APMI NIGERIA PRE-ANALYSIS PLAN

May 2020

IDinsight
DATA. DECISIONS. DEVELOPMENT.

Authors

Will Slotznick: will.slotznick@IDinsight.org
Maureen Stickel: maureen.stickel@IDinsight.org
Torben Fischer: torben.fischer@IDinsight.org
Christy Lazicky: christy.lazicky@IDinsight.org
Penny Davis: penny.davis@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.

TABLE OF CONTENTS

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 **Nigeria.**

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, then center 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**.

EVALUATION DESIGN

RESEARCH QUESTIONS

We explore the causal impact of the APMI program on SHFs, with a focus on flock performance, income, nutrition, and women's empowerment. **Table 1** outlines the primary and secondary 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 Noiler chickens affect poultry related outputs (eggs and meat)?	How do Noiler chicken breeds compare to local chicken varieties on measures of mortality and resources required to care for them (including time)?
Income	How does owning Noiler 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 Noiler chicken varieties?	To what extent are women's and children's diets more diverse as a result of owning dual purpose birds?
Women's Empowerment	Does owning 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 improved (Noiler) 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 Noiler chickens.[3] At endline, we will compare outcomes of treatment and control SHFs to estimate the causal effect of owning Noilers.

[1] The study aims to identify the effect of owning (some) Noiler chickens in comparison to not owning any Noiler 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 Noiler 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.**

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

STUDY SITE AND COMMUNITY SELECTION

We identified five states for the focus of this research. These are Kano and Katsina in Northern Nigeria, plus Ondo, Ekiti, and Kwara in South West Nigeria.

We selected particular study sites in coordination with Amo Farm. Our iterative selection process with Amo Farm produced an initial list of viable communities. We then randomly allocated treatment status to communities stratified by state and LGA, designating about half the communities as treatment (receive access to Noiler) and half as control (do not receive access to Noiler). We sought nearly equal proportions of sites across the five states, dependent on the supply of viable sites.

The resulting designations were as follow:

Table 2: State-level designations

State	Treatment	Control	Replacements
Kano	9	9	5
Katsina	8	8	8
Ondo	8	8	5
Ekiti	9	9	5
Kwara	9	10	1
Totals	43	44	24

During the baseline data collection process, we needed to replace 24 communities from the original list for two primary reasons:

- **Insecurity** – Due to reported incidents of increased levels of attacks and kidnappings, communities were replaced in order to prioritize safety for the field teams.
- **Low likelihood of Mother Unit set up** – As the selection of communities was in coordination with AFSH agents, the agents were able to determine the likelihood of setting up MUs in communities. Communities that turned out to be highly unlikely to set up an MU after initial engagements were replaced with communities in the same state (and LGA whenever possible) that exhibited similar characteristics, but were deemed to have a higher likelihood for a Mother Unit to be set up by AFSH agents.

To retain comparability between treatment and control communities, we chose replacement communities that were similar in terms of characteristics such as, but not limited to: population size, access to health care and access to financial services.[4] We also ensured that these replacements were in the same state and local government area as the communities they were compared to. Replacing communities was best undertaken in coordination with AFSH agents and the local survey firm as they were the most familiar with the surrounding communities. As such, the agents and field managers were able to provide valuable insights on communities that were similar in size and demographics to communities that needed to be replaced. After coordinating with the agents, IDinsight sent mobilizers to potential sites to scope the community to determine if they were viable replacements. This strategy allowed us to find similar communities to replace the older communities in an accelerated time, mitigating potential delays in data collection.

SAMPLE SELECTION

The population of interest comprises *buyers (treatment)* and *would-be buyers (control)* of Noiler chickens.

Treatment sample

In the majority of cases, the treatment SHF represents an **actual buyer** of Noiler.

IDinsight will randomly sample 20 treatment SHFs (plus 5 replacements) using MU owners' sales records.[5] Identifying SHFs and their location details will be a coordinated process between IDinsight, Amo Farm, and our locally-based survey partner.

We aim to survey these treatment SHFs within **10 weeks**[6] from their date of purchase, to minimize any Noiler-related costs and benefits felt by SHF owners at baseline.

In the case that Amo Farm does not establish MUs by the necessary time of data collection in a given region[7], we will sample SHFs **who are likely** to purchase Noilers[8]. In this scenario, we will attend official Amo Farm outreach events in communities and deliver a 3-4 minute eligibility survey to one participant per household. The eligibility survey gathers basic identifying information, and assesses general willingness to pay (WTP) for a new breed of poultry[9]. Of those SHFs with adequate WTP, we will

---

[4] As we will match across SHFs, balance across treatment and control communities is not strictly necessary. However, as having comparable communities increases the likelihood of finding good SHF matches, we aimed to find control communities that look similar to the replacement treatment communities.

[5] If the MU reports to have sold to less than 20 customers, we will attempt to survey all customers.

[6] SHFs purchase Noilers at 5 weeks of age. At 15 weeks of age, Noiler cocks reach maturity to sell for consumption. Noiler hens reach laying-age at 22 weeks. Some SHFs may choose to sell Noiler cocks slightly before 15 weeks, so we optimally aim to survey communities at 7 weeks after purchase. All expenses accrued from Noiler ownership in the interim (before surveying) will be backed-out in the analysis to understand true baseline levels of poultry expenditures.

[7] Based on current expansion data, we anticipate this will be the case for about 32 treatment communities.

[8] We also refer to this approach as 'prospective sampling.'

[9] The survey includes other WTP questions so as to not link the research with poultry specifically.

randomly select 28 (plus 5 replacements)[10] to participate in the main household survey (and join the baseline sample).

Control sample

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

We approach this in control communities by mimicking the process by which Amo Farm conducts outreach for Noilers. However, we do not include content that would sway participants' interest in poultry (such that they would seek out the Noiler), or affect their baseline outcomes (e.g. investing more in nutrition).

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.[11] At community events, we will deliver the same eligibility survey as above.

Respondents with adequate WTP scores are inducted into the short-list of eligible respondents as a part of our sampling frame. Of those, 30 (plus 5 replacements) will be randomly selected for inclusion in the main household survey.[12]

Altogether, the baseline sample is planned to comprise approximately 2,430 SHFs: 900 treatment SHFs and 1530 control SHFs. Ahead of endline, we will conduct a monitoring exercise to verify treatment SHFs who have actually purchased Noiler birds as well as control SHFs who did not have access to Noilers 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 treatments and 800 controls. This is also outlined in the table below:

Table 3: sample sizes in stages

	Treatment SHFs	Control SHFs	Total SHFs
1. Baseline Sample Size (with buffers)	900	1530	2430

[10] The additional number of SHFs provides buffer in case some do not ultimately purchase Noilers. If less than this number of eligible SHFs is identified in the listing exercise, all eligible SHFs will be sampled.
[11] 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.
[12] Sampling is conducted in two stages. In the first stage, we classify SHFs based on their WTP for a 5-week old Noiler 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.

2. Eligible SHFs after monitoring[13] (minimum)	700	1400	2100
3. Matched SHFs (minimum)	600	600	1200
4. Endline Sample Size (with buffers[14]) *Includes IYCF MDD sample size*	600 *(300)*	800 *(300)*	1400 *(600)*

Given the planned replacement of some of the original treatment communities described above, the numbers in Table 3 are subject to change.[15]

Table 4 provides the minimum detectable effect sizes (MDES)[16] 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.[17] 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).

Table 4: 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)

13 We anticipate the sample size to decrease slightly for both treatment and control SHFs as some treatment SHFs will likely not have purchased Noilers and some control SHFs will likely have received access to the Noiler.

14 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).

15 Update, April 2020: At baseline, we interviewed 2248 SHFs, 973 in treatment and 1273 In control communities.

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

17 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).

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. A portion of the survey (supplemental survey) is also administered to an **adult head of the opposite gender**.

- 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[18], and (3) monitor the sell prices of mature Noiler birds over time.

Baseline data collection will be conducted between June and August 2019. Monitoring data will be collected with Noiler communities throughout 2020. We anticipate endline data collection will be conducted in 2021.

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

ANALYTICAL FRAMEWORK

OUTCOMES

Table 5 below describes the primary and secondary indicators used to measure the influence of Noilers for SHFs. 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's 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 report to Gates. Statistically significant findings in secondary outcomes along with absence of statistical significance in primary outcomes will not drive main recommendations to the broader stakeholder group.

Table 5: Primary and secondary outcome indicators used in analysis

Dimension		Priority	Outcome	Indicator (at unit level)	Unit of analysis	Type[19]
Productivity / performance[20]	1.	Primary	Egg production	Number of eggs produced in the last 7 days	Household	Continuous
	2.	Secondary	Egg production (per bird)	Number of eggs produced in last 7 days per bird	Household	Continuous
	3.	Secondary	Egg production	Any Eggs produced in the last 7 days	Household	Binary
	4.	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
	5.	Secondary	Financial expenditure on birds	Financial expenditure on chickens in last 30 days	Household	Continuous

[19] 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 Noilers in the planting season); index to a range-specified value (oftentimes aggregated from numerous sub-indicators).

[20] Indicators relating to poultry performance, production, and income will be assessed for both local and Noiler varieties (if owned by household).

	6.	Secondary	Time expenditure on birds	Time spent rearing chickens in last 7 days	Household	Continuous
Income from poultry	7.	Primary	Income from poultry produce	Income from egg and chicken sales in last 30 days	Household	Continuous
	8.	Secondary	Income from eggs	Income from egg sales in last 30 days	Household	Continuous
	9.	Secondary	Income from chicken	Income from chicken sales in last 6 months	Household	Continuous
	10.	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
	11.	Secondary	Proportion of income from chicken	Perception of role of poultry income in overall household revenue is greater than 25%	Household	Binary
	12.	Secondary	Egg seller	Households who sold eggs in last 30 days	Household	Binary
	13.	Secondary	Chicken seller	Households who sold chickens in last 6 months	Household	Binary
Nutrition	14.	Primary	Child meets minimum diet	Child minimum dietary diversity (24 hour recall)	Individual	Binary[21]
	15.	Primary	Woman meets minimum diet	Woman minimum dietary diversity (24 hour recall)	Individual	Binary[22]
	16.	Secondary	Egg consumption	Number of eggs consumed by the household in the last 7 days	Household	Continuous
	17.	Secondary	Chicken consumption	Number of chickens consumed by the household in last 30 days	Household	Continuous
	18.	Secondary	Children's egg consumption	Number of eggs consumed by children aged under 5 in the last 7 days	Household	Continuous
	19.	Secondary	Children's egg consumption	Child under 5 consumed eggs in the last 24 hours	Individual	Binary
	20.	Secondary	Children's Chicken Meat consumption	Child under 5 consumed chicken meat in last 24 hours	Individual	Binary
	21.	Secondary	Total food expenditures	HH food expenditure in last 7 days	Household	Continuous

[21] See Appendix E. Only calculated if children in the relevant age group are present in the household.
[22] See Appendix E. Only calculated if an adult woman is present in the household.

	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[23]	Household	Count
Women's empowerment[24]	23.	Primary	Female empowerment score (3DE)	Female weighted empowerment score	Individual	Index, scale [0,1]
	24.	Secondary	Gender Parity Index	Household has equal empowerment scores between male and female	Individual	Index, binary
	25.	Secondary	Female empowerment score (3DE): Input into productive decisions domain	Female empowerment related to input into productive decisions	Individual	Index, binary
	26.	Secondary	Female empowerment score (3DE): Ownership of assets	Female empowerment related to ownership of assets	Individual	Index, binary
	27.	Secondary	Female empowerment score (3DE): Control over use of income	Female empowerment related to control over use of income	Individual	Index, binary
	28.	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
	29.	Secondary	Gender Parity Index: Ownership of assets:	Household has equal empowerment scores between male and female related to ownership of assets	Individual	Index, binary
	30.	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

[23] 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.
[24] Indicators adapted from the A-WEAI index. See Appendix C and D for more calculation details.

	31.	Secondary	Women's ownership in poultry-rearing	Women and poultry ownership score (WPOS)	Individual	Index, scale [0,1]
Child Health	25.	Secondary	Childhood morbidity	Child morbidity score	Individual	Index, scale [0,1]

ANALYTICAL MODEL

We will estimate the impact of Noiler chicken ownership on the above outcomes using a difference-in-difference specification. This is our preferred specification as it allows us to difference out time invariant unobservable traits of SHFs that might be correlated with purchasing a Noiler, thus reducing potential bias.[25] Given that 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 Noiler and outcomes, this feature of difference-in-difference is advantageous. 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.[26]

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[27].

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[28]
- $\varepsilon_{s,c}$ is the error term for SHF *s* in community *c,* clustered at the community-level

---

[25] 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

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

[27] 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 Noilers, 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.

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

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.[29] For these households, baseline outcomes would be missing, invalidating the Differences-in-differences 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.

SUB-GROUP ANALYSIS

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

Table 6: Sub-groups used in analysis

	Sub-group definition	**Justification**
1.	Geographic location of household (Southwest states versus North states)	Systematic socio-cultural feature differences between the North and the South that may influence poultry rearing.
2.	Poultry owner's educational attainment [above/below secondary]	Educational level may be correlated with poultry rearing knowledge
3.	PPI scores (Quintiles of PPI score)	Wealth and stability may be correlated with productivity achievements with flock
4.	Female versus male primary care-taker of poultry	Gender of primary poultry care taker 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 Noiler within 6 months of endline	Consistent renewal of flocks may affect benefits seen from poultry

If a given sub-group definition in Table 6 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.

[29] 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).

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. While there might be selection into sub-groups, we still believe that stakeholders are interest in understanding how program effects might differ for these sub-groups.

All sub-group analyses are considered exploratory.

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 Noiler chickens. If control SHFs purchase Noiler chickens, they are likely to experience the same effects as treatment SHFs, leading us to underestimate the program's effects.

In order to mitigate this likelihood, we will closely monitor MU registration to track which villages receive access to Noilers ahead of the endline survey. If fewer than 15% of respondents in a control community report having purchased Noilers, we will drop the contaminated respondents from the sample (up to four respondents). If greater than 15% of respondents report having purchased Noilers, 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 Noilers 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. This allows 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 Noilers.[30]

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).[31] This adjustment treats hypothesis tests of similar outcomes as correlated rather than independent.[32] Conventional, uncorrected, p-values will also be reported. Table 7 below summarizes

[30] 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.
[31] Westfall, P., and S. Young. 1993. Resampling-based Multiple Testing: Examples and Methods for p-value Adjustment. John Wiley & Sons, Inc.
[32] 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.

the primary outcomes for which we will adjust standard errors, and indicates the number of hypotheses being tested with each “family” of outcomes.[33]

Table 7: Outcome families for multiple hypothesis correction

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

NEXT STEPS

We welcome stakeholder feedback on this document and will aim to finalize it by end of August. Once the PAP is finalized, we will register it with Registry for International Development Impact Evaluations (RIDIE).

Finally, we will share a baseline report with stakeholders by the end of the year.

[33] 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 sampled into the proposed FWER correction.

APPENDIX A: PROGRAM OVERVIEW AND TOC

Under the APMI portfolio, the Bill and Melinda Gates Foundation (BMGF) and the World Poultry Foundation (WPF) are partnering with the Nigerian poultry supplier Amo Farm Sieberer Hatchery (AFSH) to scale-up the distribution of **low-input, Noiler chickens** to rural farming households in Nigeria. These improved chickens (termed "Noiler") are more productive and exhibit lower mortality rates than local breeds of chickens. In Nigeria, APMI's primary objectives are to:[34]

1. Increase poultry production and productivity: By increasing access to low-input Noiler 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.

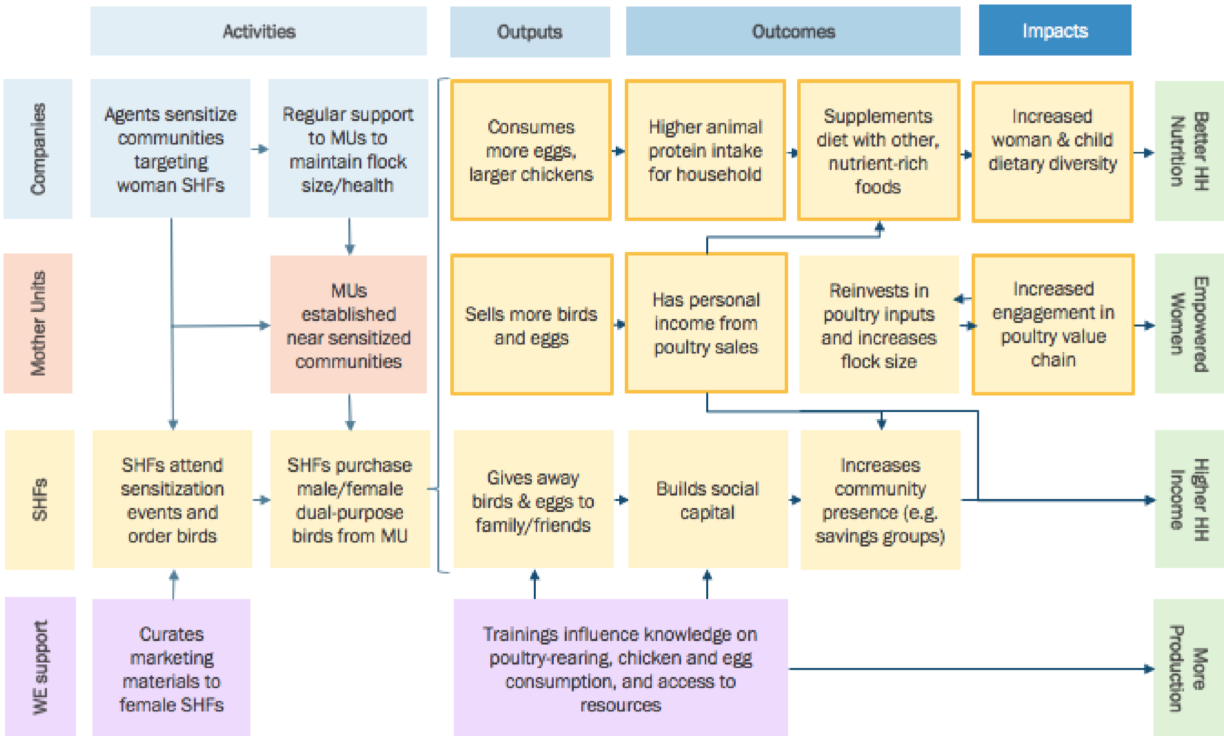
Figure 1 below maps the Theory of Change (TOC) -- the expected pathways to impact for the APMI program. Highlighted in the diagram are the roles of various stakeholders, pathways, and assumptions necessary for the program to lead to improved outcomes, as well as key indicators by which to measure program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

The TOC begins when a company establishes Mother Units (MUs) in communities to distribute Noiler chickens to smallholder farmers (SHFs). Noiler chickens allow SHFs to sell, consume, or gift more meat and eggs than would be possible with local

[34] http://worldpoultryfoundation.org/projects/the-african-poultry-multiplication-initiative-apmi-in-tanzania-and-nigeria/

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.[35]

Figure 1: Draft Theory of Change of the APMI Evaluation



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

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 conducting a monitoring exercise to verify which SHFs have purchased Noilers.

The matching process establishes pairs of SHFs who are similar on the variables on which we matched, excluding ownership of Noilers ("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 state 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, 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 state and outcome values) in the matching algorithm.

Table 7: Covariates in matching analysis

Attribute data	Outcome data
State (North/Southwest)	Egg production
Gender	Egg consumption
Religion	Meat consumption
Language	Currently selling eggs
Household size	Currently selling chickens
Flock composition[36]	Expenditure on chickens (excluding Noiler expenditure)[37]
Poverty level (PPI Score)	Food expenditures
Land ownership	Expenditure on dairy
Presence of children (2-5)	Expenditure on fruit

[36] We will exclude Noilers from flock composition calculations to match on SHF characteristics pre chicken exposure. We will increase SHFs income levels by the amount they purchased Noilers for to account for this.

[37] We will add Noiler chicken expenditure to SHF income to account for excluding this expenditure.

Presence of infants (0-2)	Expenditure on vegetables
Gender of primary care-taker of poultry	Expenditure on meat
Acres of land owned	Infant dietary score (if applicable)
Currently plant cash crops	Child dietary score (if applicable)
Currently formally employed	Woman dietary score (if applicable)
Household income over past 12 months	Gender empowerment index[38]
Livestock ownership (apart from chickens)	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.[39] As such we likely will need to rematch households following the endline survey based on children ages at endline.[40] 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 SHF, we are considering the below matching procedures:

- Propensity score matching: We estimate each SHFs likelihood of purchasing a Noiler (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 range in these scores (i.e. common support). Finally, we use nearest neighbor matching without replacement to match treatment to control SHFs on these scores.[41]

- Covariate genetic matching: This is an iterative approach that attempts to match as accurately as possible on all matching variables simultaneously.[42] 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

[38] See Appendix F for how this will be constructed.
[39] It's 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.
[40] As such a household may both be a match for a household in the main sample as well as the child sample.
[41] We would aim to match SHFs on these scores within states to account for geographical factors.
[42] 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.

to be on that variable.[43] 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.

- Coarsened exact matching: This matching approach aims to achieve balance across covariates by coarsening or binning variables by defined cutpoints. Stratums of observations are defined by each unique combination of bins. Observations within these strata are then matched to each other.[44] 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.

[43] 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.

[44] 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 9: 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 / noiler_egg_7days) In the last 7 days, how many eggs did your local / Noiler chickens lay in total?	Egg_production = local_egg_7days + noiler_egg_7days
	2.	Secondary	Egg production (per bird)	(local_egg_7days / noiler_egg_7days) In the last 7 days, how many eggs did your local / Noiler chickens lay in total? (num_local_chickens / num_noiler) How many local chickens/Noilers do you and your household own?	Eggs_per_noiler = Noiler_egg_7days / num_noiler Eggs_per_local = local_egg_7days / num_local_chickens Eggs_per bird = Egg_production / (num_noiler + num_local_chickens)
	3.	Secondary	(Any) Egg production in the last 7 days	(local_egg_prod / noiler_egg_prod) Did any of your LOCAL/NOILER chickens lay eggs in the last 7 days?	Egg_prod = local_egg_prod + noiler_egg-prod
	4.	Secondary	Bird mortality rate	(local_disease / noiler_disease) In the last 6 months, how many of your local / Noiler chickens died from disease?	bird_mortality = (local_disease + noiler_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 (Naira) 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? (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?	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 other_expenses = other_chickencost_amount / 6

				(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)?	**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_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)?	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 **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/noiler_egg_sell), (local_egg_price / noiler_egg_price), (num_local_sold / num_noiler_sold), (local_price / noiler_price)	Total_poultry_income = total_egg_income + total_chicken_income / 6

	8.	Secondary	Income from egg sales in last 30 days	(local_egg_sell/noiler_egg_sell) In the last 7 days, how many local/Noiler chicken eggs did you sell or plan to sell? (local_egg_price / noiler_egg_price) In the last 30 days, what is the most common price that you received for one local/noiler egg?	income_local_egg = (local_egg_sell * local_egg_price) *4.28 income_noiler_egg = (noiler_egg_sell * noiler_egg_price) *4.28 total_egg_income = income_local_egg + income_noiler_egg
	9.	Secondary	Income from chicken sales in last 6 months	(num_local_sold / num_noiler_sold) In the last 6 months, how many local/noiler chickens did your household sell? (local_price / noiler_price) On average, how much money did you receive per local/noiler chicken?	income_local_chicken = num_local_sold * local_price income_noiler_chicken = num_noiler_sold * noiler_price total_chicken_income = income_local_chicken + income_noiler_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_protit_excl = total_egg_income + total_chicken_income / 6 - total_chicken_expenses Value_own_consumption = num_home_local * local_price + num_home_noiler * noiler_price[45] 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 pool (income from >25% of total household income)”
	12.	Secondary	Household sold eggs in last 30 days	(local_egg_sale_frequency / noiler_egg_sale_frequency) In the last 30 days, how many times did you sell local/noiler eggs?	sold_eggs = 1 if local_egg_sale_frequency >= 1 or noiler_egg_sale_frequency >= 1
	13.	Secondary	Household sold chickens in last 6 months	(num_local_sold / num_noiler_sold) In the last 6 months, how many local/noiler chickens did your household sell?	sold_chickens = 1 if num_local_sold >=1 or num_noiler_sold >= 1
	14.	Primary	Child minimum dietary diversity score (24 hours)	**See Appendix E**	

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

Nutrition	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-5 years) in the last 7 days	(eggs_eaten_children) How many eggs did the children under or equal to 5 in your household eat?	N/A
	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	Female empowerment related to input into productive decisions		

		into productive decisions domain			
	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? (diarrhea_fever) In the last seven days, at how many days did your child have fever?	Anderson index covariate weighting algorithm (see page 30)

APPENDIX D: CALCULATING THE POVERTY PROBABILITY INDEX

The Poverty Probability Index (PPI) is a standardized poverty metric that converts to poverty likelihoods[46].

The PPI 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 Nigeria Questionnaire

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:	

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
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

46 https://www.povertyindex.org/country/nigeria

APPENDIX E: CALCULATING THE MINIMUM DIETARY DIVERSITY SCORE

The Minimum Dietary Diversity Score estimates the micronutrient adequacy of household members' diets.[47] In this study, we estimate dietary diversity for all women primary respondents, and for **one** of the children in the household if available (either aged >6 months to below 2 year 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 10: 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

[47] http://www.fao.org/3/a-i1983e.pdf

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**.[48]

The weighted average of the three indicators produces a 3DE score between 0 and 1, where 1 indicates fully empowered.[49] 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 11** below.

Respondents are considered "empowered" if their respective 3DE score is larger or equal than 0.75.[50]

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

[48] 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.
[49] 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.
[50] 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.

(between 0 and 1), we record a GPI score of 1, and 0 if not.[51] GPI scores are then averaged across the sample to produce:

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

51 Similar parity scores could be generated for the domain-specific empowerment indices.

Table 11: Modified A-WEAI

Dimension	Empowerment indicator	Survey Questions	Achievement cut-off [52] (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

[52] 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 12** below.

The composite score (WPO) is 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.**[53]

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

Table 12: 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/Noiler 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)
Women's use of income from poultry	1. How much input did you have in decisions on the use of income generated from the Noiler 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)

[53] 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