APMI TANZANIA PRE-ANALYSIS PLAN

IDinsight | Revised February 24, 2020

Authors

Sylvia Jebiwott: sylvia.jebiwott@IDinsight.org
Maureen Stickel: maureen.stickel@IDinsight.org
Torben Fischer: torben.fischer@IDinsight.org
Ari Gandolfo: ari.gandolfo@IDinsight.org

About IDinsight

IDinsight uses data and evidence to help leaders combat poverty worldwide. Our collaborations deploy a large analytical toolkit to help clients design better policies, rigorously test what works, and use evidence to implement effectively at scale. We place special emphasis on using the right tool for the right question, and tailor our rigorous methods to the real-world constraints of decision-makers.

IDinsight works with governments, foundations, NGOs, multilaterals and businesses across Africa and Asia. We work in all major sectors including health, education, agriculture, governance, digital ID, financial access, and sanitation.

We have offices in Bengaluru, Dakar, Johannesburg, Lusaka, Manila, Nairobi, New Delhi, San Francisco, and Washington, DC. Visit www.IDinsight.org and follow on Twitter @IDinsight to learn more.

CONTENTS

Introduction 4

Evaluation Design 5

Analytical Framework 10

Corrections to the Analysis 15

Next Steps 17

Appendix A: Program Overview and TOC 18

Appendix B: Matching Approach 20

Appendix C: Survey Questions by Indicator 23

Appendix D: Calculating the Poverty Probability Index 27

Appendix E: Calcuating the Minimum Dietary Diversity Score 28

Appendix F: Calcuating the Modified A-WEAI 30

Appendix G: Calculating the Women and Poultry Ownership Score 32

ACRONYMS

3DE	3 Dimensions of Empowerment
5DE	5 Dimensions of Empowerment
STL	Silverlands Tanzania Limited
APMI	African Poultry Multiplication initiative
A-WEAI	Abbreviated Women's Empowerment in Agriculture Index
BMGF	Bill and Melinda Gates Foundation
DDS	Dietary Diversity Score
GPI	Gender Parity Index
IEDD	Impact Evaluation Design Document
IYCF MDD	Infant and Young Children Feeding Minimum Dietary Diversity
MU	Mother Unit
PAP	Pre-Analysis Plan
PPI	Poverty Probability Index
RIDIE	Registry for International Development Impact Evaluations
SHF	Small Holder Farmer
TOC	Theory of Change
WPF	World Poultry Foundation
WPO(S)	Women and Poultry Ownership (Score)
WTP	Willingness to Pay

INTRODUCTION

OBJECTIVE

The enclosed Pre-Analysis Plan (PAP) outlines the objectives, methods, and analytical framework for the IDinsight impact evaluation of the Africa Poultry Multiplication Initiative (APMI). Here, we focus on the evaluation of the APMI program in **Tanzania,** specifically looking at the impact of owning Sasso (an improved chicken breed distributed by Silverlands Tanzania Limited) on SHFs.

AUDIENCE

This PAP is meant for distribution with stakeholders to the APMI program, and will be registered with the Registry for International Development Impact Evaluations (RIDIE).

The preceding deliverable, the Impact Evaluation Design Document (IEDD), discusses in detail our objectives, research questions, study design, and sampling methods. We review these in brief in the PAP and then focus on the technical execution of our analysis, such as site selection, indicator lists, and impact estimators.

For reference, a summary of the program and our Theory of Change (ToC) are provided in **Appendix A**.

NOTE TO STAKEHOLDERS

This PAP is largely similar to what was previously shared for the Amo Farms evaluation in Nigeria. The main sections that differ include the study community and sample selection sections. These sections have been modified given the different approach used to select study communities and the study sample in Tanzania.

EVALUATION DESIGN

RESEARCH QUESTIONS

We explore the causal impact of the APMI program on smallholder farmers (SHFs), with a focus on flock performance, income, nutrition, and women's empowerment. **Table 1** outlines the primary and secondary research questions for each of these four dimensions.

Table 1: Primary and secondary research questions[1]

Category	Primary Question	Secondary Question
Flock Performance	To what extent does owning Sasso chickens affect poultry related outputs (eggs and meat)?	How do Sasso chicken breeds compare to local chicken varieties on measures of mortality and resources required to care for them (including time)?
Income	How does owning Sasso chickens affect income generated from poultry related outputs?	What portion of the household income is from poultry production?
Nutrition	Has household consumption of eggs and chicken meat increased as a result of owning Sasso chicken varieties?	To what extent are women's and children's diets more diverse as a result of owning Sasso chickens?
Women's Empowerment	Does owning Sasso chickens impact women's empowerment?	To what extent do women SHFs have more ownership and agency in poultry decisions?

These research questions consider the **direct effects** of Sasso chickens on SHFs who own them (termed "Treatment on the Treated"). For this reason, we propose an evaluation design that focuses on estimating these direct effects.

STUDY DESIGN

To establish a counterfactual, IDinsight will **match one-to-one** treatment SHFs[2] to control SHFs who are comparable on a set of baseline characteristics. The baseline characteristics will be a combination of baseline outcomes and other characteristics expected to predict the propensity to purchase Sasso chickens.[3] At endline, we will compare outcomes of matched treatment and control SHFs to estimate the causal effect of owning Sasso chickens.

This design assumes that outside the observable characteristics on which we match, there are no other unobservable factors which correlate with both purchasing Sasso chickens and our outcomes of interest. We mitigate this possibility by identifying from the outset control SHFs who would be likely to purchase Sasso chickens (see more on sampling details below).

[1] The study aims to identify the effect of owning (some) Sasso chickens in comparison to not owning any Sasso chickens, accounting for characteristics such as baseline ownership of local chickens and average flock size.
[2] We define treatment SHF as any SHF confirmed to be a first-time buyer of Sasso chickens.
[3] We will choose the matching technique that generates the best balance across treatment and control SHFs on baseline outcomes and other covariates predictive of endline outcomes. This matching analysis will be conducted once baseline data collection is complete. However, potential matching techniques and the set of baseline outcomes and covariates we are considering can be found in **Appendix B.**

STUDY SITE AND COMMUNITY SELECTION

In conjunction with Silverlands, we identified ten regions in Tanzania for the focus of this research. The ten regions are: Arusha, Kilimanjaro, Tanga, Ruvuma, Mtwara, Lindi, Dodoma, Morogoro, Njombe and Mbeya. The ten regions represent diverse geographical zones and are of strategic importance to Silverlands.

Within each region, treatment communities were identified based on where Silverlands teams planned to market Sasso. In selecting where to market Sasso, the teams prioritized 1) communities within 50km radii of Silverlands distribution centres to ensure SHFs have uninterrupted access to Sasso during the two years before the endline evaluation, and 2) communities with high demand for Sasso where teams would find new Sasso customers. Using these criteria and working closely with Silverlands field staff, we identified 110 treatment communities that are evenly distributed across the ten regions (11 treatment communities per region).

Control communities were identified in a step-wise process using administrative data, information on the location of treatment communities, and feedback from field teams. Starting with ward-level administrative data, we dropped all wards that had Sasso exposure as defined by 1) the presence of AKM Glitters[4] or Silverlands Mother Units (MUs)[5] in the ward, and/or 2) the likelihood that Silverlands disseminated information about Sasso chickens to surrounding communities. The resulting list of unexposed wards was further trimmed by only retaining wards similar to treatment wards in terms of location, population size, population density and ward type (rural, urban or mixed).[6]

The tentative control community wards were then verified with field teams to ensure there was no uncaptured Sasso exposure. We also checked with the teams whether they would consider marketing Sasso in these wards if they had the opportunity. To get the final control communities, we randomly sampled villages from the wards that did not have Sasso exposure and where teams would consider marketing Sasso. In total we selected 70 control communities distributed evenly across the ten regions (seven control communities per region).

The list of study communities was modified slightly before the start of baseline data collection, mostly due to changes in administrative units.[7] For example, we dropped three communities that are presently in Songwe, a region carved out of Mbeya in 2016. We re-selected replacement treatment and control communities in Mbeya following the same process outlined above. We therefore do not

---

[4] We excluded wards where AKM Glitters has a presence because ownership of Kuroiler (an improved breed of chicken distributed by AKM Glitters) by control households would reduce impact of Sasso that we would observe.

[5] Mother units receive day-old chicks from Silverlands, rear them for five weeks and then sell them to SHFs.

[6] For some treatment communities we could not find matching control wards. In these cases, we identified the closest control wards or oversampled communities in matching wards that had already been selected.

[7] The administrative data used is from the 2012 census (link: https://data.humdata.org/dataset/2012-census-tanzania-wards-shapefiles). Some administrative units have since been created or names modified.

expect these re-selected communities to exhibit different characteristics and/or comprise farmers with different unobservable characteristics.[8]

SAMPLE SELECTION

The population of interest comprises prospective *buyers (treatment)* and *would-be buyers* of Sasso chickens *(control)* if they had access.

Treatment sample

In treatment communities, we aim to sample SHFs who are likely to purchase Sasso chickens.[9] To establish our sampling frame, we will attend official Silverlands outreach events in communities and deliver a 3-4 minute eligibility survey to one participant per attending household. The eligibility survey gathers basic identifying information, and assesses general willingness to pay (WTP) for a new breed of poultry.[10] Of those SHFs with adequate WTP[11], we will randomly select 40 (plus 5 replacements)[12] to participate in the main household survey.[13]

In the minority of cases, the treatment SHF represents an actual buyer of Sasso chickens. First time buyers of Sasso who order Sasso during the Sasso sensitization events will automatically make our sample.

We aim to survey treatment SHFs within 6 weeks[14] from their date of attending marketing events. As most treatment SHFs will be first time buyers of Sasso, we need to minimize any Sasso-related costs and benefits felt by SHF owners at baseline.[15]

Control sample

We aim to establish a control sample of SHFs who would be likely to purchase Sasso, if given access.

We identify the sampling frame in control communities by mimicking the process by which Silverlands conducts outreach for Sasso. However, we do not include content that would sway participants' interest in poultry (such that they would seek out the Sasso), or affect their baseline outcomes (e.g. investing more in nutrition).

---

[8] As this an individual matching design, it is not imperative to have totally comparable treatment and control communities, though we will aim to have communities that are as similar as possible to increase the likelihood of good matches.
[9] We also refer to this approach as 'prospective sampling.'
[10] The survey includes other WTP questions so as to not link the research with poultry specifically.
[11] A five-week old Sasso sells for 4500 Tanzania shillings. Respondents willing to pay >4500 are automatically included in our sample.
[12] This number includes a generous buffer in case some treatment SHFs do not ultimately purchase Sasso chickens. If less than this number of eligible SHFs is identified in the listing exercise, all eligible SHFs will be sampled.
[13] In the case where <40 SHFs in a community have adequate WTP, we will oversample in other communities with >40 SHFs with adequate WTP.
[14] SHFs purchase Sasso chickens at 5 weeks of age. At 15 weeks of age, Sasso cocks reach maturity to sell for consumption. Sasso hens reach laying-age at 22 weeks. Some SHFs may choose to sell Sasso cocks slightly before 15 weeks, so we optimally aim to survey communities at 7 weeks after purchase.
[15] All expenses accrued from Sasso ownership in the interim (before surveying) will be backed-out in the analysis to understand true baseline levels of poultry expenditures.

As such, we will hold generic community events that (a) encourage attendance from more enterprising individuals and (b) avoid content directly relating to nutrition and poultry.[16] At community events, we will deliver the same eligibility survey delivered in treatment communities.

Respondents with adequate WTP scores are inducted into the short-list of eligible respondents as a part of our sampling frame. Of those, 25 (plus 5 replacements) will be randomly selected for inclusion in the main household survey.[17]

Altogether, **the baseline sample is planned to comprise approximately 6150 SHFs**: 4400 treatment SHFs and 1750 control SHFs. Ahead of endline, we will conduct a monitoring exercise to identify treatment SHFs who have actually purchased Sasso chickens as well as verify that control SHFs did not have access to Sasso during the study period. Verified SHFs will then be matched for the final endline sample. We anticipate the endline sample (after matching) to include approximately 1400 SHFs, involving 600 treatment SHFs and 800 control. This is also outlined in the table below:

Table 2: sample sizes in stages

	Treatment SHFs	Control SHFs	Total SHFs
1. Baseline sample size (with buffers)	4400	1750	6150
2. Eligible SHFs after monitoring[18] (minimum)	700	1400	2100
3. Matched SHFs (minimum)	600	600	1200
4. Endline sample size (with buffers[19]) *Includes IYCF MDD[20] sample size*	600 *(300)*	800 *(300)*	1400 *(600)*

Error! Reference source not found. provides the minimum detectable effect sizes (MDES)[21] for each of the main outcome indicators given an endline sample size of 600 matched SHF pairs (or 1,200 SHF in total). These calculations assume conventional levels of significance at 0.05 and power at 0.8.[22] Moreover, we anticipate obtaining child measurements in the relevant age group for about 50% of sampled SHFs at endline. This assumption implies that the MDES for the child minimum dietary diversity (IYCF MDD) is based on a sample size of 300 matched SHF pairs (or 600 SHF households in total).

[16] In coordination with local collaborators, we selected a module entitled Best Practices in Micro-Enterprise for delivery at community events. IDinsight can share this content with interested stakeholders upon request.

[17] Sampling is conducted in two stages. In the first stage, we classify SHFs based on their WTP for a 5-week old Sasso chicken. Respondents are eligible if they are willing to pay (slightly less than or above) the usual price for such a chicken. In stage two of the sampling process, we draw a random sample of SHFs stratified by SHFs having children in the relevant age group in their household. In communities where less than 30 eligible SHFs have been identified in step one, all SHFs will be sampled.

[18] We anticipate the sample size to decrease slightly for both treatment and control SHFs as some treatment SHFs will likely not have purchased Sasso chickens and some control SHFs will likely have received access to the Sasso chickens.

[19] Buffers refer to the extra SHFs we will include in the endline sample in anticipation of re-matching SHFs that attrit, and matching SHFs with children at endline (see Appendix B for more details).

[20] IYCF MDD refers to the Infant and Young Children Feeding Minimum Dietary Diversity

[21] MDES is the minimum difference we will be able to see between treatment and control SHFs.

[22] The level of significance determines the probability of concluding that a measured effect is statistically significant, when in fact it is not. In other words, if the program actually had no effect, there still is 5% probability that we would falsely conclude that the program was effective. Power is defined as the probability of concluding that there is not a statistically significant effect, when the program is in fact effective. 0.05 significance and 0.8 power are the standards for causal impact evaluations. Further assumptions for the sample size calculations are provided in the IEDD document(Appendix E).

Table 2: Minimum Detectable Effect Sizes (MDES) for main outcomes given a sample of 1,200 SHFs

Outcome Category	Outcome Indicator	Minimum Detectable Effect Size (MDES)
Flock productivity	Eggs produced in last 7 days	1.93 (eggs)
Income	Income chicken sales in last 6 months	3 (USD)
Nutrition (IYCF MDD)	Child Minimum Dietary Diversity	11.5 (percentage points)
Nutrition (Women)	Minimum Dietary Diversity for Women	~ 8 (percentage points)
Women Empowerment	Share of empowered women	~ 8 (percentage points)

DATA COLLECTION

IDinsight will collect (a) survey data to measure outcomes of interest and (b) monitoring data to assess program implementation:

- Survey data will be collected at the household-level on outcomes of interest. The primary respondent will be the caretaker of poultry in the home, who may or may not be the head of household. A portion of the survey (the supplemental survey) is also administered to an adult of the opposite gender, often the spouse of the primary respondent.

- Monitoring data will be collected to (1) ensure companies adhere to protocol for treatment and control areas, (2) gauge whether treatment SHFs purchase and/or replenish their flock sizes over time,[23] and (3) monitor the sell prices of mature Sasso chickens over time.

Baseline data collection is anticipated to be conducted between November 2019 and January 2020. Monitoring activities are planned for 2020. We anticipate endline data collection will be conducted in 2021.

[23] In treatment communities in which likely potential buyers of Sasso have been surveyed at baseline, the monitoring exercise allows us to confirm SHFs who have actually purchased Sasso chickens and should therefore be included in the endline sample. We shall also monitor flock size as increasing Sasso flock size over time is important to SHFs achieving the expected gains in income, nutrition, and empowerment.

ANALYTICAL FRAMEWORK

OUTCOMES

Table 4 below describes the primary and secondary indicators used to measure the impact of Sasso chickens on SHF livelihoods. **These indicators are the same as those in the Amo Farms' PAP**. The individual variables and survey questions that form each indicator are provided in Appendix C. Primary outcomes are considered of main importance (and thus included in the multiple hypotheses correction procedure), while secondary outcomes are exploratory. Primary indicators reflect IDinsight's understanding as to when the APMI program (including investment in Sasso chicken) would be considered successful from a social impact perspective. Secondary indicators allow us to answer the above research questions and provide monitoring information that can be included in result trackers reported to the client, the Bill and Melinda Gates Foundation (BMGF). Statistically significant findings in secondary outcomes alone will not drive main recommendations to the broader stakeholder group especially if there are no statistically significant findings in the primary outcomes.

Table 4: Primary and secondary outcome indicators used in analysis

Dimension		Priority	Outcome	Indicator (at unit level)	Unit of analysis	Type[24]
Productivity/performance[25]	1.	Primary	Egg production	Number of eggs produced in the last 7 days	Household	Continuous
	2.	Secondary	Egg production (per chicken)	Number of eggs produced in last 7 days per chicken	Household	Continuous
	3.	Secondary	Flock size	Total number of chickens owned	Household	Continuous
	4.	Secondary	Egg production	Any eggs produced in the last 7 days	Household	Binary
	5.	Secondary	Bird mortality rate	Number of chickens lost to disease in last 6 months divided by largest flock size in last 6 months	Household	Continuous
	6.	Secondary	Financial expenditure	Financial expenditure on chickens in last 30 days	Household	Continuous
	7.	Secondary	Time expenditure	Time spent rearing chickens in last 7 days	Household	Continuous
Income from poultry	8.	Primary	Income from poultry production	Income from egg and chicken sales in last 30 days	Household	Continuous
	9.	Secondary	Income from eggs	Income from egg sales in last 30 days	Household	Continuous

[24] We examine four types of indicators: binary refers to a (0,1) variable; continuous refers to continuous numeric variable (e.g. $5.50 in egg income); categorical to a defined string variable (e.g. preference to purchase Sasso in the planting season); index to a range-specified value (oftentimes aggregated from numerous sub-indicators).
[25] Indicators relating to poultry performance, production, and income will be assessed for both local and Sasso chicken varieties (if owned by the household).

	10.	Secondary	Income from chicken	Income from chicken sales in last 6 months	Household	Continuous
	11.	Secondary	Profitability of rearing chickens	Average profit (revenue – expenses) from chickens in the 30 days prior to endline (both including and excluding own consumption)	Household	Continuous
	12.	Secondary	Proportion of income from chicken	Perception of whether role of poultry income in overall household revenue is greater than 25%	Household	Binary
	13.	Secondary	Egg seller	Households who sold eggs in last 30 days	Household	Binary
	14.	Secondary	Chicken seller	Households who sold chickens in last 6 months	Household	Binary
Nutrition	15.	Primary	Child meets minimum diet	Child minimum dietary diversity (24-hour recall)	Individual	Binary[26]
	16.	Primary	Woman meets minimum diet	Woman minimum dietary diversity (24-hour recall)	Individual	Binary[27]
	17.	Secondary	Egg consumption	Number of eggs consumed by the household in the last 7 days	Household	Continuous
	18.	Secondary	Chicken consumption	Number of chickens consumed by the household in last 30 days	Household	Continuous
	19.	Secondary	Children's egg consumption	Number of eggs consumed by children aged under 5 in the last 7 days	Household	Continuous
	20.	Secondary	Children's egg consumption	Child under 5 consumed eggs in the last 24 hours	Individual	Binary
	21.	Secondary	Children's Chicken Meat consumption	Child under 5 consumed chicken meat in last 24 hours	Individual	Binary
	22.	Secondary	Total food expenditures	Household food expenditure in last 7 days	Household	Continuous
	23.	Secondary	Household level spending on diverse food groups	Number of food groups that the household spent a considerable amount of money on in the last 7 (30) days[28]	Household	Count
Women's empowerment[29]	24.	Primary	Female empowerment score (3DE)	Female weighted empowerment score	Individual	Index, scale [0,1]

[26] See Appendix E. Only calculated if children in the relevant age group are present in the household.
[27] See Appendix E. Only calculated if an adult woman is present in the household.
[28] We will define the per-capita expenditure cut-off as the 10% quantile of the respective food group expenditure in the control group at endline.
[29] Indicators adapted from the A-WEAI index. See Appendix C and D for more calculation details.

	25.	Secondary	Gender Parity Index	Household has equal empowerment scores between male and female	Individual	Index, binary
	26.	Secondary	Female empowerment score (3DE): Input into productive decisions domain	Female empowerment related to input into productive decisions	Individual	Index, binary
	27.	Secondary	Female empowerment score (3DE): Ownership of assets	Female empowerment related to ownership of assets	Individual	Index, binary
	28.	Secondary	Female empowerment score (3DE): Control over use of income	Female empowerment related to control over use of income	Individual	Index, binary
	29.	Secondary	Gender Parity Index: Input into productive decisions domain	Household has equal empowerment scores between male and female related to input into productive decisions	Individual	Index, binary
	30.	Secondary	Gender Parity Index: Ownership of assets:	Household has equal empowerment scores between male and female related to ownership of assets	Individual	Index, binary
	31.	Secondary	Gender Parity Index: Control over use of income	Household has equal empowerment scores between male and female related to control over use of income	Individual	Index, binary
	32.	Secondary	Women's ownership in poultry-rearing	Women and poultry ownership score (WPOS)	Individual	Index, scale [0,1]
Child Health	32.	Secondary	Childhood morbidity	Child morbidity score	Individual	Index, scale [0,1]

ANALYTICAL MODEL

We will estimate the impact of Sasso chicken ownership on the above outcomes using a difference-in-difference (DID) specification. DID is our preferred specification as it allows us to factor out time constant unobservable traits of SHFs that might be correlated with purchasing a Sasso, thus reducing potential bias.[30] In this matching design the results are valid under the assumption that conditional on the observables on which we matched, there are no other confounders correlated with purchasing a Sasso and outcomes, this feature of DID. However, if we have fewer than 600 matched pairs (our target sample size) at endline, we will use an ANCOVA specification, which has statistical power gains over a difference-in-difference specification.[31]

Our difference-in-difference specification is:

$$\Delta Y_{s,c} = \beta_0 + \beta_1 * T_{s,c} + \delta_s + \varepsilon_{s,c}$$

- $\Delta Y_{s,c}$ denotes the change in outcomes over time = $Y_{s,c\ \text{endline}} - Y_{s,c\ \text{baseline}}$ for SHF s in community *c*
- $T_{s,c}$ is the treatment status of SHF *s* in community *c*
- δ_s denotes a vector of dummy variables corresponding to each matched pair of SHFs
- $\varepsilon_{s,c}$ is the error term for SHF *s* in community *c,* clustered at the community-level[32]

Our ANCOVA specification is:

$$Y_{Es,c} = \beta_0 + \beta_1 * T_{s,c} + \beta_2 * Y_{Bs,c} + \beta^* * \gamma_{s,c} + \varepsilon_{s,c}$$

- $Y_{Es,c}$ denotes the endline outcome for SHF *s* in community *c*
- $T_{s,c}$ is the treatment status of SHF *s* in community *c*
- $Y_{Bs,c}$ denotes the baseline outcome for SHF *s* in community *c*
- $\gamma_{s,c}$ is a vector of baseline covariates that will be used in the matching algorithm[33]
- $\varepsilon_{s,c}$ is the error term for SHF *s* in community *c,* clustered at the community-level

Results will be presented in a combination of outcome-specific regression tables and bar graphs with relevant error margins.

For all child-specific outcomes, we will run an ANCOVA specification similar to the one above irrespective of sample size. We propose this analytical model for child-specific outcomes because we expect a significant proportion of households that report to have children in the relevant age group (6 months to 5 years) at endline might not have reported children in that age group at baseline.[34] For these households, baseline outcomes would be missing, invalidating the difference-in-difference approach. In the ANCOVA specification, we will replace missing baseline values for child outcomes with 0 and include an additional binary variable indicating whether the outcome is missing.

---

[30] Simulations have shown that a difference-in-difference specification has an advantage for removing confounding bias over an ANCOVA specification when the bias is large: https://declaredesign.org/blog/2019-01-15-change-scores.html.

[31] McKenzie, David. "Beyond baseline and follow-up: The case for more T in experiments." *Journal of development Economics* 99.2 (2012): 210-221.

[32] We cluster standard errors above the level of treatment, i.e. at the level of the community, to account for community-level shocks that may influence outcomes for all residing SHFs (i.e. correlated outcomes within a particular community). We do this to be able to provide a population estimate of the impact of purchasing Sasso, as it is likely that this effect might depend on community-level characteristics such as general purchasing power of the village. See more details at https://declaredesign.org/blog/2018-12-18-cluster-level.html.

[33] See Appendix B for more details about the matching algorithm.

[34] We anticipate obtaining child measurements in the relevant age group for about 50% of sampled SHFs at endline. This assumption implies that the minimum detectable effect sizes for the child minimum dietary diversity (IYCF MDD) is based on a sample size of 300 matched SHF pairs (or 600 SHF households in total).

SUB-GROUP ANALYSIS

All sub-group analyses are considered exploratory.

In addition to the above analysis, we will estimate heterogeneous treatment effects for primary outcomes indicators for the following subgroups:

Table 5: Sub-groups used in analysis

	Sub-group definition	Justification
1.	Geographic location of household (northern, coastal, and southern zones)	Systematic socio-cultural differences between zones may influence poultry rearing
2.	Poultry owner's educational attainment (above/below secondary)	Educational level may be correlated with poultry rearing knowledge
3.	PPI[35] scores (Quintiles of PPI score)	Wealth and stability may be correlated with flock productivity
4.	Female versus male primary caretaker of poultry	Gender of primary poultry caretaker may affect women's empowerment scores
5.	Flock size (greater/fewer than 20 birds at endline)	Greater flock sizes may correlate to greater income and/or nutrition from poultry
6.	Households that owned chickens prior to baseline	Prior ownership may be correlated with poultry rearing knowledge
7.	Households that own Sasso within 6 months of endline	Consistent renewal of flocks may affect benefits seen from poultry

If a given sub-group definition in Table 5 would lead to too little variation in the sample (i.e. 90% or more of respondents would fall into one of the categories), we will not conduct that sub-group analysis.

Note that subgroups 5 and 7 are not based on pre-treatment covariates and therefore do not give "heterogeneous" treatment effects in the classical sense. As a consequence of this personal choice (which might be affected by the intervention), there might be selection bias leading to different types of SHFs in the subgroups. While the results for subgroups 5 and 7 will be only indicative, we still believe that stakeholders are interested in understanding how program effects might differ for these sub-groups.

[35] The Poverty Probability Index (PPI) is a standardized poverty metric that comprises 10 social and financial indicators with score-based responses. The summation of scores calculates the probability of experiencing poverty on various national and international poverty scales

CORRECTIONS TO THE ANALYSIS

CONTAMINATION

Contamination occurs when the control group (erroneously) obtains access to the treatment intervention. In this study, SHFs in control villages are contaminated if they receive access to and/or purchase Sasso chickens. If control SHFs purchase Sasso chickens, they are likely to experience the same effects as treatment SHFs, leading us to underestimate the program's effects.

In order to mitigate the likelihood of contamination, we will share the list of control communities with Silverlands field teams and request that they do not actively market Sasso in these control communities. Throughout monitoring we will contact a random group of respondents and/or visit communities to assess if there has been exposure. If fewer than 15% of respondents in a control community report having purchased Sasso, we will drop the contaminated respondents from the sample (up to four respondents). If greater than 15% of respondents report having purchased Sasso, we will drop the entire community from the study.

RE-MATCHING

Re-matching may be necessary due to SHF attrition by the endline follow-up. Attrition can occur when study participants die, move, or refuse to participate in the study, and are therefore not included in the endline data collection.

We follow the below re-matching procedure:

1) Following the baseline, we will generate treatment-control SHF matches (see Appendix B for more details). We will also identify a group of control SHFs that are close but not final matches that we can survey at endline in anticipation of SHF attrition.
2) If at endline we find that the control SHF of the match has attrited, we will replace that SHF with a highly similar SHF from the set of buffer SHFs interviewed. In the re-matching process, we will apply the same algorithm used to establish the initial matches.
3) If the treatment SHF has attrited, we will likely need to drop that pair of SHFs given we will not have a treatment SHF buffer group. We anticipate needing to include all SHFs who purchase Sasso in our treatment sample.

MISSING VALUES

Missing values can take the form of item non-response (e.g. uncompleted surveys or missing responses), partial response (e.g. "Don't know" responses), or errors in the data. We do not anticipate major issues with item non-response due to the constraints built into our electronic data collection. Additionally, we run regular checks on the incoming data to monitor the frequency of "Don't know" / "Refusal" responses and errors in the data. These checks allow us to quickly follow up with enumerators for appropriate corrective measures.

For covariates with missing values, we will employ a dummy variable adjustment, whereby we create a dummy variable to indicate whether the value for that covariate is missing (0 = missing, 1 = non-missing), and all missing values are set to the same value (0). This method will allow us to recover an unbiased estimate of the impact of purchasing Sasso.[36]

[36] Groenwold, Rolf HH, et al. "Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis." *Cmaj* 184.11 (2012): 1265-1269.

For outcomes with missing values, if we find that a given outcome has fewer than 10% of responses missing (e.g. "Don't know" or "Refuse"), we will drop these observations from the analysis of that outcome. If there are greater than 10% of responses missing for a given outcome, we will check the sensitivity of the complete case analysis with imputation methods (e.g. multiple imputation).

MULTIPLE HYPOTHESIS CORRECTION

IDinsight's understanding is that stakeholders expect the APMI program to simultaneously shift the primary outcomes specified above. From our discussions with BMGF, we understand that positive findings for any of the nutrition, income, and empowerment outcomes might be interpreted as a success of the intervention. While performance / productivity is a key outcome for stakeholders, this outcome is an intermediate outcome that leads to the other three social impact outcomes. Consequently, we do not include performance / productivity in the family of social impact outcomes. As there are multiple outcomes that could inform success of the intervention, the hypotheses tested in this impact evaluation cannot be considered independent. This necessitates a correction for testing multiple hypotheses.

In order to reduce the risk of incorrectly rejecting the null hypothesis (Type I error), we will adjust standard errors using the family-wise error rate (FWER) using the free-step down procedure proposed by Westfall and Young (1993).[37] This adjustment treats hypothesis tests of similar outcomes as correlated rather than independent.[38] Conventional, uncorrected p-values will also be reported. Table 6 below summarizes the primary outcomes for which we will adjust standard errors, and indicates the number of hypotheses being tested with each "family" of outcomes.[39]

Table 6: Outcome families for multiple hypothesis correction

Outcome Family	Theme	Indicators	Total hypotheses tested
Social Impact Outcomes	Income from poultry	Income from poultry production	3
	Nutrition	Woman meets minimum diet	
	Women's empowerment	Female empowerment score (3DE)	

[37] Westfall, P., and S. Young. 1993. Resampling-based Multiple Testing: Examples and Methods for p-value Adjustment. John Wiley & Sons, Inc.

[38] If outcomes are correlated, the proposed FWER correction exhibits larger statistical power compared to other methods. The correction relies on re-sampling the treatment variable through reproducing the treatment assignment process. Given that the study design is a matching design and treatment is considered random conditional on the matching variables, we will permute treatment status within matched pairs.

[39] We exclude the child nutrition indicator from the social impact outcome family, despite being a primary outcome, because it is measured using a separate sample of matched SHFs. It is not easily possible to combine outcome indicators from separate samples into the proposed FWER correction.

NEXT STEPS

This PAP will be registered with the Registry for International Development Impact Evaluations (RIDIE).

We aim to share a baseline report with stakeholders by the end of the March 2020.

APPENDIX A: PROGRAM OVERVIEW AND TOC

Under the Africa Poultry Multiplication Initiative (APMI) portfolio, the Bill and Melinda Gates Foundation (BMGF) and the World Poultry Foundation (WPF) are partnering with the Tanzania poultry supplier Silverlands (STL) to scale-up the distribution of **low-input, Sasso chickens** to rural farming households in Tanzania. These improved chickens (termed "Sasso") are more productive and exhibit lower mortality rates than local breeds of chickens. In Tanzania, APMI's primary objectives are to:[40]

1. Increase poultry production and productivity: By increasing access to low-input Sasso breeds which have been properly brooded, fed, and vaccinated, the program expects an increase in productivity compared to the local breeds.

2. Increase rural household income: Through egg and chicken sales, APMI aims to increase household income.

3. Improve household nutrition: By increasing productivity and producing more chicken meat and eggs, the program anticipates that men, women, and children will consume more of these products in the household. APMI also aims to increase SHFs' knowledge of the importance of eggs and chicken meat consumption.

4. Empower women: By increasing access to quality poultry inputs and increasing women's engagement in the poultry value chain, APMI aims to increase women's empowerment.

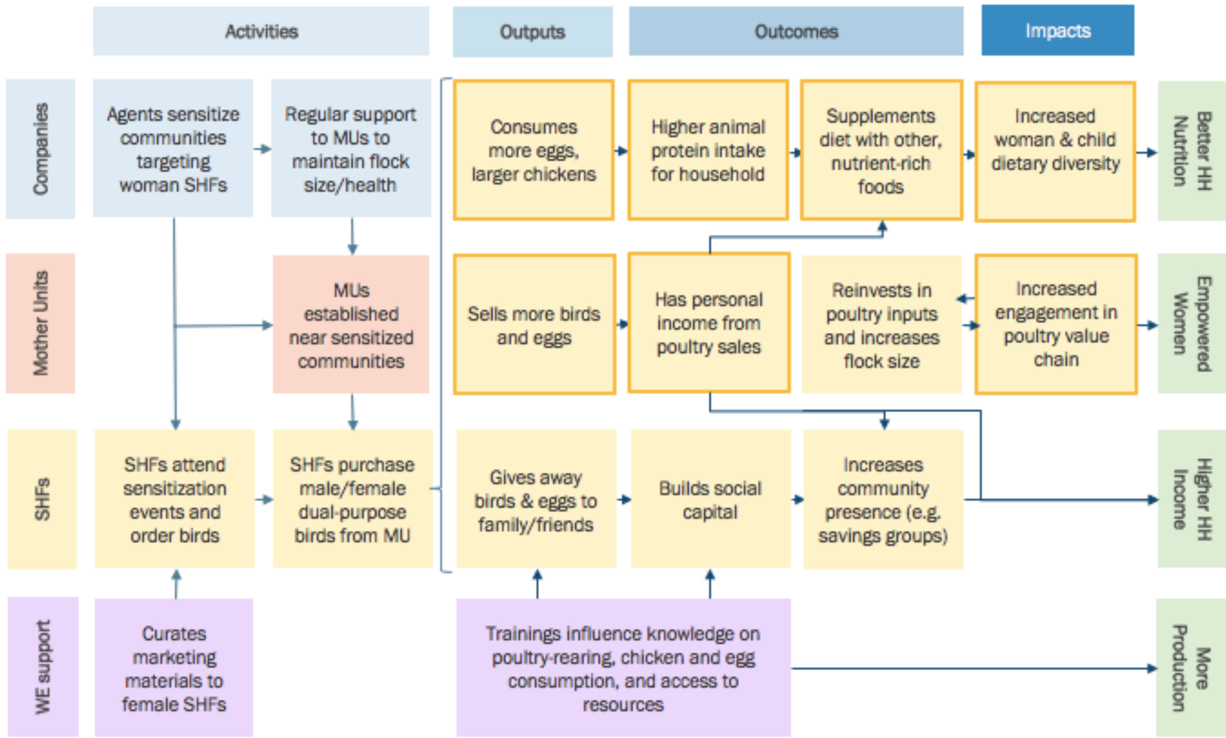
Figure 1 below maps the Theory of Change (TOC) -- the expected pathways to impact for the APMI program. Highlighted in the diagram are the roles of various stakeholders, pathways, and assumptions necessary for the program to lead to improved outcomes, as well as key indicators by which to measure program success.

The TOC begins when a company establishes Mother Units (MUs) in communities to distribute Sasso chickens to smallholder farmers (SHFs). Sasso chickens allow SHFs to sell, consume, or gift more meat and eggs than would be possible with local chickens. SHFs can use their additional income from sales to increase spending on nutrient-rich foods for themselves and their children; the extra consumption of chicken also increases the share of protein in their families' diets. As SHFs accrue benefits from chickens, they then increase the size of their flocks and multiply their outcomes. Over time, these channels lead to measurable impacts in household income, female decision-making, and women and children's nutrition.[41]

[40] To learn more about the APMI project, see http://worldpoultryfoundation.org/projects/the-african-poultry-multiplication-initiative-apmi-in-tanzania-and-nigeria/

[41] While Sasso chickens will not be exclusively sold to female SHFs, we assume based on field observations and relevant literature that the majority of buyers will be women. This will lead to particular impacts from poultry on women's nutrition and economic empowerment.

Figure 1: Draft Theory of Change of the APMI Evaluation



APPENDIX B: MATCHING APPROACH

In this appendix, we describe the various covariates and techniques we are considering for the individual-level matching process, although we will establish the final protocol after baseline data is collected and analysed. Matching will take place after we conduct a monitoring exercise to verify which SHFs have purchased Sasso.

The matching process establishes pairs of SHFs who are similar on the variables on which we matched, excluding ownership of Sasso ("treatment"). We will develop one-to-one matches between treatment and control SHFs from a combination of attribute and baseline outcome indicators. Additionally, we will aim to match treatment and control SHFs within the same region to account for environmental factors. Variables that we are considering for matching are outlined in **Table 7** below.

Attribute indicators describe households' socio-economic and demographic characteristics. Outcome indicators describe households' baseline levels on the main outcomes of the study, i.e. income from chickens, family nutrition, and women's empowerment. Given the large number of attribute variables considered, we might estimate the propensity score from a model that includes all the variables. We would then include the propensity score along with key baseline variables (such as region and outcome values) in the matching algorithm.

Table 7: Covariates in matching analysis

Attribute data	Outcome data
Region	Egg production
Gender	Egg consumption
Household size	Meat consumption
Flock composition[42]	Currently selling eggs
Poverty level (PPI Score)	Currently selling chickens
Land ownership	Expenditure on chickens (excluding Sasso expenditure)[43]
Presence of children (2-5)	Food expenditures
Presence of infants (0-2)	Expenditure on dairy
Gender of primary caretaker of poultry	Expenditure on fruit
Acres of land owned	Expenditure on vegetables
Currently plant cash crops	Expenditure on meat
Currently formally employed	Infant dietary score (if applicable)
Household income over past 12 months	Child dietary score (if applicable)
Livestock ownership (apart from chickens)	Woman dietary score (if applicable)
	Gender empowerment index[44]
	Gender Parity Index
	Women's involvement in poultry rearing

We will create a separate matched sample for child-level outcomes in order to compare treatment and control households who have children at similar ages at endline. The survey asks children's nutrition questions for households that had children aged 6 months to five years, though we will not necessarily follow up about the same child at endline that we survey at baseline.[45] As such we likely

[42] We will exclude Sasso chickens from flock composition calculations to match on SHF characteristics pre-chicken exposure. We will increase SHF income levels by the amount they purchased Sasso to account for this.
[43] We will add Sasso chicken expenditure to SHF income to account for excluding this expenditure.
[44] See Appendix F for how this will be constructed.
[45] It is likely that in some of these households we would find different children that are in the priority age range (6 months - 2 years) or no longer in the 6 month – 5 year age range at all at endline. Furthermore, some SHFs may report to have children in the relevant age group at endline, while there were no children in that age group at baseline.

will need to rematch households following the endline survey based on children ages at endline.[46] Specifically, for all households with children at endline we will match on the above matching variables plus the child's age at endline. For households with children at endline who also reported child nutrition at baseline, we will also match on child/infant diet outcomes at baseline.

To create the matched pairs of SHFs, we are considering the below matching procedures:

- <u>Propensity score matching:</u> We estimate each SHF's likelihood of purchasing Sasso chicken (i.e. propensity score) using the set of outcomes and covariates outlined above. We will explore estimating propensity scores using various models, such as logistic regressions, elastic net and random forests. We then consider only SHFs that have overlapping ranges in these scores (i.e. common support). Finally, we use nearest neighbor matching without replacement to match treatment to control SHFs on these scores.[47]

- <u>Covariate genetic matching:</u> This is an iterative approach that attempts to match as accurately as possible on all matching variables simultaneously.[48] This process allow us to determine the closeness of matches along each of the matching covariates by setting calipers, which denote how close matches need to be on that variable.[49] If one run of the matching algorithm produces imbalance on a certain input, we can adjust the caliper on that variable to ensure better balance.

- <u>Coarsened exact matching:</u> This matching approach aims to achieve balance across covariates by coarsening or binning variables by defined cutpoints. Strata of observations are defined by each unique combination of bins. Observations within these strata are then matched to each other.[50] We are able to adjust the cutpoints as needed to achieve more matches or better balance across the matching variables.

We also consider using these approaches in combination by e.g. first predicting the propensity score and then implementing a genetic matching approach that includes both covariates and the estimated propensity score.

The final matching procedure will be determined by the algorithm that generates the best balance on the matching variables across the treatment and control SHFs. To assess balance, we will compare standardized mean differences and test equality of distributions across the variables.

---

[46] As such a household may both be a match for a household in the main sample as well as the child sample.
[47] We would aim to match SHFs on these scores within regions to account for geographical factors.
[48] Diamond, Alexis and Sekhon, Jasjeet. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." The Review of Economics and Statistics. 95.3 (2013) 932-945.
[49] Calipers are measured in standard deviations. A caliper of zero will require exact matches on a variable, small calipers will require close matches, and large calipers will allow for greater latitude.
[50] If a stratum has an uneven number of control and treatment units, estimators will require weighting observations according to the size of their strata.

APPENDIX C: SURVEY QUESTIONS BY INDICATOR

Table 8: Indicators and questions

Dimension		Priority	Indicator (at unit level)	Questions (variable name)	Aggregation (at unit level)
Productivity	1.	Primary	Egg production in the last 7 days	(local_egg_7days / sasso_egg_7days) In the last 7 days, how many eggs did your local / Sasso chickens lay in total?	Egg_production = local_egg_7days + sasso_egg_7days
	2.	Secondary	Egg production (per bird)	(local_egg_7days / sasso_egg_7days) In the last 7 days, how many eggs did your local / Sasso chickens lay in total? (num_local_chickens / num_sasso) How many local chickens/sassos do you and your household own?	Eggs_per_sasso = sasso_egg_7days / num_sasso Eggs_per_local = local_egg_7days / num_local_chickens Eggs_per bird = Egg_production / (num_sasso + num_local_chickens)
	3.	Secondary	(Any) Egg production in the last 7 days	(local_egg_prod / sasso_egg_prod) Did any of your LOCAL/SASSO chickens lay eggs in the last 7 days?	Egg_prod = local_egg_prod + Sasso_egg-prod
	4.	Secondary	Bird mortality rate	(local_disease / sasso_disease) In the last 6 months, how many of your local / sasso chickens died from disease?	bird_mortality = (local_disease + sasso_disease) / (largest_flock + largest_flock_n)
	5.	Secondary	Financial expenditure on chickens in last 30 days	(feed_frequency) In the last 30 days, how many times did your household purchase chicken feed? (feed_quantity) How much feed did you buy in total (Tanzania Shillings) over the last 30 days? (eggs_transport_times) In the last 30 days, how many times did you spend money on transportation for selling eggs?	feed_expenses = feed_quantity eggs_expenses = eggs_transport_times * eggs_transport_cost chickens_expenses = chicken_transport_times * chicken_transport_cost vaccine_expenses = chicken_vaccine_costamount / 6 vet_expenses = vet_costamount / 6

				(eggs_transport_cost) On average, when you spent money on transportation for selling eggs, how much did you usually spend each time? (chicken_transport_times) In the last 30 days, how many times did you spend money on transportation for selling chickens? (chicken_transport_cost) On average, when you spent money on transportation for selling chickens, how much did you usually spend each time? (chicken_vaccine_costamount) In the last 6 months, in total, how much did you spend on vaccines, antibiotics, or other medications for your chickens? (vet_costamount) the last 6 months, in total, how much did you spend veterinary services for chickens? (other_chickencost_amount) On average, when you spent money on this expense, how much did you spend on (chickencost)?	other_expenses = other_chickencost_amount / 6 **total_chicken_expenses** = feed_expenses + egg_expenses + chickens_expenses + vaccine_expenses + vet_expenses + other_expenses
	6.	Secondary	Time expenditure on chickens in last 7 days	(time_cleaning_coop) In a typical week, how much time do you or your household spend cleaning where the chickens stay? (time_feeding_chickens) On a typical day, how much time do you or your household spend feeding the chickens?	time_coop = time_cleaning_coop time_feed = time_feeding_chickens * 7 time_source = time_sourcing_feed / 4.28 time_selling = time_selling_chickens / 4.28 time_other = time_other_chickenactivity * 7

				(time_sourcing_feed) In the last 30 days, how much time did you spend sourcing feed for your chickens? (time_selling_chickens) In the last 30 days, how much time did you spent selling your chickens? (time_other_chickenactivity) On a typical day, how much time (minutes) do you or your household spend on (chickenactivity)?	**total_time_expenses** = time_feed + time_source + time_selling + time_cleaning_coop + time_other
Income	7.	Primary	Income from poultry produce in last 30 days	(local_egg_sell/sasso_egg_sell), (local_egg_price / sasso_egg_price), (num_local_sold / num_sasso_sold), (local_price / sasso_price)	Total_poultry_income = total_egg_income + total_chicken_income / 6
	8.	Secondary	Income from egg sales in last 30 days	(local_egg_sell/sasso_egg_sell) In the last 7 days, how many local/sasso chicken eggs did you sell or plan to sell? (local_egg_price / sasso_egg_price) In the last 30 days, what is the most common price that you received for one local/sasso egg?	income_local_egg = (local_egg_sell * local_egg_price) *4.28 income_sasso_egg = (sasso_egg_sell * sasso_egg_price) *4.28 total_egg_income = income_local_egg + income_sasso_egg
	9.	Secondary	Income from chicken sales in last 6 months	(num_local_sold / num_sasso_sold) In the last 6 months, how many local/sasso chickens did your household sell? (local_price / sasso_price) On average, how much money did you receive per local/sasso chicken?	income_local_chicken = num_local_sold * local_price income_sasso_chicken = num_sasso_sold * sasso_price total_chicken_income = income_local_chicken + income_sasso_chicken
	10.	Secondary	Profitability of rearing chickens	Average profit (revenue – expenses) from chickens in the 30 days prior to endline (including, excluding own consumption), (num_home_local)	Chicken_profit_excl = total_egg_income + total_chicken_income / 6 - total_chicken_expenses Value_own_consumption = num_home_local * local_price +

					num_home_sasso * sasso_price[51] Chicken_profit_incl = Chicken_protit_excl + Value_own_consumption
	11.	Secondary	Perception of role of poultry income in overall household revenue	(poultry_income_role) How large a role do you feel your poultry income plays in your household's overall income streams?	Poultry_income_role== "Poultry contributes significantly to our income (income from chicken >25% of total household income)"
	12.	Secondary	Household sold eggs in last 30 days	(local_egg_sale_frequency / sasso_egg_sale_frequency) In the last 30 days, how many times did you sell local/sasso eggs?	sold_eggs = 1 if local_egg_sale_frequency >= 1 or sasso_egg_sale_frequency >= 1
	13.	Secondary	Household sold chickens in last 6 months	(num_local_sold / num_sasso_sold) In the last 6 months, how many local/sasso chickens did your household sell?	sold_chickens = 1 if num_local_sold >=1 or num_sasso_sold >= 1
Nutrition	14.	Primary	Child minimum dietary diversity score (24 hours)	**See Appendix E**	
	15.	Primary	Woman minimum dietary diversity (24 hours)		
	16.	Secondary	Egg consumption in last 7 days	(eggs_eaten) In the last 7 days, how many eggs did you and the other ${hh_number} members of your household consume?	N/A
	17.	Secondary	Chicken meat consumed in last 30 days	(chicken_eaten) In the last 4 WEEKS, how many chickens did you or your household consume?	N/A
	18.	Secondary	Egg consumption by children (age 6 months-	(eggs_eaten_children) How many eggs did the children under or equal to 5 in your household eat?	N/A

[51] In cases in which the respondent has not sold any local/Sasso chickens, the average selling price in the community will be used.

			5 years) in the last 7 days		
	19.	Secondary	Children's Egg Consumption	(idds/iycf_eggs) Did ${iycf/idds_name_string} consume any eggs?	N/A
	20.	Secondary	Child under 5 consumed chicken meat in last 24 hours	(Iycf_chicken) Did ${iycf_name_string} consume any poultry meat? (Idds_poultry) Did ${idds_name_string} consume any poultry?	Chicken_eaten_child = 1 if iycf_chicken = 1 or idds_poultry = 1
	21.	Secondary	HH food expenditure in last 7 days	(tubers_roots) Did the household buy any **tubers and roots** in the last 7 days? (tubers_roots_spend) How much did the household spend purchasing these foods? [**repeat** for fruit, vegetables, dairy, grains, beverages, baby food] (legumes) Did the household buy any **legumes** in the last 30 days? (legumes_spend) How much did the household spend on purchasing these foods? [**repeat** for seeds, meat, other food]	Total_food_expenditure = tubers_roots_spend + fruit_spend + vegetables_spend + dairy_spend + grains_spend + beverages_spend + baby_food_spend + [(legumes_spend + seeds_spend + meat_spend + other_food_spend)/4.28]
	22.	Secondary	Household level spending on diverse food groups	Number of food groups that the household spend a considerable amount of money on in the last 7 (30) days	
Women's Empowerment	23.	Primary	Female weighted empowerment score	**See Appendix F**	
	24.	Secondary	Household has equal empowerment scores between male and female		
	25.	Female empowerment score (3DE): Input into productive decisions domain	Female empowerment related to input into productive decisions		

	26.	Female empowerment score (3DE): Ownership of assets	Female empowerment related to ownership of assets		
	27.	Female empowerment score (3DE): Control over use of income	Female empowerment related to control over use of income		
	28.	Gender Parity Index: Input into productive decisions domain	Household has equal empowerment scores between male and female related to input into productive decisions		
	29.	Gender Parity Index: Ownership of assets:	Household has equal empowerment scores between male and female related to ownership of assets		
	30.	Gender Parity Index: Control over use of income	Household has equal empowerment scores between male and female related to control over use of income		
	31.	Secondary	Women and poultry ownership score (WPOS)	**See Appendix G**	
Child Health	32.	Secondary	Child morbidity score	(diarrhea_days) In the last seven days, how many days did your child have diarrhea? (diarrhea_blood) In the last seven days, was there blood in the stool?	Anderson index covariate weighting algorithm (see page 30)

				(diarrhea_fever) In the last seven days, at how many days did your child have fever?	

APPENDIX D: CALCULATING THE POVERTY PROBABILITY INDEX

The Poverty Probability Index (PPI) is a standardized poverty metric that comprises 10 social and financial indicators with score-based responses (Figure 2). The summation of scores calculates the probability of experiencing poverty on various national and international poverty scales. In this study, we will use the International 2011 PPP Lines as poverty benchmarks (Figure 3).

Figure 2: PPI Questionnaire[52]

Figure 3: Poverty Likelihood Look-Up Table

PPI Score	$1.90/day 2011 PPP	$3.10/day 2011 PPP
0 – 4	96.3	100.0
5 – 9	96.3	100.0
10 - 14	75.7	95.4
15 - 19	71.4	95.3
20 - 24	62.5	92.0
25 - 29	48.0	87.5
30 - 34	36.8	76.4
35 - 39	25.9	65.8
40 - 44	15.4	50.7
45 - 49	10.6	42.5
50 - 54	7.9	32.0
55 - 59	2.9	20.4
60 - 64	0.5	15.4
65 - 69	0.5	7.8
70 - 74	0.5	4.8

52 We used the same PPI questionnaire in Tanzania and Nigeria. However, the social and financial indicators used to compute poverty probability index differ slightly across both countries. As we used Nigeria indicators in the questionnaire, we cannot easily compute PPI scores using the Tanzania score card and poverty likelihood look-up table. Consequently, we are using Nigeria's scorecard and look-up table (figure 3) to approximate the probability of experiencing poverty in Tanzania. This, however, leads to results that are representative of Nigeria poverty levels. To solve this, we will apply a fixed adjustment to account for differences in poverty levels between Tanzania and Nigeria. Although this adjustment could bias our results for Individual poverty levels, we do not anticipate this process to impact our matching procedure. Further, we will compare individual survey responses from our respondents to a representative survey in Tanzania.

75 - 79	0.0	1.8
80 - 84	0.0	0.0
85 - 89	0.0	0.0
90 - 94	0.0	0.0
95 - 100	0.0	0.0

Indicators	Responses	Score
1. How many members does the household have?	A. Ten or more	0
	B. Eight or nine	5
	C. Seven	10
	D. Six	11
	E. Five	17
	F. Four	19
	G. Three	25
	H. One or two	32
2. How many separate rooms do the members of the household occupy (do not count bathrooms, toilets, storerooms, or garage)?	A. One	0
	B. Two	4
	C. Three	5
	D. Four	6
	E. Five or more	7
3. The roof of the main dwelling is predominantly made of what material?	A. Grass, clay tiles, asbestos or plastic sheets, or others	0
	B. Concrete, zinc, or iron sheets	4
4. What kind of toilet facility does the household use?	A. None, bush, pail/bucket, or other	0
	B. Uncovered pit latrine, or V.I.P. latrine	3
	C. Covered pit latrine, or toilet on water	6
	D. Flush to septic tank, or flush to sewage	15
5. Does the household own a gas cooker, stove (electric, gas table, or kerosene), or microwave?	A. No	0
	B. Yes	3
6. How many mattresses does the household own?	A. None	0
	B. One	6
	C. Two	8
	D. Three or more	10
7. Does the household own a TV set?	A. No	0
	B. Yes	8
8. How many mobile phones does the household own?	A. None	0
	B. One	2
	C. Two	5
	C. Three or more	7
9. Does the household own a motorbike or a car or other vehicle?	A. No	0
	B. Only motorbike	3
	C. Car (regardless of motorbike)	11
10. Does any member of this household practice any agricultural activity such as crop, livestock, or fish farming, or own land that is not cultivated? If so, does the household own any sprayers, wheelbarrows, or sickles?	A. Farms or has uncultivated land, but no sprayers, wheelbarrows, or sickles	0
	B. Farms or has uncultivated land, and has sprayers, wheelbarrows, or sickles	3
	C. Does not farm nor has uncultivated land	3
	Total Score:	

APPENDIX E: CALCULATING THE MINIMUM DIETARY DIVERSITY SCORE

The Minimum Dietary Diversity Score estimates the micronutrient adequacy of household members' diets.[53] In this study, we estimate dietary diversity for all women primary respondents, and for one[54] of the children in the household if available (either aged >6 months to below 2 years or aged between 2-5 years). We use a separate index for each of these three potential individuals. The score draws from 7 to 10 food groups depending on the respondent, each recorded as dichotomous variables (0 = not consumed, 1 = consumed). The summation of all food groups calculates the dietary diversity score (DDS). The **Minimum Dietary Diversity (MDD)** is a dichotomous threshold, which we calculate from the DDS as follows:

Table 9: MDD

Index and target respondent	Number of groups	Aggregation method (Stata syntax)	MDD threshold
Women's Minimum Dietary Diversity	10 Food Groups ☐ Grains, white roots and tubers, and plantains ☐ Pulses (beans, peas and lentils) ☐ Nuts and seeds ☐ Dairy ☐ Meat, poultry and fish ☐ Eggs ☐ Dark green leafy vegetables ☐ Other vitamin A-rich fruits and vegetables ☐ Other vegetables ☐ Other fruits	Gen wdd_grains_roots_tubers_0 = wdd_grains + wdd_tubers Gen wdd_grains_roots_tubers = 0 Replace wdd_grains_roots_tubers = 1 if wdd_grain_roots_tubers_0 >=1 Gen wdd_meat_poultry_fish_0 = wdd_organmeat + wdd_meat + wdd_fish + wwd_poultry Gen wdd_meat_poultry_fish = 0 Replace wdd_meat_poultry_fish = 1 if wdd_meat_poultry_fish_0 >=1 Gen w_dds = wdd_grains_roots_tubers + wdd_meat_poultry_fish + wdd_pulses + wdd_nuts + wdd_dairy + wdd_eggs + wdd_veg + wdd_vita + wdd_otherveg + wdd_otherfruit	DDS >= 5 (out of 10) Gen w_mdd = 1 if w_dds >=5

[53] http://www.fao.org/3/a-i1983e.pdf
[54] In households where there is more than one child each age range, we ask about the nutrition of the child closest to two years.

Individual Minimum Dietary Diversity for children 2-5 years	9 Food Groups ☐ Grains, white roots and tubers, and plantains ☐ Pulses (beans, peas and lentils), nuts, and seeds ☐ Dairy ☐ Meat, poultry and fish ☐ Eggs ☐ Dark green leafy vegetables ☐ Other vitamin A-rich fruits and vegetables ☐ Other vegetables ☐ Other fruits	Gen idds_grains_roots_tubers_0 = idds_grain + idds_tubers Gen idds_grains_roots_tubers = 0 Replace idds_grains_roots_tubers = 1 if idds_grain_roots_tubers_0 >=1 Gen idds_meat_poultry_fish_0 = idds_meat + idds_fish +idds_poultry Gen idds_meat_poultry_fish = 0 Replace idds_meat_poultry_fish = 1 if idds_meat_poultry_fish_0 >=1 Gen i_dds = idds_grains_roots_tubers + idds_meat_poultry_fish + idds_pulses + idds_milk + idds_eggs + idds_vegetables + idds_vitamina_fruit + idds_othervegetables + idds_otherfruit	DDS >= 5 (out of 9) Gen i_mdd = 1 if i_dds >=5
Infant and Young Child Minimum Dietary Diversity for children 6-24 months	7 Food Group ☐ 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 ☐ Other fruits and vegetables	Gen iycf_grains_roots_tubers_0 = iycf_grain + iycf_tubers Gen iycf_grains_roots_tubers = 0 Replace iycf_grains_roots_tubers = 1 if iycf_grain_roots_tubers_0 >=1 Gen iycf_meat_poultry_fish_0 = iycf_meat + iycf_fish +iycf_chicken Gen iycf_meat_poultry_fish =0 Replace iycf_meat_poultry_fish = 1 if idds_meat_poultry_fish_0 >=1 Gen iycf_other_0 = iycf_fruit + iycf_vegetables Gen iycf_other = 0 Replace iycf_other = 1 if iycf_other_0 >=1 Gen iycf_dairy_0 = iycf_milk + iycf_yoghurt Gen iycf_dairy = 0 Replace iycf_dairy = 1 if iycf_dairy_0 >=1 Gen iycf_dds = iycf_grains_roots_tubers + iycf_meat_poultry_fish + iycf_pulses + iycf_eggs + iycf_dairy + iycf_vitaminA + iycf_other	DDS >= 4 (out of 7) Gen iycf_mdd = 1 if iycf_dds >=4 Aggregate into one child indicator: *Gen child_mdd = 1 if iycf_mdd = 1 or i_mdd = 1

APPENDIX F: CALCULATING THE MODIFIED A-WEAI

IDinsight has adapted the Abbreviated Women's Empowerment in Agriculture Index (A-WEAI) as a measure of women's empowerment in this research. The original A-WEAI involves 5 dimensions and 6 sub-indicators, which comprise the weighted Five Dimensions of Empowerment (5DE) composite score and the Gender Parity Index (GPI).

We emphasize three dimensions of the 5DE that are relevant to context of the APMI study:

1. Input in productive decisions
2. Ownership of assets
3. Control over use of income

We utilize these three indicators to generate (a) a modified empowerment score which we term the **3DE,** and (b) a Gender Parity Score.

Empowerment scores

The 3DE is the weighted average of the three indicators listed above, measured for each of the male and female primary respondents in a household. For each indicator, a binary value is assigned according to the respondent's achievement in that field. A score of 1 indicates adequate empowerment. A score of 0 indicates inadequate empowerment. **The weights provided in the 3DE are proportional to those in the 5DE, adjusted for the reduction in total dimensions**.[55]

The weighted average of the three indicators produces a 3DE score between 0 and 1, where 1 indicates fully empowered.[56] We can analyze this in the aggregate as the mean score (M_{3DE}) across each of the control and treatment groups, by gender. This procedure is described in more detail in **Table 10** below.

Respondents are considered "empowered" if their respective 3DE score is larger or equal than 0.75.[57]

Gender parity scores

Gender parity is measured in binary terms at the household-level. GPI reflects the percentage of women who are equally or more empowered as the men in their households. In households where women have the same or higher 3DE score (between 0 and 1), we record a GPI score of 1, and 0 if not.[58] GPI scores are then averaged across the sample to produce:

$\mathbf{M_{GPI}}$: the percentage of households that achieve Gender Parity.
1- $\mathbf{M_{GPI}}$: the percentage of households that have not achieved Gender Parity

---

[55] The original weights were 1/5, 2/15, and 1/5 for each of the three dimensions above. These were normalized by dividing each value by the sum of weights (8/15), producing 3/8, 2/8, and 3/8 respectively. Alternatively, we could employ an Anderson index that generates weights algorithmically based on how much each variable explains overall variance in the outcome. At this time, we chose to maintain the standardized, proportional weights, in order to preserve potential for some comparability to other A-WEAI studies.

[56] Missing entries in the 3DE components will be counted as inadequate empowerment in the 3DE score aggregation. For example, respondents who do not report to participate in any productive decision are considered disempowered in the production dimension of the 3DE. This is in contrast to IFPRI's recommended approach of dropping any observation with incomplete information. We decided against this approach to mitigate the risk of reduced sample size for a main outcome variable.

[57] Table 10 illustrates that this is equivalent to saying that respondents are empowered if they are empowered at least in the income and production domains.

[58] Similar parity scores could be generated for the domain-specific empowerment indices.

Table 10: Modified A-WEAI

Dimension	Empowerment indicator	Survey Questions	Achievement cut-off [59] (indicates achievement in aspect/field)	Aggregation Method
Production	1. Input in productive decisions	How much input did you have in making decisions about: [food crop farming, cash crop farming, livestock raising?] To what extent do you feel you can make your own personal decisions regarding these aspects of household life if you want(ed) to: [food crop farming, cash crop farming, livestock raising]?	**If individual participates**, (a) must have at least 'input into some decisions' (value=2); and (b) individual makes the decision themselves (value = 1) OR individual feels at least 'to a medium extent' (value = 2) they could make their own personal decisions	Achievement in **two** decisions = 1 Weight: 3/8
Resources	2. Ownership of assets	Does anyone in your household currently have any [ITEM]? Do you own any of the [ITEM]? [Farm equip (non-mech); Farm equip (mechanized; Nonfarm business equipment; Large durables; Small durables; Cell phone; Non-ag land (any); Transport]	**If household owns the asset,** the individual owns most of that asset alone or jointly (values = 1,2)	Achievement in **one**, unless that one asset is (a) a small consumer durable or (b) a non-mechanized farm equipment = 1 Weight = 2/8
Income	3. Control over use of income	How much input did you have in decisions on the use of income generated from: [Food crop, Cash crop, Livestock, Non-farm activities, Wage& salary] To what extent do you feel you can make your own personal decisions regarding these aspects of household life if you want(ed) to: [Non-farm economic activities, your own wage or salary employment, health expenditure, education expenditure, minor household expenditure]	**If individual participates**, (a) they have at least 'input into some decisions' about income generated (value=2); AND (b) they feel they can make their own person decisions at least 'to a small extent' (value=2).	Achievement in **one**, if not only minor household expenditure = 1 Weight: 3/8

[59] In this table, achievement cut-off and aggregation methods follow IFPRI guidance as per the original A-WEAI in https://www.ifpri.org/sites/default/files/a-weai_instructional_guide_final.pdf

APPENDIX G: CALCULATING THE WOMEN AND POULTRY OWNERSHIP SCORE

IDinsight also explores women's empowerment specifically in the context of poultry rearing. IDinsight's constructed Women and Poultry Ownership Score (WPO) is a composite score of three sub-indicators. Each sub-indicator is assigned a value of 0-1, depending on a respondent's achievement in that dimension, as described in **Table 11** below.

The composite score (WPO) is a standardized index of the three sub-indicators. Thus, the WPO score has a mean of 0 and a standard deviation of 1, where higher scores indicate a respondent is more **empowered.**[60]

$\mathbf{M}_{\mathbf{WPO}}$ is the aggregated measure of WPO, as the overall percentage of female respondents who achieve empowerment.

Table 11: WPO Score

Indicator Name	Survey Questions	Aggregation Method	Achievement cut-off
Women's involvement in poultry rearing	1. When decisions are made regarding extension support (e.g. Attending trainings, seeking advice, adopting new techniques), who is it that normally takes the decision? 2. When decisions are made regarding the place where the chickens stay (e.g. Buy, repair, expand a coop), who is it that normally takes the decision? 3. When decisions are made regarding packaged feed (e.g. Purchase, feed chickens), who is it that normally takes the decision?	Achievement in **one** decision = 1	Female respondent must take the decision, but not necessarily alone (value=1)
Women's ownership of poultry	1. Who in the household decided to sell the local/Sasso chickens? 2. How much input did you have in decisions regarding buying chickens? 3. How much input did you have in decisions regarding selling eggs? 4. How much input did you have in decisions regarding eating eggs or chicken in the past 6 months?	Achievement in **two** decisions = 1	Female respondent must take the decision to sell, but not necessarily alone (value=1); individual must have at least 'some input' in the decision (value=2)

[60] The Anderson index is an algorithm that generates a standardized index based on how much each variable explains overall variance in the outcome. This is an optimal approach for non-standardized indices where we have no priors on the relative significance of its constituent parts. Details can be found at: https://are.berkeley.edu/~mlanderson/pdf/Anderson%20Preschool.pdf

Women's use of income from poultry	1. How much input did you have in decisions on the use of income generated from the Sasso chicken sales? 2. How much input did you have in decisions on the use of income generated from the local chicken sales? 3. When the eggs were sold, to what degree did you decide how the income would be used?	Achievement in **one** decision = 1	Female respondent must have at least 'some input' in the decision (value=2)