3***	2.2	1.7	2.2	2.1
Household owns livestock	58	66.9***	57	59.9	61.4	65.9+
Household decision making and gender-related attitudes and behaviors						
Any physical or sexual violence against female caregiver	18.7	20.4	25.6	24.3	19.8	20.5
Any physical violence against female caregiver	14	14.6	21	17.4	14.6	18
Any sexual violence against female caregiver	10.7	12.5	11.8	13.4	9.7	8
Joint or sole decision-making power on all six decisions	77.9	76.7	84.3	80.7	76.7	75.8
Support for harsh punishment in school or at home	46.3	32.4***	36.8	34.8	39.5	43
Youth reporting high gender equitable beliefs, per GEM Scale score	24.9	20.7	15.8	19.8	16.3	20.9
Health						
Child health						
Children ages 0–17 for whom caregivers report birth was registered at the sector level	85.2	90.8***	79.3	89.1***	84.8	91***
Nutrition knowledge	16.3	15.9**	15.9	32.1**	11.5	24.6***
Children ages 0–59 months receiving growth monitoring and nutrition services	71.1	83.2***	60.9	83.1***	66.9	77.7***
Percentage of children ages 6–59 months who meet minimum dietary diversity	32.3	61.5***	34.4	50.5+	34.3	56.4***
Caregiver's knowledge of child's HIV status	30	26.8+	18.8	14.8	16.9	10.5***
Youth knowledge of HIV prevention behaviors	40.9	35.5	32.5	36.3	39.2	34.6
Youth (10- to 17-year-olds) ever tested for HIV (youth report)	26.7	33.4*	18.4	29.1**	18.4	15.9

Outcome	Full ISVP		HES-only		Control	
	Initial	End line	Initial	End line	Initial	End line
Caregiver health						
Caregiver ever tested for HIV	63.1	68.1	42.4	60.4***	39.9	60.4***
Education						
Early childhood development						
Children ages 36–59 months attending early childhood development program	38.1	33	13.6	8.6+	11.6	15.4
Schooling						
Regular secondary school attendance among 13- to 17-year-olds	17.7	18.9	9.6	10.7	16.1	16.9
School-age children who progressed in school from previous year	90.7	79.1***	83.8	71.7***	89.1	75.7***
Youth graduating from primary school and returning to secondary school	13.0	10.0*	6.0	5.9	9.3	8.2

+ Borderline statistically significant at $p < 0.10$.

* Statistically significant at $p < 0.05$.

** Statistically significant at $p < 0.01$.

*** Statistically significant at $p < 0.01$.

REFERENCES

- Foster, M. E., Jones, D., & Conduct Problems Prevention Research Group. (2006). Can a costly intervention be cost-effective? An analysis of violence prevention. *Archives of General Psychiatry*, *63*(11), 1284–1291. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/17088509>.
- Husereau, D., Jacobs, P., Manns, B., Hoomans, T., Marshall, D., & Tamblyn, R. (2014). *Economic evaluation of complex health system interventions: A discussion paper*. Alberta, Canada: Canadian Institute of Health Economics. Retrieved from <https://www.ihe.ca/publications/economic-evaluation-of-complex-health-system-interventions-a-discussion-paper>.
- McIntosh, A., & Zeitlin, A. (2018). *Benchmarking a child nutrition program against cash: Experimental evidence from an evaluation in Rwanda*. Impact evaluation report. Retrieved from <https://www.poverty-action.org/sites/default/files/publications/Benchmarking.pdf>.

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

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. TR-19-363

ISBN: 978-1-64232-171-5

