

Turning Inward and/or Outward: Which socioemotional skills matter the most for women in agribusiness?*

Smita Das[†]

Clara Delavallade[‡]

Ayodele Fashogbon[§]

Sreelakshmi Papineni[¶]

September 2023

Abstract

Socioemotional skills programs are increasing in popularity to promote economic empowerment. Are there certain types of socioemotional skills that better spur productivity and economic growth, and do they differ by gender and social context? To provide evidence on these questions we examine the impact of a socioemotional skills intervention on agribusiness owners in a large-scale government program in Nigeria. The intervention covers seven *intrapersonal skills* and seven *interpersonal skills* including emotional regulation, self-efficacy, personal initiative, empathy, relatedness, and negotiation, among others. Using randomized assignment we test whether the economic returns to *interpersonal skills* are higher when offered in combination with foundational *intrapersonal skills* and examine how returns differ by the gender of the business owner and social norms. Baseline and randomization were completed in August 2021 and the endline survey is planned to start in September 2023. The sample consists of 4,500 individuals for the RCT and a further 1,500 individuals for a regression discontinuity design used to examine the overall effect of the government program. This plan outlines the hypotheses to be tested and the specifications to be used in the analysis.

Gender, Agriculture, Entrepreneurship, Socioemotional Skills, Nigeria
[JEL] J16, J24, O12, Q12, L26

*This pre-analysis plan is a product of the World Bank Africa Gender Innovation Lab within the office of the Africa Region Chief Economist. We thank the Agro-Processing, Agricultural Productivity Enhancement and Livelihood Improvement Support (APPEALS) project team. This work has been funded in part by Wellspring Philanthropic Fund, as well as the World Bank Umbrella Fund for Gender Equality (UFGE) and other World Bank sources. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. IRB approval was obtained from HML IRB Research and Ethics, reference 681IPA19. The trial is registered on the AEA RCT registry (AEARCTR-0006134) at <https://doi.org/10.1257/rct.6134-2.0>. The paper was Accepted, Pre-Results Review at the Journal of Development Economics on September 22nd 2023.

[†]Innovations for Poverty Action (IPA) sdas@poverty-action.org

[‡]World Bank Africa Gender Innovation Lab cdelavallade@worldbank.org

[§]World Bank Africa Gender Innovation Lab afashogbon@worldbank.org

[¶]World Bank Africa Gender Innovation Lab spapineni@worldbank.org (Corresponding author)

1 Introduction

The labor market returns to social and emotional skills or "socioemotional" skills - such as self-efficacy, assertiveness, or optimism - have been well-documented in many high-income country contexts (for example, Borghans et al. (2008); Almlund et al. (2011); Lindqvist and Vestman (2011); Heckman and Kautz (2012); Deming (2017); Edin et al. (2022)).¹ Seminal work by Heckman (1995) showed that since measured cognitive ability explains only a small fraction of the variation in earnings, productivity is likely influenced by multiple dimensions of skill.² In low- and middle-income country (LMIC) settings there is a growing body of evidence that shows psychology-focused programs that strengthen socioemotional skills can positively impact economic outcomes, especially for women (see for example, Campos et al. (2017)). While this evidence has helped to influence policymakers to implement programs that build socioemotional skills, there is limited evidence on which socioemotional skills are most important for productivity and economic outcomes, and for whom (Lippman et al., 2015).

In this study, we use an individual-level randomized controlled trial (RCT) to examine the impact of the provision of different types of socioemotional skills (SES) training delivered to both men and women who are current and aspiring agribusiness owners in Nigeria. While SES refers to a number of skills, most can be categorized into skills that focus on *intrapersonal skills* (e.g. self-awareness and self-management) and *interpersonal skills* (e.g. social awareness and relationship management). Intrapersonal involves teaching skills to turn inward within one's own self and interpersonal relates to relationships or communication skills with others. This study is designed to explore whether the provision of interpersonal socioemotional skills alone or in combination with intrapersonal skills can better support the economic empowerment of female and male recipients. While the design will examine the impact of interpersonal skills training, and the additive impact of intrapersonal skills training, it may also shine light on whether intrapersonal skills are foundational for the development of interpersonal skills. For example, self-control may improve listening ability and awareness of desires may be essential for negotiation.

The SES training is delivered to beneficiaries of a Government of Nigeria development program that promotes entrepreneurship in the agribusiness sector.³ Embedding the socioe-

¹In the economics literature, socioemotional skills are often referred to as non-cognitive skills, soft skills, or life skills.

²Subsequent work summarized in Heckman and Kautz (2012) finds that socioemotional skills explain important variation in adult outcomes.

³The project is called Agro-Processing, Agricultural Productivity Enhancement and Livelihood Improvement Support (APPEALS) that aims to enhance the agricultural productivity of small and medium scale farmers and improve value addition along 11 priority value chains (poultry, cassava, rice, aquaculture, cashew,

motional skills intervention within the APPEALS WYEP project presents a unique opportunity to investigate our research questions as the program targets both men and women (50/50) and already alleviates capital and skills constraints through the provision of a grant for inputs and technical skills training to start an agribusiness.⁴ Program coverage across six states in Nigeria (3 in the North and 3 in the South) also allows for the exploration of how returns to SES varies with more conservative versus liberal gender norms.

Socioemotional skills training is often offered as a complement to programs that focus on economic empowerment. For example, training on SES is often included within vocational, technical and business training, training for maternal health or adolescent safe spaces, micro-finance, and graduation programs, to name a few. However, isolating the impact of SES has only recently been explored through impact evaluations in LMICs. For example, training on personal initiative (PI) in Togo had three times the impact of traditional business training on profits for women entrepreneurs (Campos et al., 2017). Skills for Effective Entrepreneurship Development (SEED) program for high school students in Uganda trains on soft-skills through a “mini-MBA” program and saw a 30% increase in profits (Chioda et al., 2021). STEP, a program targeting self-efficacy among college students has shown positive impacts in Uganda, Kenya, and Mexico with 34% more business starts one year after the training (Gielnik et al., 2012). In Malawi, an increase in skills such as perseverance, passion for work, and optimism were correlated with higher rates of adoption of more profitable cash crops (Montalvao et al., 2017).

While a few studies in the US and LMIC settings have shown that women exhibit higher returns to investment in socioemotional skills than men (Heckman et al. 2006; Campos et al., 2018). Few studies have systematically explored the relative returns to socioemotional skills for women versus men, especially with respect to increasing earnings. Socioemotional skills may influence economic outcomes for women more than men via an intrahousehold channel. For instance, if training improves self-efficacy, assertive communication, and perseverance, a woman’s bargaining power may improve, and the allocation of resources within the household may be more efficient to allow for higher economic success. This may have implications for other development outcomes, as the literature suggests that increasing women’s bargaining power in the household improves overall family welfare (Duflo, 2003; Qian, 2008; Thomas,

maize, wheat, tomatoes, ginger, dairy and cocoa) in 6 states (Lagos, Enugu, Kano, Cross River, Kaduna and Kogi) in Nigeria. A component of the APPEALS project – the Women and Youth Empowerment Program (WYEP) – will promote entrepreneurship in the agribusiness sector and targets young men (aged between 18 and 40 years) and women (over age 18). The APPEALS WYEP project provides beneficiaries with business and value-chain specific technical training, mentorship, and a labelled in-kind grant (approximately USD5,000 to USD10,000) linked to a business plan to support the start-up of the beneficiary’s agribusiness.

⁴Note -the authors will explore a quasi-experimental evaluation method (regression discontinuity design) to measure the impact of the overall APPEALS WYEP program that is not the focus of this paper.

1990). In the Nigerian context, socioemotional skills may be particularly relevant in places where care and household duties disproportionately fall onto women within the household.

It is however still unclear whether women may yield higher returns from intra- or interpersonal skills. On the one hand, teaching interpersonal skills may affect a woman's ability to negotiate deviations from restrictive gender norms, build bonds to tackle constraints cooperatively, increase belongingness in male-dominated occupations, and persuade her surrounding community to support her. On the other hand, the provision of intrapersonal skills may have a multiplier effect on economic empowerment by encouraging creativity, risk taking, and perseverance in instances where social support is not available. Intrapersonal skills may affect a woman's ability to set goals, plan, regulate emotions and time, take risks, persevere, and develop the confidence to succeed despite more binding capital and network constraints. The skills may prove more important if women need support as they look for resources, prioritize their agribusiness and long-term returns. In this study we will examine the impact of two SES programs on men and women's economic outcomes covering enterprise-level outcomes such as sales and profits, as well as intrahousehold bargaining power through a decision-making index, time use on care and household chores, and individual well-being as secondary outcomes.

An additional contribution of this study is in the measurement of the socioemotional skills. In addition to self-reported measures we will implement behavioral measures such as situational judgement tests and tasks to estimate levels of skills and impacts of the program on 14 socioemotional skills (see Figure 5 in Appendix for a list of the 14 socioemotional skills and their categorization as interpersonal or intrapersonal skills). These measures were developed specifically for this study and two companion studies.

To quantify overall APPEALS WYEP program impacts, we will also contrast outcomes with a quasi-experimental counterfactual group of non-program individuals located in three of the study states. We use a regression discontinuity design (RDD) analysis to estimate a local average treatment effect by exploiting a strict cut-off criteria that determined selection into the APPEALS WYEP program. Using a score recorded during the screening interview we will exploit this cutoff to establish a comparison group near the selection threshold. The comparison group are individuals who applied for the APPEALS WYEP program but just missed out on the program based on a score at interview stage.

The APPEALS WYEP program began in 2020 and the socioemotional skills interventions were delivered through classroom-based trainings between September 2021 and February 2022 in all 6 participating states. The interventions also include reinforcement video messages of training material via phones referred to as "emovis". Evidence shows that well-designed reminder campaigns can increase rates of behavior change in a range of policy areas from

agricultural production to savings (Grabowska, 2020). For example, there is evidence of the efficacy of reminders for achieving health outcomes (Tomlinson et al., 2013; Aranda-Jan et al., 2014). Messaging that includes visualization of the future or concrete help with planning has also been shown to improve preventative health outcomes, savings, technology adoption and investments in education (Orkin, 2020). The follow-up survey planned in September 2023 will be approximately 1.5-2 years after the socioemotional skills trainings were implemented at a time when all beneficiaries are expected to have received their in-kind grant to start their APPEALS WYEP agribusiness.

The remainder of this paper is organized as follows. Section 2 describes the research design including intervention details, outcomes, sources of data and hypotheses to be tested, while Section 3 describes the empirical strategy. Section 4 discusses any limitations and challenges of the study, Section 5 describes how impact results will be interpreted, and we provide administrative information in Section 6.

2 Research Design

2.1 Basic methodological framework

2.1.1 Randomized Controlled Trial (RCT) Design

The experimental design is a stratified randomized controlled trial (RCT) impact evaluation of the socioemotional skills trainings. Beneficiaries will be divided into groups, with individual random assignment stratified by gender of the business owner, by value chain selected for the APPEALS WYEP project, and by state geographic location:⁵

- T1: Treatment 1: Receives interpersonal socioemotional skills only through classroom instruction and reinforcement messages.
- T2: Treatment 2: Receives a combination of intrapersonal and interpersonal socioemotional skills through classroom instruction and reinforcement messages.
- C: Control Group: Receives no classroom SES training and receives placebo messaging on COVID-19 information.

⁵Note we have both male and female agribusiness owners across 6 states in the sample. The eligibility criteria to the APPEALS WYEP for men included being aged 18 to 40 years, unemployed, graduates of universities or colleges of agriculture or completed secondary education with minimum of 18 months experience in agricultural value chains. The same criteria were set for women with some exceptions: women above 40 years with less than senior secondary education but with experience in the agricultural value chain were still eligible for the program.

The treatment arms are mutually exclusive. Comparing average outcomes for the treatment and control group arms post-intervention will provide rigorous evidence on whether each program can generate impact.

2.1.2 Regression Discontinuity (RD) Design Analysis

A regression discontinuity (RD) design will be used to evaluate the overall impact of the APPEALS WYEP program. The APPEALS WYEP program was oversubscribed (applications were either made online or physically). The project team shortlisted candidates at interview stage based on specific eligibility criteria (youth, unemployed and educated). During the interview stage each state interviewed the shortlisted candidates and selected applicants for the APPEALS WYEP program using a score based on a screening interview where each candidate was scored within the range of 0 to 10. The project established a cut-off point for eligibility where a candidate's score in the interview must be 7 or higher to qualify for the APPEALS WYEP program. For the RD analysis we will exploit this cutoff by surveying a sample of shortlisted candidates at endline who scored above and below the cutoff score of 7 (ranging from 5-8) to establish a comparison group that represents non-APPEALS beneficiaries near the selection threshold. The analysis will be used to examine the overall impact of the APPEALS WYEP program (technical training and input grants). Further details of the screening scores and RD design are shared below.

2.2 Hypotheses

The study will examine how each training program affects business performance and productivity, measuring key enterprise-level outcomes such as annual and monthly profits and sales, capital investment, employment, and hours worked. It will also examine changes in the business owner's socioemotional skills, household income, consumption, intrahousehold power relations, and women's empowerment and decision-making within the household. The theory behind the SES training is that investment in socioemotional skills leaves the respondent with a greater ability to make more effective use of the capital and technical skills provided by the APPEALS WYEP program. This leads to an improvement in business practices, and/or increased women's decision-making power and agency. This in turn improves business performance, making them and their household wealthier.

There are six key hypotheses to test:

Hypothesis 1: Neither interpersonal SES training (T1) nor the combination of interpersonal + intrapersonal SES training (T2) has an impact on economic outcomes, well-being outcomes, and decision-making power.

Hypothesis 2: Neither the interpersonal SES training (T1) nor the combination training (T2) has an impact on socioemotional skill constructs.

Hypothesis 3: The interpersonal SES training (T1) and combination training (T2) has no differentiated impact on socioemotional skills, decision-making power, economic outcomes, and well-being outcomes.

Hypothesis 4: The impact of either treatment arm will not vary by gender.

Hypothesis 5: The impact of either treatment arm will not vary with gender norms.

Hypothesis 6: The APPEALS WYEP (technical skills and input grants only) has no impact on economic outcomes, socioemotional skills, well-being outcomes, and decision-making power on beneficiaries right above the eligibility score threshold versus a comparison group who applied but right below the eligibility cutoff and therefore did not receive the APPEALS WYEP program.

2.3 Outcome variable(s)

The key outcome variables that will be captured from the various sources of data are listed in Tables 5-7. The primary outcomes organized by category are listed in Table 5, the potential mediator variables are in Table 6 and secondary outcomes are in Table 7 in the Appendix.

Our primary outcomes are enterprise-level outcomes (e.g. profits, sales and employees); and wellbeing (income, consumption and women’s empowerment). For most of the monetary variables such as profits and sales we will use the inverse hyperbolic sine (IHS) transformation and include the mean of the outcome variable for the control group in Naira in levels at the bottom of the tables. We will test for the sensitivity of our results on profits and sales to outliers by estimating winsorized versions of these variables at the 99th percentile.

The psychological constructs will include 14 socioemotional skill outcomes such as indices measuring generalized self-efficacy, problem solving and decision-making, relatedness, emotional regulation, and influence. The baseline survey was limited to a phone-based survey due to COVID-19 so we could only include self-reports of a few skills. Prior to the endline survey we will test behavioral measures of the socioemotional skills, such as situational judgement tests and tasks. These behavioral measures will be collected alongside the self-reports in separate sessions held in a central location for a subsample of the total impact evaluation sample (1,200 respondents evenly sampled by gender, state and treatment status). The endline survey will be approximately 2 hours conducted in the respondent’s business or household and will target the full RCT IE sample of 4,500 respondents. A further 1,500 respondents with screening scores shared will be sampled from 3 states only to conduct the regression discontinuity analysis.

2.4 Intervention(s)

2.4.1 Program Details

The socioemotional skills intervention was delivered through classroom-based training in local state training centers. All individuals assigned to treatment were invited to 4 consecutive days of classroom training. The intervention also included reinforcement video messages of training material referred to as "emovis" that were sent via WhatsApp and SMS messages to reinforce skills taught in the classroom and to increase accountability among training participants.

An experienced firm (Alkimia) who is specialized in SES training for youth in Sub-Saharan Africa was contracted to develop and design the curriculum and materials for the socioemotional skills training, as well as to conduct the training of trainers and facilitation experience with trainers. Materials were specifically designed for low-literacy learners and are intended for implementation in low-resource settings. The training was sensitive to gender, language, and culture of each of the states. Based on whether the skills are directed at observing and managing emotions in oneself or others, SES skills were divided into two training modules "Inward" and "Outward with Others". Alkimia was requested to include training on seven specific skills in each module: "Inward" included training on emotional awareness, self awareness, emotional regulation, self control, perseverance, and problem solving; "Outward with Others" included training on listening, empathy, expressiveness, interpersonal relatedness, interpersonal influence, negotiation, and collaboration.

While the skills are organized in these two themes, in this impact evaluation we will test the average impact of the interpersonal skills alone, compared to a combination of both the inter- and intra-personal skills. As both training types will be equal in length, the team will be able to evaluate whether two pieces of training of the same length would be better spent on 7 interpersonal skills, or 7 interpersonal and 7 intrapersonal. Definitions of the skills and modules of the training can be found in the Appendix.

Training facilitators were selected within each state, and were interviewed by the research team and implementing partner for their ability to speak clearly and translate SES concepts, previous training experience, and thoughts on mobilization and motivation to reduce dropout of trainees (a common issue in training programs, particularly among women). Where possible, facilitators' experience in agriculture and entrepreneurship was also considered. In each state, 4 facilitators were selected. All facilitators attended 8 days of training which included 4 days of Training of Trainers (ToT) and 4 days of Facilitation Experience - where facilitators did mock facilitation and were provided feedback by the lead trainers. During the actual training of beneficiaries, each training facilitator was assigned to a classroom of

no more than 30 individuals at a time and the training ran for 4 days per batch. On average, each state had 4 batches and 4 parallel classes per batch. The locations of the training were organized using the support of APPEALS project coordinators in each state, and the training typically took place in a training institute or hotel.

The frequency of the reinforcement messaging delivered approximately one year after the classroom training included 2-3 messaging touch-points per week. These messages were interactive and incorporated behavioral change exercises into the messaging to make the information more participatory rather than passive (Lubega et al., 2021).⁶

2.4.2 Details of Screening Scores for APPEALS WYEP

Pre-intervention screening interviews for the APPEALS WYEP project were conducted between August-November 2019 for 11,691 individuals (5,469 men and 6,222 women) across 5 states Lagos, Kaduna, Kano, Kogi and Cross River.⁷ A short survey was administered during the interview stage to all applicants of the APPEALS WYEP project. The APPEALS project will provide economic support to beneficiaries of the program through the provision of technical and business training and input grants equivalent to approximately 2million Nigerian Naira (approximately USD 5,300). The short survey at screening stage collected information from all APPEALS WYEP applicants on their chosen value chain out of 11 possible value chains and elicited information on social norms towards women working and socio-emotional skills, as well as decision-making, demographic characteristics and current and past employment status. Due to the constraints on the time to complete the survey, the questions measuring a set of socio-emotional skills were collected only from a random sub-sample of 50% of applicants whereby the randomization was automatically completed within the enumerators tablet stratified by gender and state. The survey duration was approximately 15 minutes and was conducted during the interview stage for the APPEALS WYEP project. In all the five states, trained enumerators administered the short survey to WYEP applicants who had verified their identity and submitted the relevant documents to be considered for the project.

Each state then had to shortlist the applications for interviews. All applicants were considered eligible based on unemployment status, age and education eligibility criteria. Criteria for shortlisting was that beneficiaries had to be gainfully unemployed, had to have a business

⁶Interactive messages take the form of reflective questions, offer reminders/encouragement, encourage planning and the selection and commitment to goals, and games or exercises to encourage the practice of any skills learned.

⁷Enugu did not conduct a screening stage since they were delayed in the start of the programming versus the other states.

idea, plus age and education criteria.⁸ Also if the applicants had already benefited from other government programs (e.g. CADP, FADAMA) they were disqualified from participating in the APPEALS WYEP project.

The interview for the project consisted of a few short questions from program officials and involved meeting project personnel, officers from the Ministry of Agriculture and some well-established agribusiness owners in the state. The project was advertised in several media outlets including newspapers, social media, radio and TV stations in Nigeria. Depending on each state, prospective applicants were required to complete either paper or web-based applications with evidence of fulfilling eligibility criteria. All applications that satisfied the eligibility criteria set by the program (11,691) were invited for face-to-face panel interviews held in their state and approximately 50% of the total applicants were eventually selected for the APPEALS WYEP program.

The project selected beneficiaries into the APPEALS WYEP program by conducting face-to-face interviews with the applicants and a panel of 6-8 experts in each state. The experts were drawn from the state project team, federal support team, Permanent Secretary Ministry of Agriculture of the state and farmers' associations in the states. The applicants were asked 5 random questions from a pool of questions and scored a maximum of 2 points per question. All applicants were graded by the members of the panel and scores were averaged across the panel members. Within the score range of 0 to 10, the project set an eligibility cut-off of 7 points. Therefore, to be eligible for the program an applicant must have scored 7 or above. The applicants who were selected for the APPEALS WYEP program, on average, do not differ along observable characteristics compared with the applicants who were not selected for the program.

Pool of potential questions that may have been asked at interview stage:

1. How did you get to know about the APPEALS WYEP? e.g. Advert/TV/newspaper
2. Tell me about yourself
3. What do you understand about agribusiness?
4. What do you know about agriculture?
5. What are you interested in exploring for the project? i.e. describe value chain and business idea
6. If you had the opportunity to advise the commissioner of agriculture in Kano what would you suggest? (Trying to address the challenges specific to the state they were resident of e.g. road networks etc)

⁸For example, in Kaduna they had approximately 40,000 applications that they shortlisted to 3,000 for the interview stage, and they eventually recruited 1,700 for the APPEALS WYEP.

7. Ask specific technical knowledge around the chosen value chain. E.g. if choose poultry then ask “how would you manage the waste from your poultry business?”

The interview was more of a discussion (no right or wrong answer per se) to gauge an individual’s general understanding of agriculture and technical knowledge around their business idea’s chosen value chain. The interview panelists rated each candidate, averaging their collective evaluations to determine the results. The project established a cut-off point for eligibility of 7 points within the range of 0 to 10. Therefore, a candidate’s score must be 7 or higher to qualify for the APPEALS WYEP program. The panel decided on a cutoff of 7 prior to the interview stage because that was sufficiently above the 50% scoring mark but there was no other reason than that.

Candidates who scored 0 were those who just heard about project and tried their luck - they are assumed to have limited knowledge about agriculture and no specific business plan. Those who scored between 1-6 did not qualify for the APPEALS WYEP program but may have had some knowledge on agriculture and/or demonstrated interest in the project. Those who scored 7 and above were invited to be part of APPEALS WYEP (each state targeted 1,700 participants) where they received 2 weeks of technical training and grants.

Screening scores were shared for 8,500 applicants across 4 states, and we will restrict to 3 states for the regression discontinuity analysis.⁹ Given that the interview was more of a subjective decision (rather than based on objective facts) and given our budget restriction of only being able to survey 1,500 for the RD analysis we decided to interview those closest to the cutoff (scoring 5, 6, 7 or 8 only). We randomly selected a sample from non-APPEALS below the cutoff (5s and 6s), and APPEALS beneficiaries above the cutoff (7s and 8s) stratified by state, gender and screening score. See Table 8 in the Appendix that shows that among applicants to APPEALS WYEP there was no evidence of selection on observable characteristics into the APPEALS WYEP program.

2.4.3 Training Take-up Rates

The average take-up rate of the classroom training across all 6 states was 79%. This is very high compared to other business training impact evaluations (McKenzie, 2020). On average, we find no significant statistical differences on take-up across the T1 and T2 treatment arms. In Table 3 in Appendix A we show the correlates of take-up by estimating differences in baseline characteristics between those who took up the training and those who did not.

⁹Enugu did not have a screening stage and Kano did not accurately record the screening scores. Additionally, Lagos interviewed fewer candidates overall meaning there were only a few candidates not selected.

Table 1: Training Take-up Rates

Take-up Rate of SES Training By State	Numbers and Percentage Trained in Each State			Percentage Trained by Treatment Assignment and Gender				
	Numbers Trained in Each State			Total Percentage Trained	T1 Inter only Male	T2 Combination inter_intra Male	T1 Inter only Female	T2 Combination inter_intra Female
State	No	Yes	Total	% Trained out of those Assigned				
CrossRiver	87	413	500	82.6%	80.0%	86.4%	83.2%	80.8%
Kaduna	79	421	500	84.2%	86.4%	86.4%	87.2%	76.8%
Kano	132	368	500	73.6%	74.4%	72.8%	72.8%	74.4%
Kogi	103	397	500	79.4%	80.0%	78.4%	78.4%	80.8%
Lagos	133	367	500	73.4%	67.2%	73.6%	75.2%	77.6%
Enugu	96	404	500	80.8%	79.2%	84.0%	76.8%	83.2%
Total	630	2,370	3,000	79.0%	77.9%	80.3%	78.9%	78.9%

Notes: Take-up rates calculated using administrative data collected by the implementing partner and beneficiary surveys collected during the trainings.

2.4.4 Impact Evaluation Design and Randomization Strategy

The impact evaluation design is summarized in Figure 1. This impact evaluation is an individual-level randomized controlled trial (RCT) design, to identify the effects of psychology-based interpersonal and intrapersonal socioemotional skills training. As the goal of the evaluation is to also understand the gender-disaggregated impact of the program, the eligible beneficiaries are stratified by gender. The RCT participants in the control group are still beneficiaries of the APPEALS WYEP bundled program but will not be offered socioemotional skills training. The use of randomization will allow us to attribute differences between groups to the socioemotional skills intervention itself. Comparing outcomes for the treatment and control group arms will provide rigorous evidence on whether this program can generate impact.

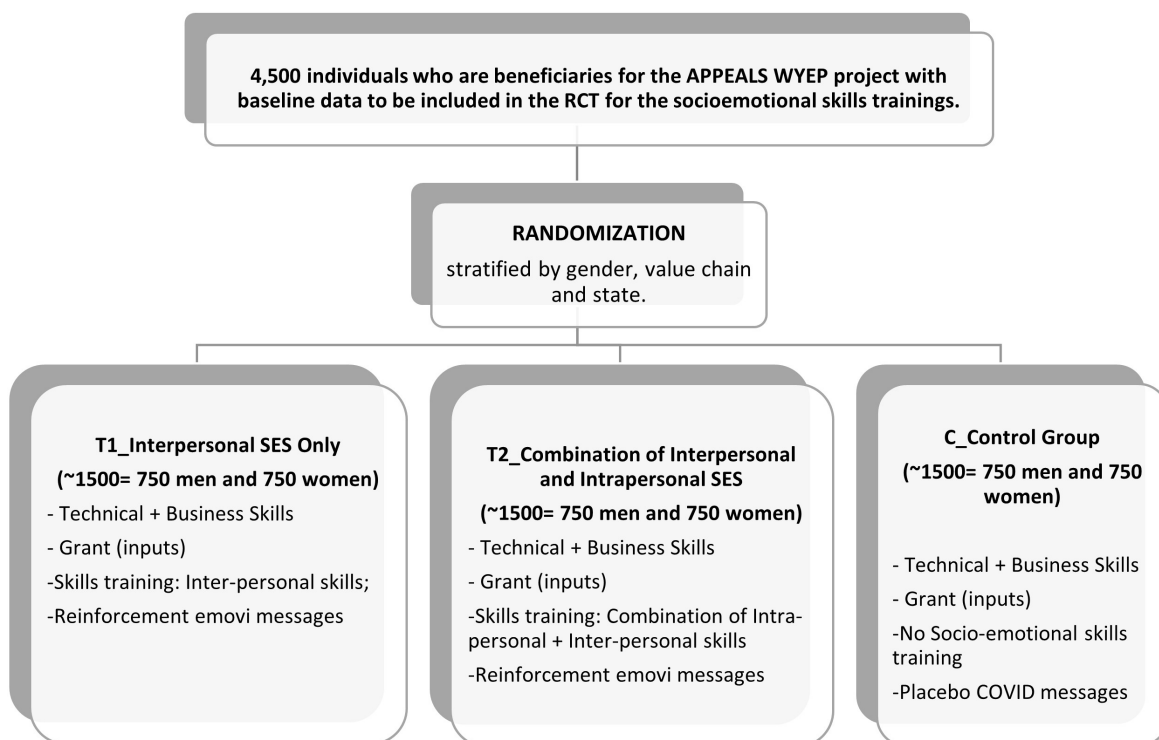


Figure 1: Impact Evaluation Design

The randomization is stratified by state, gender, and value chain.¹⁰ In each stratum, the sample is randomly assigned to one of the treatment arms. We randomly allocate the APPEALS WYEP beneficiaries into T1 (1/3), T2 (1/3), and C (1/3). The impact of T1

¹⁰Value chain is that selected for their APPEALS WYEP agribusiness out of eleven priority value chains identified by the project

(interpersonal skills only) will be assessed by comparing T1 outcomes to C outcomes while the marginal impact of T2 (intrapersonal + interpersonal skills) will be estimated by comparing T2 outcomes to T1 outcomes. The combined impact of the combination of intra- and interpersonal skills will be estimated by comparing T2 outcomes to C outcomes. This experiment will enable us to identify the impact of offering interpersonal skills; plus the marginal impact of intrapersonal skills to aspiring agribusiness owners.

The control group will not be invited to attend any SES training for the duration of the study. Since the reinforcement messages are to be sent over the phone, in order to minimize sharing of video content between treatment and control groups, the research team also developed placebo video messages containing information about COVID-19 to send to the control group and treatment group who did not attend a classroom-based training. The treatment group who did attend the classroom training will get specific skills-related content depending on the treatment assignment i.e. T1 receives only the interpersonal skills content and T2 receives both inter- and intra-personal skills content.

2.5 Theory of Change

The socioemotional skills (SES) intervention is expected to impart social and emotional skills such that beneficiaries profit more effectively from the economic components of the APPEALS WYEP program.

The APPEALS WYEP aims to address major constraints to market entry faced by potential agribusiness entrepreneurs, thereby providing business opportunities for women and young men as well as combating unemployment and low agricultural output. The combination of in-kind grants, training and business registration certificates will give the beneficiary improved access to both inputs and sales markets for the beneficiary to start their business activity. Beneficiaries may potentially obtain (additional) credits or grants from informal or formal sources to expand their business(es) and employ additional workers, making use of their improved access to inputs and sales markets. Sales generated by the firm's outputs will result in profits which can be either re-invested in the firm or used as income to support household consumption or investments, thus improving economic well-being. Higher household income might furthermore lead to an improvement in social welfare indicators, such as access to health services or children's school attendance. By employing additional workers, the new enterprises may also contribute to employment creation in the local community.

The set of socioemotional skills included in the curricula was designed to span a range of skills while capturing and differentiating between skills that have previously been found

to matter for labor market outcomes. Training on intrapersonal skills may improve: (1) awareness and emotional regulation, such that individuals can set goals, determine obstacles and areas for growth, and remain calm in the face of obstacles, (2) confidence and ability to persevere and focus despite obstacles, (3) self control to better manage attention, time, and funds, and (4) personal initiative and problem-solving and decision-making, to plan for opportunities, innovate and pro-actively tackle challenges, and search and access available resources. Training on interpersonal skills may improve beneficiaries' (1) sense of empathy, and thus improve their ability to build trust and resolve issues with collaborators and customers, (2) social perceptiveness and discovery of opportunities, (3) ability to listen and communicate clearly and assert their needs and wants, (4) ability to proactively initiate relationships and grow business networks, and (5) ability to persuade, negotiate, and collaborate with competitors for market share, with suppliers and customers for pricing and synergies, and with household members for access to inputs, assets, capital, and time. The Theory of Change is summarized in Figure 2.

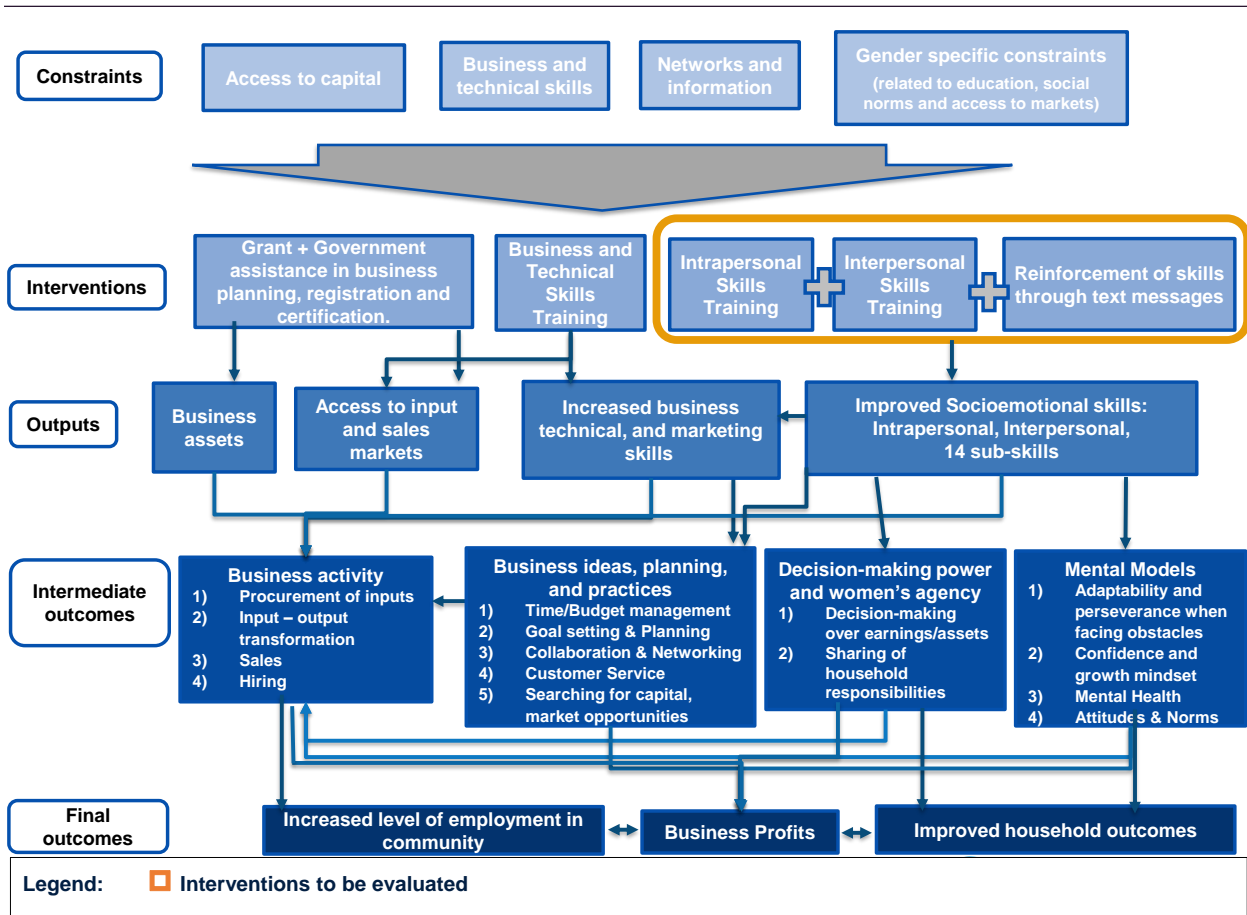


Figure 2: Theory of Change

An important contribution of our study will be to test whether returns to socio-emotional skills are higher for women than for men. Several studies suggest that SES may account for a portion of the gender gap in education and income (Jacob, 2002; Mueller and Plug, 2006). Reinforcing particular SES may make it possible to build aspirations and counter traditional norms, to the benefit of women more than men who do not face similar barriers. A few studies in the US and low- and middle-income countries have shown that women exhibit higher returns to investment in socio-emotional skills than men (Heckman et al., 2006; Campos et al., 2017). But little work has systematically explored the relative returns to socio-emotional skills for women versus for men, especially with respect to (i) comparing relative returns to various socio-emotional skills, and (ii) increasing men and women’s earnings and profits.

We may expect gender norms to restrict women’s economic choices. For example, norms around the acceptability of women’s work itself, women’s mobility, communication with men in the household, capacity to interact with men outside of the household and/or the male breadwinner status could all be at play when it comes to a woman’s choice of which agricultural value chain to operate in or her ability to grow a business. Additionally, the unequal burden of domestic responsibilities for women may limit the time available for work or her ability to seek work far from home; the communal belief that women should not earn more than their husbands may limit a woman’s aspirations, agency, and ability to obtain capital to grow a business; norms surrounding unequal land inheritance, and property rights may limit women’s ownership and control over land and other assets; limited decision-making power may affect women’s access to capital, labor, choice of work, and use of transport; customs that prevent men from talking to women without the permission or presence of her husband may make men less open to collaborating, teaching, sharing information, and creating contracts with women (Friedson-Ridenour and Pierotti, 2019; Goldstein and Udry, 2008). Gender norms which restrict female interaction with the opposite sex may limit development of social and business networks and access to information as well as input/output markets (O’Sullivan et al., 2014). Lower access to inputs such as land, capital, labor, fertilizer extension services, tenure security, and information, drives much of the gender differential in agricultural yield (Croppenstedt et al., 2013).

Socioemotional skills may improve women’s ability to defy these norms or overcome their resulting barriers. Women may use interpersonal skills to negotiate in the household, build networks and bonds to tackle constraints cooperatively, lead and manage teams, increase belongingness in male-dominated occupations, and persuade her family to support her. Meanwhile intrapersonal skills may be used to set goals when aspirations are limited, plan strategically and persevere in the face of adversity, regulate emotions and time. Personal

initiative may be required for women to explore new products and businesses (Campos et al., 2017), expand social networks, and obtain access to information and capital. Conceptually, higher self-confidence may be a prerequisite for being willing to defy norms and shed fears of social judgement (Campos et al., 2015) and upgrade career aspirations (Correll, 2004).

Specifically, an intrahousehold bargaining channel may predict that socioemotional skills to women has a greater influence on her economic outcomes than those to men. For instance, if training improves assertive and persuasive communication, a woman may be able to increase her bargaining power, and the allocation of resources within the household may be more efficient to allow for higher economic success. This may have implications for other development outcomes, as the literature suggests that increasing women’s bargaining power in the household improves overall family welfare (Duflo, 2003; Qian, 2008; Thomas, 1990). Skills such as expressiveness, influence, and perseverance may help women materialize preferences on childbearing, education, the sharing of household responsibilities and assets, and labor market participation.

In the Nigerian context, interpersonal skills may be particularly relevant since care and household duties disproportionately fall onto women within the household. Gaining socioemotional skills may prove important if women also need support from husbands as they look for resources, prioritize their business and long-term returns of work. With formal credit markets being scarcer in Sub-Saharan Africa, married female entrepreneurs often rely on their husbands to provide them with seed investment and working capital. In such contexts, interpersonal skills may help negotiate deviation from restrictive gender norms and, more generally, communication – within-household and within-community – could prove to be a determinant tool to women’s entrepreneurial success. There is little empirical evidence of the impact of negotiation skills training programs. A training program in Zambia that focused on adolescent girls significantly improved educational outcomes over a three-year period (Ashraf et al., 2018).

Defying norms may come with costs that range from violence to social ostracism or poor marriage prospects and mental health. These costs have generally been termed the “backlash effect” (Rudman and Fairchild, 2004). While some socioemotional skills may help women overcome these norms and their associated backlash, the use of certain skills such as expressiveness and negotiation may cause this backlash. Even simple deviating behaviors, such as advocating for oneself can have repercussions on the economic returns and likeability of women (Bowles et al., 2005). Gender inequitable norms have been found, in some instances, to have a negative correlation with self-esteem (Stark et al., 2018).

2.6 Sample

The unit of analysis is at the individual-level and the firm-level.¹¹

2.6.1 Sample Selection

The project was advertised in several media outlets including newspapers, social media, radio and TV stations in Nigeria. The Women and Youth Empowerment Program (WYEP) component of the APPEALS project targeted both men and women. Men had to be 18-40 years, unemployed graduates of universities or colleges of agriculture or have completed secondary education with a minimum of 18 months experience in agricultural value chains. Women had to be older than 18 years (without an upper age limit) with the same education and work experience eligibility criteria, except that women above 40 years of age with less than senior secondary education but with experience in the agricultural value chain were also eligible for the program. Individuals with a disability were also encouraged to apply.

Depending on each state, prospective applicants were required to complete either paper or web-based applications with evidence of fulfilling eligibility criteria. A total 133,584 completed applications were received and screened using project eligibility criteria. Shortlisted applicants who had verified their identity and submitted the relevant documents were invited to a face-to-face screening interview held in their state. As part of a listing exercise for the socioemotional skills RCT we conducted a short 15 min quantitative survey in August 2019-February 2020 from 11,691 individuals (5,469 men and 6,222 women). The listing survey was administered during the interview stage to all applicants to the project. Approximately 10% of the total applicants were invited for face-to-face interviews with a panel in each state, with 6% (10,300) of the total applicants eventually selected for the APPEALS WYEP (see beneficiary numbers per state in Table 5).

The interview for the project consisted of a few short questions from program officials and involved meeting project personnel, officers from the Ministry of Agriculture and some well-established agribusiness owners in the state. Participants scoring at least 7 (range 0-10) in the screening interview were deemed eligible for the APPEALS WYEP and therefore invited to participate. Applicants with a score lower than 7 were not deemed eligible for the program. The assignment of participants to the APPEALS WYEP program at the 7 score cutoff was largely adhered to by the government implementer.¹² In total, approximately 50%

¹¹The respondents are individuals who are all aspiring business owners as they are beneficiaries of the APPEALS WYEP and would have received a grant to start an agribusiness through the program. Many of the socioemotional skills outcomes are collected at the individual-level and we are interested in a number of firm-level outcomes.

¹²In a supplementary paper we intend to use the eligibility criteria cutoff to exploit a sharp regression

of the interviewed applicants were eventually selected for the APPEALS WYEP program (approximately 1,700 in each state) which is equivalent to no more than 0.5% of the eligible population of a given state receiving the APPEALS WYEP program.

The applicants who were classified by the APPEALS WYEP as beneficiaries based on interviews were eligible for the baseline survey for the SES study. Among beneficiaries and those who opted-in for training (7,796) we conducted a longer baseline survey (45 mins) in June-July 2020 over the phone. Finally, for the impact evaluation we randomly selected 4,500 individuals (2,250 men and 2,250 women) based on power calculations described below.

Data for this impact evaluation will be collected through quantitative tools such as survey instruments, administrative data (from project Monitoring and Evaluation systems), and qualitative interviews. A survey among the trainers will also be conducted. The research team develops the questionnaires for the various surveys and provides guidance for field data collection. A data collection firm will be hired to administer the individual and enterprise surveys for each of the survey rounds. The role of the firm is to select and train enumerators and collect the data based on the questionnaire. The data is analyzed by the research team.

As part of a listing exercise for the socioemotional skills RCT we conducted a short 15 min quantitative survey in August 2019-February 2020 from 11,691 individuals (5,469 men and 6,222 women) across Lagos, Kaduna, Kano, Kogi, Enugu and Cross River states in Nigeria. The listing survey was administered during the interview stage to all applicants to the APPEALS WYEP project.¹³ The applicants who were classified by the APPEALS WYEP as beneficiaries based on interviews were eligible for the baseline survey for the SES study. Among beneficiaries and those who opted-in for training (7,796) we conducted a longer baseline survey (45 mins) in June-July 2020 over the phone. Finally, for the impact evaluation we randomly selected 4,500 individuals (2,250 men and 2,250 women) based on power calculations as described below.¹⁴

discontinuity design (RDD) to estimate the local average treatment effect (LATE) of the overall APPEALS WYEP program in 3 states. Here we use a data-driven selected window close to the 7 cut-off to establish a comparison group i.e. the counterfactual.

¹³The project was advertised in several media outlets including newspapers, social media, radio and TV stations in Nigeria. Depending on each state, prospective applicants were required to complete either paper or web-based applications with evidence of fulfilling eligibility criteria. A total 133,584 completed applications were received and screened using project eligibility criteria, 9% of the total applicants were invited for face-to-face interviews with panel interviews constituted in each state, with 6% (8,080) of the total applicants selected for the APPEALS WYEP.

¹⁴Note, an additional 1,500 were randomly sampled from the screening database above and below the program screening cutoff for the RD.

2.6.2 Sample Size and Power Calculations

Power calculations were conducted to determine the minimum sample size required to detect a treatment effect of 20% over the control mean at 85% power. 85% is used instead of 80% (the traditional rule of thumb) to account for potential attrition in data collection between the baseline and follow-up rounds. Power calculations were based on the following assumptions:

- Survey rounds: one baseline survey and follow-up survey will be conducted.
- Variation in outcome measures: the ratio of the standard deviation to the mean for both treatment and control group often differs by variable type. While variables tied to earnings and consumption are particularly noisy, past studies indicate that (1) standard deviations for measures of socioemotional skills are far lower than other outcomes, and (2) it may be important to detect smaller effect sizes. As socioemotional skills are key to this intervention, and new measures will be utilized, selected sample sizes are particularly conservative. Thus, Table 4 indicates required minimum sample sizes for different variable types. For profits and consumption measures we assume the standard deviation to the mean for both treatment and control group is calculated using a ratio of 1; and for socioemotional skills a ratio of 0.25 is assumed.
- Correlation across survey rounds: the correlation in the outcome variable between baseline and follow-up survey is 0.2. We assume constant autocorrelation across the pre- and post-follow up surveys. This is a conservative estimate given the examination of similar studies in the literature (McKenzie and Woodruff, 2013). The assumption of low auto-correlation in the outcome variable is sensible given that business outcomes such as profits and sales exhibit low correlation over time and that most beneficiaries will start up new businesses, which implies significant changes in business outcomes between baseline and endline.
- Take-up: 80% take-up was achieved. However, in determining sample size for the study we assumed 70% take-up.

Based on the power calculations (see Table 4 in the Appendix), we chose a sample size of 750 per treatment arm per gender. The stratification by state helps to reduce noise by accounting for some of the (location dependent) variance in outcome measures. The proposed sample size will allow for an assessment of the overall impact of the socioemotional skills on beneficiaries, the marginal impact of foundational intrapersonal skills, as well as an analysis by gender.

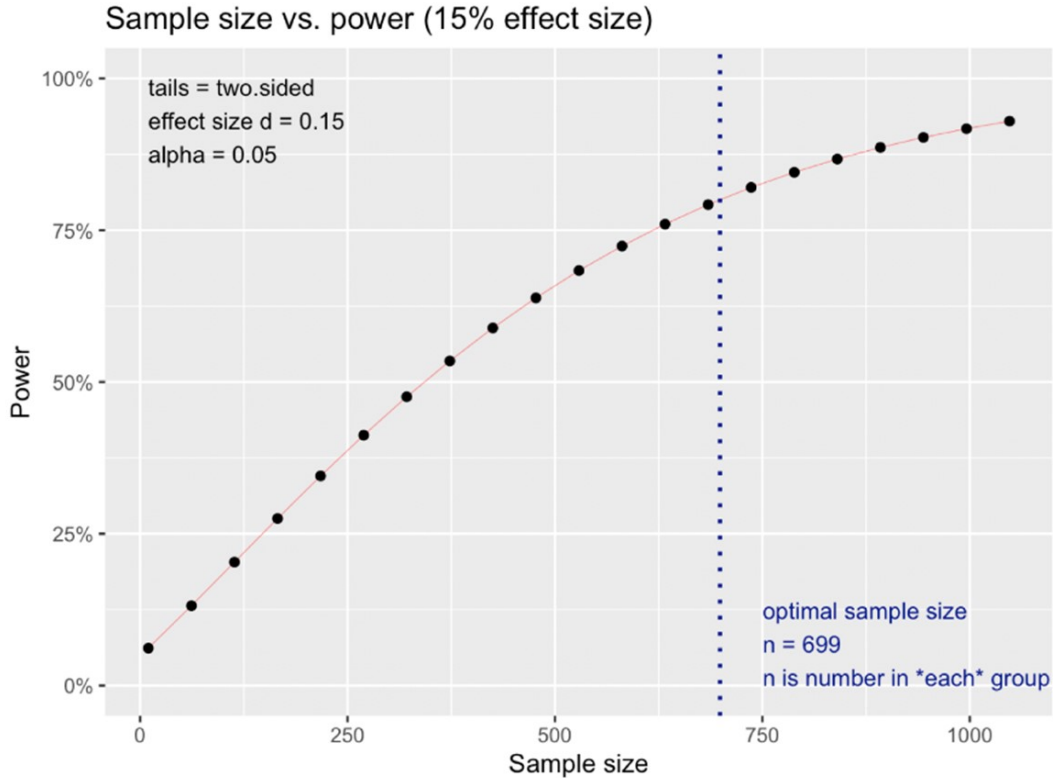


Figure 3: Sample Size Estimate for 15% Effect Size with Ratio of SD to Mean equals 1

2.7 Data collection and processing

The main source of data for this impact evaluation are quantitative surveys administered to the business owners. These surveys will be completed at two points in time: one baseline (completed in June-August 2020) and follow-up survey (endline) expected to start in September 2023. Data will be collected in private from each respondent using trained enumerators and informed consent from respondents will be collected before the surveys begin.

In the baseline survey conducted in June-August 2020, we collected data on personal characteristics of the respondent (demographic characteristics), employment and income status, mental health, quality of the relationship with their spouse, households demands including providing childcare, caring for sick/elderly, time use, household chores. Adults were asked about decision-making and attitude towards intimate partner violence (IPV), gender perceptions (gender attitudes and gender roles), gender norms, behavioral characteristics (conformism, risk preferences), aspirations for their children, and opinions on FGM and early marriage. The baseline household module included: a household listing, demographics, food security, household assets and consumption module.

The endline survey is planned to start in September 2023. The follow-up survey will

collect information from the business owners on the areas listed above, as well as information on the sharing of household duties, support from family members, decision-making power, women's economic empowerment, business behaviors linked to socioemotional skills- goal setting, searching for opportunities, negotiation, collaboration; program participation for both APPEALS and other programs, and gender-based violence.¹⁵

During the time of the endline survey we will also collect data from a random subsample of 1,200 respondents in separate sessions to capture behavioral measures of SES. These sessions will be used to capture demographic and economic indicators and will include situational judgement tests and tasks, in addition to self-reported skill measures. These sessions will last approximately 2-2.5 hours. This data will be used to supplement the endline data collected through one-on-one surveys from all 4,500 respondents. In terms of tracking, data collection will only be terminated after each respondent has been called to be interviewed at least 6 times over the phone; all phone numbers have been tried; and the number of observations is such that the survey completion rate is over 85%-90% of the targeted sample.

During the SES training, we also collected supplementary data from the trainers and the training participants. The facilitator interview took place the day before training started and consisted of a 1.5-hour survey. Surveys with training participants included skills measures for personal initiative (PI), perseverance, collaboration, and negotiation that was self-administered on tablets on day 1 of the training.

¹⁵We will use administrative data from the implementing partner's monitoring and evaluation activities to serve as a robustness check for program participation.

3 Empirical analysis

3.1 Randomized Controlled Trial Statistical Model

We estimate the intention-to-treat (ITT) impacts. We will use an analysis of covariance (ANCOVA) estimator to assess the impact of the interpersonal only SES training and any marginal impacts of the foundational intrapersonal skills on outcomes of interest for which we have both baseline and follow-up data.¹⁶ We estimate the treatment impact for individuals in the following regression specification:

$$Y_{i,t} = \beta_0 + \beta_{T1}InterOnlySES_i + \beta_{T2}InterIntraSES_i + \beta_3Y_{i,0} + \beta_4X'_{i,0} + \lambda_s + \varepsilon_{i,t} \quad (1)$$

$Y_{i,t}$ is the outcome variable for individual i measured at time t ($t = 1$ at endline post-treatment), and $Y_{i,0}$ is the baseline value of the outcome variable. *InterOnlySES* is a dummy variable for random assignment to the T1: Interpersonal Only SES training, and *InterIntraSES* is a dummy variable for random assignment to the T2: Combination Interpersonal and Intrapersonal SES training. β_{T1} and β_{T2} will measure the treatment effects relative to the control group (i.e. those who were not assigned to receive any SES training). $X'_{i,0}$ is a vector of baseline controls, λ_s are randomization strata fixed effects, and $\varepsilon_{i,t}$ is the error term. Huber-White robust standard errors are used throughout. We will report the estimates and significance for $\beta_{T1} = \beta_{T2}$ to test the null hypothesis that there is no difference in the effect of the *InterOnlySES* and *InterIntraSES* trainings.

For outcome variables that were only collected during the endline survey, we will rely on the random assignment of treatment status and use ordinary least squares (OLS) estimation as outlined in equation 2 to compare outcomes for treatment and control groups.

$$Y_{i,t} = \beta_0 + \beta_{T1}InterOnlySES_i + \beta_{T2}InterIntraSES_i + \beta_3X'_{i,0} + \lambda_s + \varepsilon_{i,t} \quad (2)$$

Equation 1 and equation 2 will provide the intention-to-treat (ITT) estimates, which is the effect of being assigned to attend the T1 or T2 training sessions among the sample. To estimate the impact by gender we will interact the treatment dummy variables in equations 1 and 2 with a dummy variable for *Female* equal to 1 if the individual is female; 0 if male.

We will also conduct several robustness checks for our main treatment results: in the regression analysis we will control for a **social desirability scale (SDS)** and interact the scale with treatment to examine if any treatment effects are driven by responding with a so-

¹⁶The ANCOVA estimator typically has more statistical power than a difference-in-differences estimator (McKenzie, 2012).

cially desirable answer.¹⁷ We will include a set of baseline characteristics that are unbalanced across treatment and control groups at baseline and a set of core demographic variables such as age and marital status. We will also test whether effects are robust to including additional control variables selected using the **double-LASSO-selected controls** procedure of Belloni et al. (2014).

3.1.1 Outliers

For variables measured as a value, such as household income, we will use winsorized versions of these variables at the 99th percentile. Where appropriate, we will apply the inverse hyperbolic sine (IHS) transformation to variables measured as a value. We plan to document how many zeros are in the data and the magnitude of the effect at the extensive margin to determine the appropriateness of interpreting the IHS transform as a percentage change in our results.

3.1.2 Missing Values

To preserve statistical power we will perform an interpolation for missing values due to item non-response by inserting medians and including dummy variable for missing variables.

3.2 Heterogeneous effects

For the heterogeneity analysis, we will look at the distribution of effects for different subgroups. We will estimate heterogeneous treatment effects by interacting treatment status with the outcome of interest in equation 1 (for outcome variables collected at both baseline and follow-up) and equation 2 (for outcome variables collected only at endline). To test interaction effects, we will utilize multiple variable regression analyses and include the product of centered variables as the interaction term (Aiken et al., 1991). We will use specific pre-specified variables collected at baseline including: marital status, index of gender attitudes, and gender norms. We will also examine if there are any differences in impacts based on location by comparing the North vs South of Nigeria by creating a dummy variable North equal to 1 for Kano, Kaduna, and Kogi and 0 otherwise. In this regional analysis we will split the sample for men and women.¹⁸ Finally, we will examine effects based on the chosen value chain and segment of the APPEALS WYEP business (for example, by creating

¹⁷We use the 13-item social desirability scale, which Reynolds (1982) shows is a viable substitute for the 33-item Marlowe-Crowne scale.

¹⁸Gender norms are expected to be, on average, more restrictive for women in the Northern states than in the South.

a dummy variable equal to 1 for the poultry and aquaculture value chains; and 0 for the other crop-related value chains).

3.3 Multiple hypothesis testing

Our survey instrument includes several questions related to a single skill, behavior or dimension, therefore we account for multiple hypothesis testing in the following ways:

1. Index measures: We aggregate the primary outcome variables into an index or composite variable. We use outcome variable indices for several outcome variables including the main business outcomes. These indices combine multiple measures to reduce the total number of tests conducted. We can compute the average standardized effects where we divide each variable by its standard deviation and take the average of these normalized variables [Kling et al. \(2007\)](#) or else, normalize the values and do a principal component analysis with them to construct the index ([Anderson, 2008](#)).

2. Q-values: We can adjust the statistical test for each hypothesis and present sharpened False Discover Rate (FDR) q-values which uses a simple method proposed by [Benjamini et al. \(2006\)](#) to calculate the smallest level of significance at which the null hypothesis would be rejected as described in [Anderson \(2008\)](#). The [Benjamini et al. \(2006\)](#) sharpened two-stage q-values will be presented in our main regression tables in square parentheses below standard errors. The sharpened q-values will be shown for our main results for the outcomes presented in Table 5: P1, P2 and P3.

3.4 Mediation Analysis

We will conduct mediation analysis to understand how much of the total treatment effect is due to: (i) an indirect effect operating through one or several observed mediators, and (ii) the direct effect of the intervention not captured by these observed mediator(s). In the table of outcomes in Appendix A we list several observable variables that are expected potential mechanisms which mediate the training-success relationship. We are particularly interested in using mediation to test whether socioeconomic outcomes can plausibly be explained by changes in the psychological variable constructs and which socioemotional skills explain the most. We will calculate the Average controlled direct effect (ACDE) which refers to the effect that the interventions would have on an outcome if the mediators are fixed at some particular value.

To conduct mediation analysis we will follow the two step procedure as described in [Acharya et al. \(2016\)](#). Step 1: regress the outcome on the mediator, the treatment variable(s), a set of controls, and the interaction between the mediator and all other variables. Obtain the

predicted value of the outcome fixing all mediators to zero. This is the ‘demediated’ outcome. Step 2: regress the demediated outcome on the treatment variable(s). The coefficients from this regression give the estimate of the average conditional direct effect (ACDE). We will use a nonparametric bootstrapping procedure for the standard errors.

The mediation analysis will allow us to investigate the role played by observable indicators in channeling the effects of SES trainings on economic outcomes. SES trainings are meant to improve the participants’ levels of SES. Both interpersonal and intrapersonal skills are deemed important for entering and succeeding in the labor market. We therefore expect both trainings to increase economic outcomes by either making the participant more motivated and skilled, and thus more likely, to start a business or by being better at managing and growing their business. However, different SES trainings may have different outcomes. For example, intrapersonal skills may improve the participants’ ability to concentrate, motivate themselves and overcome obstacles in attaining their business goals. Interpersonal skills may improve the participant’s ability to handle conflict with others, to listen and collaborate in teamwork, to assert themselves, build and grow supportive networks. The different SES trainings are expected to increase economic outcomes through either interpersonal or intrapersonal skills or both.

More schematically, the mediation analysis will test the hypotheses that SES trainings may improve economic outcomes through the following channels: SES trainings may increase SES levels, which in turn may translate into improved (i) business ideas, planning and management practices, (ii) decision-making power and agency, (iii) mental models driving individuals’ motivation and success on the labor market (see Theory of Change in Figure 2). These outcomes may then mediate effects on labor market outcomes (employment, income and profits). The full set of mediators are listed in Appendix Table 6.

We ask does SES training improve economic outcomes, even when business practices are comparable or across entrepreneurs with similar mental models or for women with similar decision-making power and agency? The sequential g-estimator proposed by [Acharya et al. \(2016\)](#) will only produce a consistent estimate of the controlled direct effect of SES training on economic outcomes under the assumptions that (i) there are no omitted variables for the effect of SES training on economic outcomes, conditional on pre-treatment confounders (e.g. variables affecting the treatment, the outcome, and possibly the mediator), (ii) there are no omitted variables for the effect of the mediator on the outcome, conditional on the treatment, pre-treatment confounders, and intermediate confounders. In our context however, we cannot entirely rule out omitted variables such as trust in existing markets, business institutions, entrepreneurs’ motivation and concentration that may bias our mediation analysis. We try to mitigate these risks by analyzing four families of mediators which encompass much

of the theorized mechanisms for potential impact: SES levels (inter- and intra-personal skill indices), business ideas and practices, decision-making power and agency, and mental models. Of course, we cannot assign any causal interpretation to the mediation results, but the analysis provides useful suggestive evidence on which channels might contribute more significantly to any overall effects on economic outcomes. Through mediation analysis we are also unable to disentangle the sequence of influence e.g. whether SES levels change and that leads to behavioral changes.

In Table 6 of potential mediators we have 3 levels of outcome variables: families, categories and outcomes. Within each family of mediators (see M1 to M3) we will build categories by aggregating mediators into indices. For example for Table 6 M1 Psychological Constructs we will create a standardized aggregate index of SES, an aggregate interpersonal skills index, an aggregate intrapersonal skills index, and analyze each skill separately given the core research question of the paper. For Table 6 M2 Business ideas, planning and practices we will analyze the categories of business practices, business ideas and aspirations by creating 3 separate aggregate indices of these mediators. For Table 6 M3 Intrahousehold Decision Making, Time Use and Mental Models we will create aggregate indices across the 5 separate categories. For the main results we will use aggregated indices of the underlying variables in specific categories since there may be some nuance to the specific skills and behaviors that are affected by SES training. However, for the mediation analysis we will use aggregated index of the family of mediators.

3.5 Regression Discontinuity Design (RDD)

Additionally, we use the eligibility criteria cutoff that determined APPEALS WYEP participation to run a regression discontinuity design (RDD) analysis to examine the overall effects of the APPEALS WYEP program. We exploit a sharp regression discontinuity design to estimate the local average treatment effect (LATE) of the APPEALS WYEP program. In the study, participants scoring at least 7 in the screening interview were deemed eligible for the APPEALS WYEP. Applicants with a score lower than 7 are not eligible. The assignment of participants to the APPEALS WYEP program at the 7 score cutoff was largely adhered to by the government implementer. The score could not be manipulated by any of the potential beneficiaries – the applicant either was invited to participate or not. We may consider applicants in states to be more plausibly exogenously sorted around the cutoff of 7. We sample within a smaller window close to the cutoff (applicants scoring 5, 6, 7 or 8). The difference in average outcomes between APPEALS WYEP beneficiaries and non-program is an unbiased estimator of the effect of WYEP, assuming local randomization within the

selected window around the cutoff.

Since the assignment variable is discrete we use a local randomization-based RD design method as proposed by Cattaneo et al. (2015, 2016) rather than conventional nonparametric local polynomial techniques that rely on large-sample approximations. Cattaneo et al. (2016) propose a data-driven, finite-sample, method to find a bandwidth or “window” around the cutoff where the local randomization assumption is assumed to hold. In the following RD approach we use only observations that are between $c - h$ and $c + h$, where h is the bandwidth that determines the size of the neighborhood around the cutoff, c , and the sample for which the empirical RD analysis is conducted. In a smaller bandwidth around the cutoff, units above and below are more likely to be comparable. In the RD approach we test for the assumption of local randomization close to the cutoff point, that is, beneficiaries closest to the 7 score cutoff threshold can be viewed as being part of a local randomized experiment. In the following we present analysis for the bandwidth selection and then present the local average treatment effect (LATE) using only observations close to the cutoff.

In order to determine the bandwidth, h , we utilize the Stata command `rdwinselect`, from the `rdlocrand` package which conducts finite-sample inference in regression discontinuity (RD) designs under a local randomization assumption (Cattaneo et al. 2016). The command uses a data-driven algorithm to find the largest window in which a vector of baseline covariates are found to be balanced within the window based on a joint test using a large-sample approximation of Hotelling’s T-squared statistic. In general we have covariate balance between APPEALS WYEP program and non-program applicants (Table 8). We use a linear local regression function to select a bandwidth for the RD analysis. We include variables collected at screening (pre-intervention) that includes income, business ownership, years of experience and land ownership to determine an optimal bandwidth following Cattaneo et al. 2016. The recommended window from this method is [5; 8] using all observations for which a score was recorded (8,654).¹⁹

To examine treatment effects we estimate the effect of APPEALS WYEP on our main outcomes of interest relative to a comparison group that did not participate in APPEALS WYEP using observations in the data-driven selected windows around the 7 cutoff where local randomization holds.²⁰

¹⁹The optimal bandwidth selection has well known problems that requires a trade-off between precision and bias.

²⁰Note a smaller bandwidth would severely truncate the sample size.

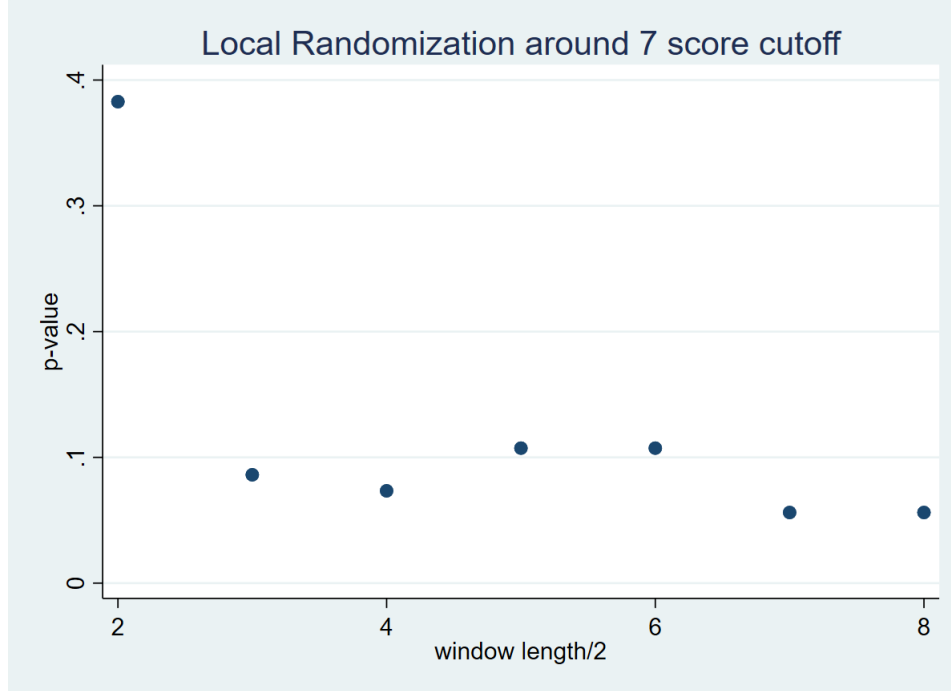


Figure 4: Bandwidth selection around 7 screening score cutoff

3.6 Variations from the intended sample

3.6.1 Threat to internal validity: Survey attrition

Survey attrition in follow-up surveys is a risk where the individuals will be unavailable or not found at the time of the follow-up survey. However, with the applicants being part of the APPEALS WYEP project we expect contact information to be up-to-date which increases the chance of tracking the respondents during follow-up surveys.²¹ Given the research team’s experiences of collecting data in Nigeria and the thorough tracking information collected, we expect the level of attrition to be no more than 15% of the total sample size i.e. a minimum of 85% of the study sample will be resurveyed at follow-up.

If attrition is random, it does not introduce bias but it affects power. If attrition is non-random (i.e. correlated with treatment), it may generate a bias. We will test whether survey attrition is related to treatment status using the following equation:

$$A_i = \beta_0 + \beta_{T1}InterOnlySES_i + \beta_{T2}InterIntraSES_i + \beta_3X'_{0i} + \lambda_s + \varepsilon_i \quad (3)$$

A_i represents whether individual i attrited from the study, $InterOnlySES$ and $InterIntraSES$

²¹We also increased the sample size by 25% to limit the effect that attrition could have on the power of the experiment.

are the treatment dummy variables. Thus, β_{T1} and β_{T2} will reflect whether assignment to treatment T1 and T2 significantly affects the likelihood of an individual attriting from the survey, respectively. X'_{0i} is a set of baseline controls. λ_s are randomization strata fixed effects, and ε_i is the error term.

Additionally, we will test for the joint effect of the coefficient on the interactions to see if attrition is differential across treatment arms. If treatment status does not affect survey attrition at the 5 percent significance level, we will not adjust the estimates for attrition. If treatment status does have a statistically significant effect on survey attrition, we will test the robustness of our results using Lee bounds (Lee, 2009) and Inverse Probability Weighting which, in effect, re-weights the selected sample in order to make it representative of the population.

3.6.2 Threat to internal validity: Partial compliance

In addition to estimating the ITT effects shown in equation 1 and equation 2, we will also estimate the treatment-on-the-treated (TOT) effect using instrumental variable (IV) estimation to estimate the local average treatment effect (LATE). The LATE can be interpreted as the average treatment effect (ATE) for compliers (i.e. those assigned to the treatment group who actually attend the SES trainings and those in the control group who do not receive treatment). We will instrument attendance in the SES training with the random assignment to the treatment groups. This estimate will enable us to control for non-compliance with treatment assignment as not everyone who is offered attends (79% take-up).

$$Y_{1i} = \beta_0 + \beta_{T1} \textit{AttendedInterOnlySES}_i + \beta_{T2} \textit{AttendedInterIntraSES}_i + \beta_3 X'_{0i} + \lambda_s + \varepsilon_i \quad (4)$$

Where *AttendedInterOnlySES* and *AttendedInterIntraSES* are dummy variables indicating whether the respondent attended the SES training sessions, respectively. All other variables are the same as in equation 1. We use the assignment to training as an instrument for attending the SES training sessions. The first stage IV regressions are:

$$\textit{AttendedInterOnlySES}_i = \gamma_0 + \gamma_1 \textit{AssignedInterOnlySES}_i + \gamma_3 X'_{0i} + \lambda_s + \varepsilon_i \quad (5)$$

$$\textit{AttendedInterIntraSES}_i = \gamma_0 + \gamma_1 \textit{AssignedInterIntraSES}_i + \gamma_3 X'_{0i} + \lambda_s + \varepsilon_i \quad (6)$$

We will use the predicted values from these regressions in the second stage IV regression, which is specified in equation 7.

$$Y_{1i} = \beta_0 + \beta_{T1} \widehat{AttendedInterOnlySES}_i + \beta_{T2} \widehat{AttendedInterIntraSES}_i + \beta_3 X'_{0i} + \lambda_s + \varepsilon_i \quad (7)$$

The LATE is only valid under the assumption that the very act of being invited to the SES trainings has no impact on outcomes even if you do not end up attending the training. Since for most respondents this assumption is likely to hold true we will estimate the LATE in addition to the ITT and present the results in an Appendix.

4 Limitations and challenges

4.1 Challenges in the study implementation

The study is designed to provide effects of the individual treatment arms (T1 and T2) compared to the control group (C) and the marginal effect of intrapersonal skills - comparing T2 with T1. However, a few potential circumstances might pose challenges to the study implementation. First, imperfect compliance of project activities with the randomization design may lead to an underestimation of the impact of the intervention. To minimize this risk we conducted an impact evaluation workshop with the implementing partner to help the project team understand the research in order to keep research fidelity. The workshop helped gain buy-in from key decision-makers on the project for the study which ensured project support to help encourage take-up of the SES trainings. Overall, average take-up rates of the SES trainings was close to 80% of those invited which is high relative to other business trainings. Second, spillover or contamination effects may result from sharing the reinforcement messages about the skills learned in the training to other APPEALS WYEP beneficiaries via WhatsApp who are in the control group. This could lead to an underestimation of the impact of the intervention. We mitigate this potential issue by sending placebo messages to the control group containing information about COVID-19. Third, attrition in follow-up surveys - there is a risk that individuals will drop out of the survey and will be unavailable for the follow-up survey, and this will differ by treatment and control. However, with the study sample all being part of the APPEALS WYEP project we expect contact information to be good which increases the chance of finding the respondents during follow-up survey rounds. We have also increased the sample size by 25 percent, to limit the effect that attrition could have on the power of the experiment. Finally, poor data quality that results from poorly implemented data collection surveys can threaten the evaluation's ability to measure impact accurately. The research team will mitigate these risks by ensuring a survey firm with good experience and staff is engaged in the data collection and will provide careful supervision of all data collection activities. Back-checks on certain sections will also be carried out.

4.2 Addressing Possible Spillover or Contamination Effects

Spillovers across treatment arms are a risk with individual-level randomization and classroom trainings. However, they are expected to be unlikely in this study setting through either word-of-mouth or through direct observation (e.g. knowledge, network or social spillovers) for several reasons:

- **Wide geographic spread of beneficiaries across six states in Nigeria:** Our experiment unfolds across 6 distinct, mostly non-contiguous states across Nigeria (3 in the North and 3 in the South). The program is restricted to individuals based on specific eligibility criteria within the 6 states, rather than being limited to within specific communities. As a result, individuals in our study sample live relatively far apart from each other and have limited chances of interaction. We estimate that no more than 0.5% of the eligible population of a given state received the APPEALS WYEP program, or 0.03% of the total population of a given state. In Table AX in Appendix we present the assumptions used in the calculation of these estimates. We will use GPS coordinates collected during the follow-up survey to corroborate this wide geospatial coverage.
- **Sample selection process** suggests that the APPEALS WYEP applicants were likely to be relatively scattered wide geographically and are unlikely to build connections across the different trainings.
- **Separate locations and timings of training limits the potential for cross-treatment contamination:** the two SES trainings (inter-only and combination) were conducted in classroom settings in training institutes at different points in time (not simultaneously). In other words, the relatively modest scale and sequential timing of both interventions leave little room for individuals participating in different SES trainings to cross paths and exchange information on the differing content of their trainings.
- **Underlying principles of the SES training facilitation:** Rather than simply providing instruction, the SES trainings were designed to incorporate a facilitation and reflection approach. Training facilitators were trained to model positive behavior techniques and facilitate reflection on experiences, reinforce lessons, and/or help the beneficiaries to define meaning in an experience. This is not a type of training that lends itself to individuals easily passing skills to one another and the skills are difficult to be imitated. We concede that there is a small risk that beneficiaries could practice their skills with individuals assigned to other treatment arms. However, the incidence of this is expected to be rare. If spillovers occur between treated and control households then our results are likely to be biased downward.
- **Joint business partnerships across treatment arms a low risk:** APPEALS WYEP beneficiaries were not explicitly encouraged or incentivized to set up joint business partnerships (that could potentially lead to greater information sharing) and in the follow-up survey, we will check the prevalence of joint business partnerships

across the interventions.

While spillovers are expected to be limited in this setting for the reasons described above, in the follow-up survey we will collect information on exposure to the various APPEALS WYEP interventions as well as GPS data for all individuals in our sample. We will check for any correlation in socioemotional skills of the control group and distance to a nearest treatment household. This will help us rule out any direct contamination and better assess risks of social externalities. Given that we did not collect baseline links between treatment and control identifiers, an analysis on nearest neighbors would be very difficult to use to shed light on spillovers in this context.

5 Administrative information

5.1 Funding

This work was supported by the Wellspring Foundation and the World Bank Umbrella Fund for Gender Equality (UFGE) and other World Bank sources.

5.2 Ethics approval

All necessary ethics approvals are in place. Institutional Review Board (IRB) approval was obtained from Health Media Lab (HML) IRB Research and Ethics, reference 681IPA19.

5.3 Declaration of interest

No conflicts of interest to declare.

5.4 Acknowledgments

We thank Wale Ogunleye and Emeka Nweke for excellent research assistance and field management. We thank the editor and 2 anonymous reviewers at the Journal of Development Economics Pre-Results Review for helpful comments.

References

- Acharya, A., Blackwell, M., , and Sen, M. (2016). Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review*, 110(3):512–529.
- Aiken, L. S., West, S. G., and Reno, R. R. (1991). Multiple regression: Testing and interpreting interactions. *Sage*.
- Almlund, M., Duckworth, A. L., Heckman, J., and Kautz, T. (2011). Personality psychology and economics. *Handbook of the Economics of Education*, 4:1–181.
- Anderson, M. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484):1481–1495.
- Aranda-Jan, C., Mohutsiwa-Dibe, N., and Loukanova, S. (2014). Systematic review on what works, what does not work and why of implementation of mobile health (mhealth) projects in africa. *BMC Public Health*, 14(188):1–15.
- Ashraf, N., Bau, N., Low, C., and McGinn, K. (2018). Negotiating a better future: How interpersonal skills facilitate inter-generational investment. *The Quarterly Journal of Economics*, 135.
- Belloni, A., Chernozhukov, V., and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608–650.
- Benjamini, Y., Krieger, A. M., and Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3):491–507.
- Borghans, L., Duckworth, A., Heckman, J., and Weel, B. (2008). The economics and psychology of personal traits. *The Journal of Human Resources*, 43.
- Bowles, H. R., Babcock, L., and Lai, L. (2005). It Depends Who Is Asking and Who You Ask: Social Incentives for Sex Differences in the Propensity to Initiate Negotiation. Working paper series, Harvard University, John F. Kennedy School of Government.
- Campos, F., Frese, M., Goldstein, M., Iacovone, L., Johnson, H., and McKenzie, D. (2017). Teaching personal initiative beats traditional training in boosting small business in west africa. *Science*, 357:1287–1290.

- Campos, F., Goldstein, M., McGorman, L., Muñoz Boudet, A. M., and Pimhidzai, O. (2015). Breaking the metal ceiling: female entrepreneurs who succeed in male-dominated sectors. Policy Research Working Paper Series 7503, The World Bank.
- Chioda, L., Contreras-Loya, D., Gertler, P., and Carney, D. (2021). Making entrepreneurs: Returns to training youth in hard versus soft business skills. Working Paper 28845, National Bureau of Economic Research.
- Correll, S. J. (2004). Constraints into preferences: Gender, status, and emerging career aspirations. *American Sociological Review*, 69(1):93–113.
- Croppenstedt, A., Goldstein, M., and Rosas, N. (2013). Gender and agriculture: Inefficiencies, segregation, and low productivity traps. *World Bank Working Paper Series*.
- Deming, D. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, 132(4):1593–1640.
- Duflo, E. (2003). Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in south africa. *The World Bank Economic Review*, 17:1–25.
- Edin, P.-A., Fredriksson, P., Nybom, M., and Öckert, B. (2022). The rising return to noncognitive skill. *American Economic Journal: Applied Economics*, 14(2):78–100.
- Friedson-Ridenour, S. and Pierotti, R. S. (2019). Competing priorities: Women’s microenterprises and household relationships. *World Development*, 121:53–62.
- Gielnik, M. M., Frese, M., Kahara-Kawuki, A., Katono, I. W., Kyejjusa, S., Munene, J., and Dlugosch, T. J. (2012). Action and action-regulation in entrepreneurship: Evaluating a student training for promoting entrepreneurship. *Academy of Management Learning and Education*, 14:69–94.
- Goldstein, M. and Udry, C. (2008). The profits of power: Land rights and agricultural investment in ghana. *Journal of Political Economy*, 116(6):981–1022.
- Grabowska, M. (2020). Reminders to promote behaviours. *Oxford CSAE COVID-19 policy brief*.
- Heckman, J. (1995). Lessons from the bell curve. *Journal of Political Economy*, 103(5):1091–1120.
- Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4):451–464.

- Heckman, J. J., Stixrud, J., and Urzua, S. (2006). Effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.
- Jacob, B. A. (2002). Where the boys aren't: non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education Review*, 21(6):589–598.
- Kling, J., Liebman, J., and Katz, L. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.
- Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3):1071–1102.
- Lindqvist, E. and Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the swedish enlistment. *American Economic Journal: Applied Economics*, 3(1):101–128.
- Lippman, L. H., Ryberg, R., Carney, R., and Moore, K. A. (2015). Workforce connections: Key “soft skills” that foster youth workforce success: toward a consensus across fields. *Washington, DC: Child Trends*.
- Lubega, P., Nakakawa, F., Narciso, G., Newman, C., Kaaya, A., Kityo, C., and Tumuhimbise, G. (2021). Body and mind: Experimental evidence from women living with hiv. *Journal of Development Economics*, 150:102613.
- McKenzie, D. (2012). Beyond baseline and follow-up: The case for more t in experiments. *Journal of Development Economics*, 99(2):210–221.
- McKenzie, D. (2020). Small business training to improve management practices in developing countries: Re-assessing the evidence for “training doesn’t work”. *Oxford Review of Economic Policy*.
- McKenzie, D. and Woodruff, C. (2013). What are we learning from business training and entrepreneurship evaluations around the developing world? *The World Bank Research Observer*, 29(1):48–82.
- Montalvao, J., Frese, M., Goldstein, M., and Kilic, T. (2017). Soft skills for hard constraints evidence from high-achieving female farmers. *World Bank Working Paper Series*.
- Mueller, G. and Plug, E. (2006). Estimating the effect of personality on male and female earnings. *Industrial and Labor Relations Review*, 60(1):3–22.

- Orkin, K. (2020). Behavioural insights for covid-19 communications. *Oxford CSAE COVID-19 policy brief*.
- O’Sullivan, M., Rao, A., Banerjee, R., Gulati, K., and Vinez, M. (2014). Levelling the field: Improving opportunities for women farmers in africa. *World Bank*.
- Qian, N. (2008). Missing women and the price of tea in china: The effect of sex-specific earnings on sex imbalance. *Quarterly Journal of Economics*, 123(3):1251–1285.
- Reynolds, W. M. (1982). Development of reliable and valid short forms of the marlowe-crowne social desirability scale. *Journal of Clinical Psychology*, 38(1):119–125.
- Rudman, L. A. and Fairchild, K. (2004). Reactions to counterstereotypic behavior: the role of backlash in cultural stereotype maintenance. *Journal of personality and social psychology*, 87 2:157–76.
- Stark, L., Asghar, K., Seff, I., Cislighi, B., Yu, G., Gessesse, T. T., Eoomkham, J., Baysa, A. A., and Falb, K. (2018). How gender- and violence-related norms affect self-esteem among adolescent refugee girls living in ethiopia. *Global Ment Health*, 5:e2.
- Thomas, D. (1990). Intra-household resource allocation: An inferential approach. *The Journal of Human Resources*, 25(4):635.
- Tomlinson, M., Rotheram-Borus, M. J., Swartz, L., and Tsai, A. C. (2013). Scaling up mhealth: Where is the evidence? *PLoS Medicine*, 10(2):1–5.

A Appendix

A.1 Randomization Balance Table

Table 2 shows that the treatment assignment was balanced across C, T1 and T2 on observable characteristics, except for a few variables such as: the number of employees in the business, generalized self-efficacy and attitudes and norms towards the male breadwinner status. The normalized differences suggest that any statistically significant differences are small in magnitude. Importantly, the F-test for joint significant of covariates by treatment assignment is not significant (see bottom of Table 2) suggesting the randomization created similar groups based on observables.

A.2 Correlates of Training Take-Up

In Table 3 we show the correlates of take-up by estimating differences in baseline characteristics between those who took up the training and those who did not. Table 3 column 2 indicates the characteristics of those who were invited but did not attend the training, and column 3 indicates those who did attend; and in (2)-(3) we show the test for a difference in the means across these groups. On average, those who attend the training are slightly older, less likely to be married, less likely to be active on their own farm in the past 30days, and have lower average household income, relative to those who did not attend a training.

A.3 Power Calculations MDEs for Key Outcomes

In Table 4 we present the minimum detectable effect (MDE) sizes using data from previous impact evaluation studies. Our study sample size was selected assuming only 70% training take-up. However, we successfully achieved 80% take-up.

In the following we present the full set of outcomes, along with descriptions and details on construction are in Appendix section A4 and A5.

Table 2: Randomization Balance Table

Variable	(1)	(2)	(3)	(4)	(5)	(6)	t-test Difference						Normalized Differences					
	Control Male	Interpersonal Only Male	Combination inter_intra Male	Control Female	Interpersonal Only Female	Combination inter_intra Female	(1)-(2)	(1)-(3)	(2)-(3)	(4)-(5)	(4)-(6)	(5)-(6)	(1)-(2)	(1)-(3)	(2)-(3)	(4)-(5)	(4)-(6)	(5)-(6)
Age (Number)	31.50	31.27	31.35	34.61	34.25	34.56	0.22	0.15	-0.08	0.36	0.05	-0.31	0.04	0.03	-0.01	0.04	0.01	-0.03
	[0.19]	[0.20]	[0.19]	[0.35]	[0.34]	[0.33]												
Married (Yes = 1)	0.56	0.53	0.54	0.67	0.65	0.69	0.03	0.03	-0.01	0.03	-0.02	-0.04*	0.06	0.05	-0.01	0.06	-0.04	-0.09
	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]												
Worked for pay in the last week (Yes = 1)	0.63	0.63	0.64	0.54	0.51	0.49	0.00	-0.01	-0.01	0.03	0.05*	0.02	0.00	-0.03	-0.03	0.06	0.10	0.04
	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]												
Owens a Business (Yes = 1)	0.78	0.75	0.77	0.78	0.78	0.79	0.02	0.01	-0.01	0.01	-0.01	-0.02	0.05	0.02	-0.03	0.01	-0.03	-0.04
	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.01]												
Average Monthly Business Profits unconditional (NGN)	34518.83	31446.35	29925.07	17160.35	17788.90	20100.26	3072.48	4593.76	1521.27	-628.55	-2939.91	-2311.36	0.04	0.07	0.03	-0.02	-0.08	-0.06
	[2791.98]	[2340.81]	[2009.93]	[1078.19]	[1147.50]	[1543.45]												
Number of Employees in Business unconditional (Number)	1.37	1.49	1.25	0.60	0.67	0.75	-0.12	0.13	0.25**	-0.07	-0.16*	-0.08	-0.05	0.06	0.10	-0.04	-0.09	-0.05
	[0.08]	[0.09]	[0.08]	[0.06]	[0.06]	[0.07]												
Generalized self-efficacy (1-5)	4.20	4.18	4.20	4.14	4.16	4.12	0.02	0.00	-0.01	-0.02	0.02	0.04*	0.04	0.01	-0.03	-0.05	0.04	0.09
	[0.02]	[0.02]	[0.02]	[0.01]	[0.02]	[0.02]												
Participated in Own Farm Activity in past 30days?	0.55	0.58	0.57	0.39	0.39	0.40	-0.03	-0.02	0.01	0.00	-0.01	-0.01	-0.06	-0.05	0.02	0.00	-0.02	-0.02
	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]												
Participated in Non-Farm Enterprise Activity in past 30days?	0.53	0.54	0.53	0.54	0.54	0.53	-0.02	-0.00	0.01	-0.01	0.00	0.01	-0.03	-0.00	0.03	-0.01	0.01	0.02
	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]												
Participated in Livestock Activity in past 30days?	0.36	0.38	0.37	0.33	0.31	0.31	-0.02	-0.02	0.00	0.02	0.02	0.01	-0.05	-0.04	0.01	0.03	0.05	0.01
	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]												
Participated in Wage Employment Activity in past 30days?	0.13	0.11	0.14	0.07	0.09	0.08	0.02	-0.01	-0.03*	-0.02	-0.01	0.01	0.06	-0.03	-0.09	-0.08	-0.04	0.04
	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]												
Average monthly household income (NGN)	65234.03	61281.33	61167.87	58128.35	54106.39	56990.53	3952.70	4066.15	113.45	4021.96	1137.82	-2884.14	0.03	0.03	0.00	0.03	0.01	-0.03
	[5624.59]	[3643.11]	[3210.99]	[5413.50]	[3297.45]	[4558.77]												
Hours spent on work for pay in a day (number)	7.03	7.15	7.00	5.82	5.69	5.77	-0.12	0.03	0.15	0.13	0.05	-0.08	-0.04	0.01	0.05	0.04	0.02	-0.02
	[0.11]	[0.11]	[0.11]	[0.12]	[0.12]	[0.11]												
Hours spent on care in a day (number)	3.93	3.78	4.04	5.29	5.44	5.31	0.15	-0.11	-0.26**	-0.14	-0.01	0.13	0.06	-0.04	-0.10	-0.04	-0.00	0.04
	[0.10]	[0.09]	[0.10]	[0.13]	[0.13]	[0.12]												
Hours spent on household chores in a day (number)	1.99	1.95	2.05	3.13	3.17	3.20	0.05	-0.06	-0.10	-0.03	-0.06	-0.03	0.04	-0.04	-0.08	-0.02	-0.04	-0.02
	[0.05]	[0.05]	[0.05]	[0.06]	[0.07]	[0.06]												
Personally agree with the male breadwinner norm? (Agree =1)	0.37	0.41	0.38	0.30	0.29	0.26	-0.04	-0.01	0.02	0.02	0.04*	0.02	-0.08	-0.03	0.05	0.04	0.09	0.05
	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]												
Perceived male breadwinner norm (Out of 10 neighbors, agree..)	5.17	5.44	5.33	4.67	4.57	4.39	-0.27*	-0.16	0.11	0.10	0.28**	0.18	-0.10	-0.06	0.04	0.04	0.10	0.07
	[0.10]	[0.10]	[0.10]	[0.10]	[0.10]	[0.10]												
Number of Observations	750	750	750	750	750	750												
F-test of joint significance (p-value)							0.72	0.99	0.64	0.95	0.55	0.83						

Notes: ***, **, and * indicate significance at the 1, 5, and 10 percent critical level. The value displayed for t-tests are the differences in the means across the groups. The value displayed for F-tests are p-values. All variables are measured during the baseline.

Table 3: Training Take-up Correlates

Variable	(1)	(2)	(3)	t-test Difference			Normalized difference		
	Control Group (Not Invited to a Training)	Treatment and Did Not Attend Training	Treatment and Attended the Training	(1)-(2)	(1)-(3)	(2)-(3)	(1)-(2)	(1)-(3)	(2)-(3)
Gender of respondent (Female = 1 / Male = 0)	0.500 [0.013]	0.502 [0.020]	0.500 [0.010]	-0.002	0.000	0.002	-0.003	0.001	0.004
Age (Number)	33.054 [0.200]	31.456 [0.289]	33.231 [0.158]	1.598***	-0.177	-1.776***	0.209	-0.023	-0.232
Married (Yes =1)	0.618 [0.013]	0.569 [0.020]	0.610 [0.010]	0.049**	0.008	-0.041*	0.100	0.016	-0.084
Worked for pay in the last week (Yes = 1)	0.585 [0.013]	0.552 [0.020]	0.573 [0.010]	0.032	0.012	-0.021	0.065	0.024	-0.042
Business (Yes =1)	0.779 [0.011]	0.768 [0.017]	0.775 [0.009]	0.011	0.005	-0.006	0.027	0.011	-0.015
Average Monthly Profits Unconditional (Naira)	25839.589 [1512.672]	26703.746 [1889.867]	24313.112 [1043.482]	-864.157	1526.476	2390.634	-0.016	0.028	0.048
Number of Employees	0.987 [0.051]	1.056 [0.085]	1.038 [0.043]	-0.069	-0.051	0.018	-0.034	-0.025	0.008
Generalized self-efficacy score	4.168 [0.011]	4.169 [0.018]	4.162 [0.009]	-0.001	0.006	0.007	-0.003	0.014	0.016
Farming activities on your own farm (Yes = 1)	0.469 [0.013]	0.424 [0.020]	0.502 [0.010]	0.045*	-0.033**	-0.078***	0.090	-0.066	-0.155
Farming activities on someone else's farm (Yes = 1)	0.127 [0.009]	0.118 [0.013]	0.129 [0.007]	0.009	-0.002	-0.011	0.027	-0.007	-0.034
Non-farm activities (Yes = 1)	0.531 [0.013]	0.528 [0.020]	0.538 [0.010]	0.003	-0.007	-0.010	0.006	-0.015	-0.020
Livestock rearing (Yes = 1)	0.343 [0.012]	0.358 [0.019]	0.340 [0.010]	-0.014	0.003	0.018	-0.030	0.007	0.037
Wage Employment (Yes = 1)	0.099 [0.008]	0.089 [0.011]	0.109 [0.006]	0.010	-0.010	-0.020	0.035	-0.032	-0.066
Household Income (Naira)	61681.186 [3903.042]	68528.084 [4634.560]	55693.729 [2000.579]	-6846.898	5987.457	12834.355***	-0.048	0.049	0.126
Number of observations	1500	630	2370						

Note: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table 4: Minimum Detectable Effect (85% power, 5% significance level)

Ratio of SD to Mean	Effect Size in standard deviations	Variable Category	Data sources	Sample size per treatment arm	Adjustment for take-up of 80%	Adjustment for take-up of 70%	Adjustment for take-up of 60%
1.2	0.2	Income/ Profits/Revenue/ Consumption	Nigeria LSMS 2015	363	567	741	1,008
1	0.2		Nigeria YouWIN!	252	394	516	700
0.8	0.2		Personal Initiative, Togo	161	252	329	447
0.3	0.05	Socioemotional skills	ELA, Uganda	363	567	741	1,008
0.25	0.05		Sisters of Success, Liberia	252	394	516	700
0.2	0.05		Personal Initiative, Togo	161	252	329	447

A.4 Definitions of Socioemotional Skills and Categorization

	Intrapersonal	Interpersonal
Awareness	<p>Emotional Awareness: identifying and accepting one's emotions</p> <p>Self Awareness: identifying and interpreting one's own thoughts and behaviors and to evaluate one's strengths and weakness</p>	<p>Listening: attending to what other people are saying, taking time to understand other's point of view, asking clarifying questions and not interrupting at inappropriate times</p> <p>Empathy: understand another's viewpoint or thoughts and have emotional concern for another's situation or experience</p>
Management	<p>Emotional Regulation: maintaining or changing one's own emotions by controlling one's thoughts and behavioral responses</p> <p>Self Control: focusing one's attention, staying on task, breaking habits, restraining impulses and keeping good self-discipline</p> <p>Personal Initiative: developing long-term goal, to seek opportunities to improve one's self and to be motivated to put these plans and goals into action</p> <p>Perseverance: sustaining effort despite setbacks</p> <p>Problem Solving: approaching a problem by gathering information, generating a number of solutions and evaluating the consequences of these solutions before acting</p>	<p>Expressiveness: explaining ideas in a way that others will understand and openly expressing one's opinion</p> <p>Interpersonal Relatedness: taking actions intended to build trust and benefit others, initiating and maintaining relationships and being respectful, encouraging and caring for others</p> <p>Interpersonal Influence: communicating in a manner that changes other's perspectives and adapting one's behavior in situationally appropriate ways to influence others</p> <p>Negotiation: identifying own and others interests to brainstorm mutually beneficial solutions and maximize own self interest</p> <p>Collaboration: considering different perspective, listening and communicating in groups of two or more people, identifying situations involving group problem-solving and decision-making, and organizing and coordinating team members to create shared plans and goals</p>

Figure 5: Definitions of Intrapersonal and Interpersonal Socioemotional Skills

A5 Description of Outcomes and Mediators

P1: Economic and Business Outcomes

P2. Agricultural outcomes

P3. Women's Empowerment

M1. Socioemotional skills

M2. Business ideas, planning, and practices

M3. Intrahousehold Decision Making and Mental Models

S1. Financial variables

S2. Marital Satisfaction and Intimate Partner Violence

S3: Asset Investments

S4. Consumption

S5. Subjective Wellbeing

Table 5: Primary Outcomes

Category	Outcome	Description
P1: Economic and Business Outcomes		
<i>Business Performance</i>	1 Business activity in the past 30 days	Binary variable that indicates if the respondent conducted any business activity in the past 30 days? Number of hours worked in the business in the past week.
	2 Enterprise Profits and Revenues	Annual and monthly recall periods to be asked directly. Include the inverse hyperbolic sine (IHS) transformation of profits and sales $\log(y+(y2 +1)1/2)$ – which can deal with zero values. Profits monthly and annual recall and revenues only asked annually.
	3 Business Costs	Measured as a summation of individual costs and a total monthly measure.
	4 Value of capital investment	Value of owned and rented machinery, inventory stocks, and raw materials.
	5 Number of Employees	Number of workers, number of hours worked. Number of hours business open per week.
	6 Business Survival rate	The business owner still has at least one business?
<i>New business</i>	7 New business activity	Any business started in addition to APPEALS WYEP?
<i>Household wealth</i>	8 Household income	Average monthly income earned by all members of the household
	9 Share of income contributed to the household	Proportion of income contributed by the respondent
P2. Agricultural outcomes		
<i>Agricultural and Livestock Activity and Income (Yields and Inputs)</i>	1 Farming activity in the past 30 days	Binary variable that indicates if the individual within the household practice any farming activity (including livestock, aquaculture or poultry) in the past 30 days?
	2 Total hired labor on plots in last agricultural season	How many hours did your household hire men, women and children for activities such as land preparation, planting, ridging, weeding and fertilizing?
	3 Inputs expenditures	Total expenses on inputs (fertilizer) and seeds
	4 Crop variety	Number of crops. Shift to any cash crops.
	5 Agricultural Revenues and Income	Agricultural revenue and income (in the past 12 months) – PPP adjusted US Dollar value of all revenue from crops, land rentals, and other earnings. (For income minus land rental costs and input expenditures.)
	6 Livestock Revenues and Income	Livestock revenue and income (in the past 12 months) – PPP adjusted US Dollar value of all revenue from milk, egg, bird and meat sales. (For income minus costs.)
P3. Women's Empowerment		
<i>Modified project-level Women's Empowerment in Agriculture Index (Pro-WEAI)</i>	1 Women's empowerment index based on modified Pro-WEAI indicators	Index of potential indicators from the Pro-WEAI guide: Self-efficacy Attitudes about intimate partner violence against women Respect among household members (optional) Input in productive decisions Ownership of land and other assets Access to and decisions on financial services Control over use of income Work balance Visiting important locations Group membership <i>Modified version of the project use WEAI index will be collected and used to measure women's economic empowerment.</i>
<i>Decision Making Power</i>	2 Index of decision-making power over livelihood	Standardized index of all decisions collected at endline.
<i>Access to Financial and Productive Resources</i>	3 Index of access to resources	Index of whether the respondent has access to ... household labor (whether hh/spouse helps with agricultural production), financial services and productive capital

Table 6: Potential Mediators

Category	Outcome	Description
MI. Psychological Constructs		
Socioemotional Skills	1 Generalized Self-Efficacy Scale (GSE)	Standardized Index: Higher score indicating higher self-efficacy (Schwarzer, R, 1995) Sum of responses from 1 to 4 for all 10 statements (GSE index will be standardized) 1. I can always solve difficult problems if I try hard enough. 2. If someone is against me, I can find means and ways to get what I want. 3. It is easy for me to stick to my aims and accomplish my goals. 4. I am confident that I could deal appropriately with unexpected events. 5. Thanks to my skillful and creative thinking, I know how to handle unforeseen situations. 6. I solve most problems if I put in the necessary effort. 7. I can remain calm when I am facing difficulties because I can rely on my abilities to cope. 8. When I am faced/confronted with a problem, I can usually find several solutions. 9. If I am in trouble, I can usually think of a solution. 10. I can usually handle whatever comes my way. Possible responses: Not at all true (1), Hardly true (2), Moderately true (3), Exactly true (4)
	2 Emotional Awareness	Defn: identifying and accepting one's emotions <i>Self Report Scale (Original items & Adapted portion of Cognitive and Affective Mindfulness Scale-Revised (Feldman et al., 2007))</i> 1. I know why my feelings change from one moment to another. 2. I recognize what I am feeling. 3. I can usually describe what I am feeling at the moment in great detail. 4. When I am stressed, I take time to reflect on the true cause of my feelings 5. I notice when my emotions cause physical sensations in my body 6. I try to notice my thoughts without judging them. 7. I am able to accept the thoughts and feelings I have. <i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale</i>
	3 Self Awareness	Defn: identifying and interpreting one's own thoughts and behaviors and to evaluate one's strengths and weaknesses <i>Self Report Scale (Original items)</i> 1. I understand my own behaviors 2. I am aware of my thoughts 3. I monitor my thinking to ensure it is accurate 4. I know the skills I have that other people do not have 5. I assess my strengths and weaknesses in new situations. 6. I examine my own abilities to better understand myself. 7. I review how I am thinking when I make a mistake 8. I have a clear sense of who I am <i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale</i>
	4 Emotional Regulation	Defn: maintaining or changing one's own emotions by controlling one's thoughts and behavioral response <i>Self Report Scale (Adapted portions of Schutte Self-Report Emotional Intelligence Test (Schutte et al., 1998), Emotion Regulation Questionnaire (Gross & John, 2003), Emotion Regulation of Self and Others (Niven et al., 2011), mapped portion of GSE, Emotional Control Sub-Scale of Psychological Style assessment by the Australian Temperament Project)</i> 1. When I feel nervous, I know what to do to feel more relaxed. 2. When I feel sad, I know how to take my mind off my problems. 3. When I am angry at someone, I can calm down before talking to them. 4. When I'm faced with a stressful situation, I make myself think about it in a way that helps me stay calm. 5. I control my emotions by changing the way I think about my situation. 6. When I want to feel better, I ask others for help. 7. When I am upset, I ask others for emotional support 8. When I want to feel better, I share my feelings with someone I trust 9. When I want to feel better, I do something I enjoy. 10. When I want to feel hopeful, I change my thinking so I am more positive. 11. When I am facing difficulties, I can remain calm because I rely on my abilities to cope. <i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale</i>
	5 Self Control	Defn: focusing one's attention, staying on task, breaking habits, restraining impulses and keeping good self-discipline <i>Self Report Scale (Adapted portions of Brief Self Control Scale (Tangney et al., 2004) & Cognitive and Affective Mindfulness Scale-Revised (Feldman et al., 2007))</i> 1. Pleasure and fun sometimes keep me from getting work done. 2. I do things that feel good in the moment, but I will regret later on. 3. Sometimes I can't stop myself from doing something, even if I know it is wrong. 4. I often act without thinking through all the alternatives. 5. I have trouble concentrating. 6. I am easily distracted. 7. Sometimes I cannot focus on the present moment 8. It is difficult to pay close attention to one thing for a long period of time. <i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale + Continuous Performance Test</i>

Table 6 cont...: Potential Mediators

Category	Outcome	Description
MI. Psychological Constructs	6 Perseverance	<p>Defn: sustaining effort despite setbacks</p> <p><i>Self Report Scale (Adapted from Short Grit Scale (Duckworth & Quinn, 2009) & mapped GSE items)</i></p> <ol style="list-style-type: none"> 1. I finish whatever I begin. 2. Setbacks don't discourage me. 3. I am diligent. 4. When work is difficult, I keep up my effort. 5. If someone is against me, I keep working to fix the problem until I get what I want. 6. It is easy for me to put in the effort necessary to achieve my goals. 7. I am confident that I could deal appropriately with unexpected events. 8. Even when goals require years of patience, I keep working on them. <p><i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale + Triangle test adapted from Alan et al., 2019</i></p>
	7 Personal Initiative	<p>Defn: developing long-term goals, to seek opportunities to improve one's self and to be motivated to put these plans and goals into action</p> <p><i>Self Report Scale (Adapted (Frese et al., 1997) & Personal Growth Initiative Scale-II (Robitschek et al., 2012))</i></p> <ol style="list-style-type: none"> 1. I actively tackle problems. 2. Whenever something goes wrong, I search for a solution immediately. 3. Whenever there is a chance to get actively involved, I take it. 4. I take action immediately even when others don't. 5. I am quick to take advantage of opportunities to reach my goals. 6. I am particularly good at making my ideas a reality. 7. I seek opportunities to learn more. 8. I am constantly trying to improve myself by learning new skills. 9. I look for opportunities to improve myself personally and professionally. 10. If I get feedback on ways I can improve, I immediately try to do better. 11. I watch people who are successful at work and try to adopt their approaches. <p><i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale</i></p>
	8 Problem Solving	<p>Defn: approaching a problem by gathering information, generating a number of solutions and evaluating the consequences of these solutions before acting</p> <p><i>Self Report Scale (Original items & relevant GSE items)</i></p> <ol style="list-style-type: none"> 1. If I am in trouble, I can usually think of a solution. 2. I solve most problems if I put in the necessary effort. 3. I can always solve difficult problems if I try hard enough. 4. I plan tasks carefully 5. When making a decision, I analyze my options and their consequences before I act. 6. When making a decision, I look for as much information as I can before I decide what to do. 7. I compare all my options before making a decision. 8. When I have a problem, I can find several ways to solve it. 9. If someone needs input on a problem, I can come up with many suggestions 10. If my first solution does not work, I can come up with another way to solve my problem. 11. I can find creative solutions to unplanned problems. 12. I come up with solutions that are unique. 13. I develop new solutions to problems that surprise other people. <p><i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale</i></p>
	9 Listening	<p>Defn: attending to what other people are saying, taking time to understand other's point of view, asking clarifying questions and not interrupting at inappropriate times</p> <p><i>Self Report Scale (Original items & Adapted portions of Active-Empathetic Listening Scale (Bodie, 2011), Active Listening Attitude Scale (Kourmoussi et al., 2017; Mishima et al., 2000), & Competencias Personales y Sociales (Brea, n.d.))</i></p> <ol style="list-style-type: none"> 1. When I am listening to someone, I wait to talk until the other person finishes talking. 2. When I am listening to someone, I will not interrupt the other person even if I have something important to say. 3. Even when I want to share my opinion, I can listen to others' opinions first. 4. I ask questions to understand the other person's position on an issue. 5. When I am listening to someone, I make sure they know I am interested in what they are saying. 6. When I am listening to someone, I show them that I am open to their ideas. 7. When I am listening to someone, I ask questions that show my understanding of what they are saying. <p><i>Alternative Measure: Listening comprehension questions + Enumerator observations of active listening</i></p>
	10 Empathy	<p>Defn: understand another's viewpoint or thoughts and have emotional concern for another's situation or experience</p> <p><i>Self Report Scale (Original items & Adapted portions of Interpersonal Reactivity Index (Davis, 1983) & Multi-Dimensional Emotional Empathy Scale (Caruso & Mayer, 1998b))</i></p> <ol style="list-style-type: none"> 1. I try to understand the perspective of others before making a decision that affects them. 2. When I'm upset at someone, I usually try to imagine myself in their situation to better understand them. 3. Before judging somebody, I try to imagine how I would feel if I were in their place. 4. I ask questions to understand the other person's position on a given issue. 5. I always try to understand the feelings of people I trust 6. If I see someone is hurt, I feel upset 7. I feel good when I help someone in need 8. If others are happy, I feel good 9. I can quickly sense when someone in the group is uncomfortable <p><i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale</i></p>

Socioemotional Skills

Table 6 cont...: Potential Mediators

Category	Outcome	Description
M1. Psychological Constructs		
Socioemotional Skills	11 Expressiveness	<p>Defn: explaining ideas in a way that others will understand and openly expressing one's opinion</p> <p><i>Self Report Scale (Original items)</i></p> <ol style="list-style-type: none"> 1. I can explain complicated ideas in a clear way. 2. Others understand my thoughts when I express them 3. I find it easy to explain my perspective to others 4. I communicate in a way that others will understand. 5. I share my opinion with others without hesitation 6. I ask for what I need when I need it. 7. I think it's good to ask for what I want 8. I share my thoughts even if others do not agree with them. 9. I directly communicate what I need from others. <p><i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale</i></p>
	12 Interpersonal Relatedness	<p>Defn: taking actions intended to build trust and benefit others, initiating and maintaining relationships and being respectful, encouraging and caring for others</p> <p><i>Self Report Scale (Original items)</i></p> <ol style="list-style-type: none"> 1. When others are sad, I try to comfort them. 2. I listen patiently when people tell me their problems. 3. When I see that someone is going through a difficult time, I help out the best I can. 4. I give my friends and family encouragement when they need it. 5. I support my friends and family by praising them. 6. I am able to forgive my friends and family if they do something that frustrates me. 7. I stay connected with people who are important to me. 8. I am good at building relationships with people I don't know. 9. I know how to make a good impression 10. I find it easy to get people to trust me 11. If I'm looking for help, I'll ask my contacts if they can recommend someone. 12. I am able to introduce myself to people I don't know well (e.g., strangers, new or unfamiliar people). 13. I am good at getting to know people. <p><i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale</i></p>
	13 Interpersonal Influence	<p>Defn: communicating in a manner that changes other's perspectives and adapting one's behavior in situationally appropriate ways to influence others</p> <p><i>Self Report Scale (Original items)</i></p> <ol style="list-style-type: none"> 1. I can communicate my ideas in a way that convinces people to agree with me. 2. Other people do what I ask them to do. 3. When someone disagrees with me, I know how to adjust my argument to change their opinion 4. I am good at getting people to help me when I need it. 5. People pay attention when I express my ideas. 6. People trust what I say. 7. I can motivate people to do what I ask. 8. I evaluate social situations to decide the best way to act. 9. I observe social situations carefully before deciding how to present an idea to others. 10. I am able to adjust my behavior to make a good impression. <p><i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale</i></p>
	14 Negotiation	<p>Defn: the ability to identify one's own and other's interests during a disagreement and to change one's own behaviors, thoughts and feelings as a strategy for resolving interpersonal problems and achieving one's goals</p> <p><i>Self Report Scale (Original items)</i></p> <ol style="list-style-type: none"> 1. When I disagree with someone, I try to manage my anger so I do not make the situation worse. 2. When I disagree with someone, I am able to give up some things I want to solve our disagreement. 3. Even when I disagree with someone, I still listen to them share their thoughts and views 4. When I disagree with someone, I think about the long-term consequence of my actions on the relationship before I do anything. 5. When I disagree with someone, I can come up with as many possible solutions to solve our problems. 6. When I disagree with someone, I can find solutions to the problem that help both me and the other person. 7. When I am in a disagreement with someone, I problem-solve to ensure we have a resolution that satisfies everyone. 8. When I am in a disagreement with someone, I can find solutions that improve the relationship for the future. <p><i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale</i></p>
	15 Collaboration	<p>Defn: considering different perspectives, listening and communicating in groups of two or more people, identifying situations involving group problem-solving and decision-making, and organizing and coordinating team members to create shared plans and goals</p> <p><i>Self Report Scale (Original items)</i></p> <ol style="list-style-type: none"> 1. When I work with others, I admit when I am wrong and apologize for my mistakes. 2. When I work with others, I tell others my ideas and ask for theirs in return. 3. When I don't know a solution to a problem, I can brainstorm with a group of people to get better ideas. 4. When working with others, I try to make sure that we agree on what actions each person will take to help us reach our goal. 5. When my team is having difficulty making a decision, I know what to do to help the team work together more effectively. 6. When I work with others, I clarify the problem we are trying to solve. 7. When I work with others, I summarize the information the group has agreed upon. <p><i>Alternative Measure: Scenario-based questions with similar theoretical structure to self report scale</i></p>
	16 Aggregate indices of SES	<p>Standardized indices of each skill will be examined individually and aggregated into the following categories: Intrapersonal Awareness, Intrapersonal Management, Interpersonal Awareness, Interpersonal Management, Intrapersonal, Interpersonal, All</p>

Table 6 cont...: Potential Mediators

M2. Business ideas, planning, and practices			
<i>Business Practices</i>	1	Business practices indices - Record keeping	Index: Has a written business plan; Has a written annual budget; Keeps financial records. Coded as 1 for each that the business has done in the last 12 months. (Note: coded as zero if not applicable)
	2	Business practices indices - Marketing Practices	Index: Visited at least one of its competitor's businesses to see what prices they are charging; Visited at least one of its competitor's businesses to see what products he or she offers; Asked existing customers whether there are products they would like you to offer; Asked a supplier about which products are selling well in this business' industry; looked for ways to improve your marketing and advertising strategies; Advertised in any form. Coded as 1 for each that the business has done in the last 12 months. (Note: coded as zero if not applicable)
	3	Business practices indices - Stock Control Practices	Index: Negotiated with a supplier for a lower price on raw material; Compared the prices or quality offered by your supplier's product/service with other suppliers. Coded as 1 for each that the business has done in the last 12 months. (Note: coded as zero if not applicable)
	4	Business practices indices - Financial Planning Practices	Index: analyzed if the sales of your most important product/services have increased, decreased or remained the same; looked for additional financial resources for your business; looked for new markets. Coded as 1 for each that the business has done in the last 12 months. (Note: coded as zero if not applicable)
<i>Business Ideas and Planning</i>	1	Number of business ideas	In the past twelve months, how many business ideas did you have?
	2	Entrepreneurial activity	In the past twelve months, how many times have you... (Gielnik et al., 2015) a. <input type="checkbox"/> visited one of your competitors in order to learn about the products he/she offers b. <input type="checkbox"/> asked your clients if there are products or services they would like you to offer c. <input type="checkbox"/> compared the prices or quality of your supplier's product/service with other suppliers d. <input type="checkbox"/> looked for new markets e. <input type="checkbox"/> analyzed if the sales of your most important product/services have increased, decreased or remained the same f. <input type="checkbox"/> looked for ways to improve your marketing and advertising strategies g. <input type="checkbox"/> looked for additional financial resources for your business h. <input type="checkbox"/> discussed with other entrepreneurs in your sector about production techniques, suppliers or new products
	3	Introduction of new/innovative product	Have you introduced new products or services in the 12 months? Is there a product/ service which was new in the locality at the moment you introduced it for the first time? Why have you introduced this new product /service? Can you name the two most important reasons? Is the new product/ service... variation of a brand or a new product. Has the new product/ service been introduced... as a result of own idea, or demand from customers etc.
<i>Aspirations</i>	1	Aspirations measure	Expectations for business growth: How do you see your business evolving in the next 12 months?

Table 6 cont...: Potential Mediators

M3. Intra-household Decision Making, Time Use and Mental Models		
<i>Decision-making power</i>	1	Decision-making: Female adult is (sole or joint) decision-maker related to use/sale/purchase/transfer of household resources to household expenditures. Sole or joint decision making power across a broad category of decisions: Who usually makes decisions about health care for yourself? Who usually makes decisions about making major household purchases? Who usually makes decisions about making purchases for daily household needs? Who usually makes decisions about visits to your family or relatives?
	2	Female has sole decision-making power on any decision Binary variable if female has decision making power on purchase of assets, plot or crop decisions etc.
<i>Time Use</i>	1	Share of time spent on household activities versus economic activities for respondent and their spouse. Calculate total time spent on household activities (care and domestic chores) and total time spent on economic activities and leisure. Respondent report for themselves and for their spouse. For economic activities and leisure we ask about time spent on the following activities: Tasks on non-farm business activities Paid (remunerated) work or activities outside of household or for someone not in the household Task on family farm, home garden, cattle herding, shepherding or fishing Leisure (including self-care and hobbies) Compare share of respondent with share of spouse. Include dummy indicators for if the respondent receives any assistance at home from anyone and/or spouse.
<i>Mental Models</i>	1	Social Norms Perceptions of norms around women's work and expression.
	2	Attitudes Index of Locus of control and Social Desirability Index (BIDR measure)
	3	Agricultural self-efficacy Index of agricultural self-efficacy
	4	Mental Health and Anxiety Index of anxiety

Table 7: Secondary Outcomes

Category	Outcome	Description
S1. Financial variables		
<i>Searching for finance/capital</i>	1 Applied for a business loan.	Binary variable indicate applied for a business loan in past 12months. Cash or in-kind through FIs or suppliers.
	2 Approved a business loan.	Binary variable indicate approved for a business loan in past 12months. Cash or in-kind through FIs or suppliers.
	3 Size of loan.	Total Amount of loans (self-reported)
S2. Marital Satisfaction and Intimate Partner Violence		
<i>Acceptability of Violence Against Women (VAW)</i>	1 Index of acceptability of violence against women in the household 0= not acceptable in any situation 1= acceptable in all stated situations	In your opinion, is a husband justified in hitting or beating his wife in the following situations. Is it okay for a man to hit his wife if she goes out without telling him? ... if she neglects the children? ... if she burns the food?
<i>Intrahousehold Relationship Quality</i>	1 Satisfaction with married life	Index of satisfaction with married life scale
	2 Trust within marriage	Index of trust scale
	3 Empathy in relationship with husband/wife	Index of relationship quality
S3: Asset Investments		
<i>Productive assets (Value and access)</i>	1 Livestock	Total number and current value
	2 Farming Assets	Total number and current value
	3 Household Assets	Total number and current value
	4 Other productive assets	Total number and current value
<i>Savings (Value and access)</i>	5 Savings	Any savings. Types of formal and informal savings used by households and current value by type.
S4. Consumption		
<i>Food Consumption</i>	1 Food own production	Total value consumed by household in last 7 days
	2 Food purchased	Total value on food expenditures during past 7 days.
<i>Non-food Expenditures</i>	3 Total amount spent in categories of frequent expenditures in last 7 days	Regular expenditures: Transport, leisure.
	4 Total amount spent in categories of infrequent expenditures in last month and six-month timeframe.	Communication, personal care goods, services, other
	5 Education expenditure	School fees in past 6 months
	6 Health expenditures	Medical expenses in past 6months
	7 Household Expenditures	Children clothing, adult's clothing, household utensils, etc.
	8 Social Expenditures	Ceremonies/festivals (weddings, funerals etc.). House decorations. Donations to church, mosque, other religious groups.
	9 Other Expenditures	Temptation goods – expenditures on tobacco and alcohol in past 7 days
<i>Index of total expenditures</i>	10 Index of daily adult equivalence consumption.	Total of food and non-food consumption rescaled by household size for adults and children and converted to a daily amount.
S5. Wellbeing		
<i>Subjective Wellbeing</i>	1 Happiness	Index from World Value Survey indicator.
	2 Life Satisfaction	Index from World Value Survey indicator.

Table 8: Program Selection - Comparison of APPEALS WYEP Program Applicants

	(1)	(2)	(3)	(4)	(5)	(6)
	Male Not Selected	Male Selected WYEP APPEALS	Female Not Selected	Female Selected WYEP APPEALS	Normalized difference	Normalized difference
Variable	Mean/SE	Mean/SE	Mean/SE	Mean/SE	(1)-(2)	(3)-(4)
Number of Children	1.476 [0.051]	1.483 [0.049]	2.404 [0.043]	2.454 [0.040]	-0.002	-0.022
Marital Status - Never Married	0.524 [0.010]	0.514 [0.009]	0.274 [0.008]	0.265 [0.008]	0.02	0.019
Work Experience-Total(years)	7.653 [0.112]	7.556 [0.103]	7.193 [0.130]	7.74 [0.130]	0.017	-0.075
Work Experience-Value Chain (Years)	4.426 [0.090]	4.261 [0.083]	3.797 [0.083]	3.514 [0.072]	0.037	0.066
Household Owns Land (0/1)	0.81 [0.008]	0.811 [0.007]	0.794 [0.008]	0.772 [0.007]	-0.004	0.052
Currently own a Business (0/1)	0.572 [0.010]	0.59 [0.009]	0.577 [0.009]	0.585 [0.009]	-0.037	-0.015
Household Average Monthly Income (Naira)	81337.383 [2149.062]	86000.002 [2123.517]	69301.03 [1814.095]	74048.581 [1666.355]	-0.042	-0.049
Value Chain Choice: Aquaculture (0/1)	0.104 [0.006]	0.127 [0.006]	0.098 [0.006]	0.153 [0.006]	-0.069	-0.164
Value Chain Choice: Cashew (0/1)	0.012 [0.002]	0.012 [0.002]	0.013 [0.002]	0.008 [0.002]	0.001	0.054
Value Chain Choice: Cassava (0/1)	0.103 [0.006]	0.11 [0.006]	0.147 [0.007]	0.101 [0.005]	-0.024	0.141
Value Chain Choice: Cocoa (0/1)	0.047 [0.004]	0.031 [0.003]	0.024 [0.003]	0.021 [0.003]	0.082	0.019
Value Chain Choice: Dairy (0/1)	0.057 [0.004]	0.053 [0.004]	0.032 [0.003]	0.034 [0.003]	0.017	-0.011
Value Chain Choice: Ginger (0/1)	0.014 [0.002]	0.006 [0.001]	0.008 [0.002]	0.008 [0.002]	0.083	0.001
Value Chain Choice: Maize (0/1)	0.052 [0.004]	0.039 [0.004]	0.033 [0.003]	0.021 [0.002]	0.065	0.076
Value Chain Choice: Poultry (0/1)	0.513 [0.010]	0.51 [0.009]	0.56 [0.009]	0.573 [0.009]	0.005	-0.025
Value Chain Choice: Rice (0/1)	0.087 [0.005]	0.104 [0.006]	0.078 [0.005]	0.073 [0.005]	-0.057	0.022
Value Chain Choice: Tomatoes (0/1)	0.01 [0.002]	0.008 [0.002]	0.004 [0.001]	0.008 [0.002]	0.024	-0.05
Value Chain Choice: Wheat (0/1)	0.001 [0.001]	0.001 [0.001]	0.001 [0.001]	0.002 [0.001]	0.001	-0.004
Number of observations	2653	2773	2892	3280		

Notes: The normalized difference in column (5) provides a scale-invariant measure of the size of the difference between males selected and not selected into the WYEP APPEALS program. The normalized difference in column (6) provides a scale-invariant measure of the size of the difference between females selected and not selected into the WYEP APPEALS program.