

Can Trainings on Women’s Voice and Agency Raise Women’s Aspirations, Entrepreneurial Activities, and Well-being? Evidence From Rural Nigeria

Long-term Followup Survey Amendment to Pre-Analysis Plan

Katrina Kosec, Jordan Kyle, Cecilia Hyunjung Mo

July 23, 2026

Abstract

This document is a follow-up amendment to the pre-analysis plan (PAP) for “Can Trainings on Women’s Voice and Agency Raise Women’s Aspirations, Entrepreneurial Activities, and Well-being? Evidence From Rural Nigeria.” With additional funding secured, we were able to conduct a second endline survey approximately two years after the first, allowing us to examine three sets of questions that the original study was not designed to answer. First, we test whether the intervention produced effects on economic outcomes—including earned income, assets, and savings—that were unlikely to emerge in the short-run but may have developed meaningfully over a longer time horizon. Second, we examine a set of outcomes conceptually related to our primary outcomes, including financial inclusion, ability to access to state/NGO assistance, food security, time allocation, engagement in better coping strategies, and ability to recover from shocks. We also examine child well-being outcomes, on the premise that any improvements in women’s economic circumstances may extend to investments in children in the long-run (e.g., as seen in [Duflo, 2003](#)). This amendment pre-registers the analyses we will conduct using these second endline data.

1. Introduction

This document is a follow-up amendment to the pre-analysis plan (PAP) for “Can Trainings on Women’s Voice and Agency Raise Women’s Aspirations, Entrepreneurial Activities, and Well-being? Evidence From Rural Nigeria.” It pre-registers the analysis plan for a second endline survey to be fielded approximately two years after the first endline, focusing on a set of economic outcomes that could not plausibly have changed in the short run but may develop meaningfully over a longer time horizon, as well as the durability of effects among a small set of measures collected at both the first and second endline surveys.

The study evaluates whether training interventions designed to strengthen women’s individual and collective efficacy and political participation influence their economic behaviors and outcomes—specifically their aspirations, entrepreneurial activities, and well-being. It further examines whether allyship training for male partners amplifies these effects. In a randomized controlled trial spanning 450 communities in southwestern Nigeria, we recruited 5,850 politically unaffiliated women into newly formed women’s action committees (WACs) trained by ActionAid Nigeria. All communities first participated in an identical burn-in process in which women were recruited into WACs and received a basic civic education session. The trainings began in May 2023. Following this, communities were randomly assigned to one of three conditions: a control group that received no further programming; a treatment group (T1) that received additional intensive training in advocacy, leadership, and organizing designed to build women’s individual and collective efficacy; and a second treatment group (T2) that received the same training as T1, with parallel allyship training offered to the male partners of participating women focused on men’s role in supporting women’s empowerment and gender equality.¹

¹We pre-registered a parallel study to look at the effects of the intervention on a set of outcomes related to the political empowerment of women. We found that both T1 and T2 significantly improved the level and quality of women’s political participation and enhanced their perceptions of the responsiveness of community leaders. Beyond perceptions, there is suggestive evidence of real shifts in leader behavior: treated women were more successful in having their community grant applications endorsed than women who were not treated with the training, pointing to at least some degree of increased responsiveness among local leaders stemming

The original PAP on economic outcomes pre-registered four hypotheses based on measures collected in an endline survey conducted in January–February 2024: that T1 and T2 would increase women’s economic aspirations (H1), economic activities (H2), and economic well-being (H3), and that effects would be larger for T2 than T1 (H4). The first endline found encouraging evidence in line with our pre-registered hypotheses. In the short-term, we found that the trainings produced meaningful early-stage gains. Significant improvements in personal income aspirations, savings aspirations, and social status aspirations were also observed. We also observed evidence of increased economic activities and economic well-being: trained women were more likely to talk with family members about income-generating ideas, inquire about loans, start saving toward business investment, seek information about non-farm enterprises, open bank accounts, borrow to invest, try new crops, and apply for government or NGO support. We observe improvements in women’s collective, political, and economic self-efficacy and perceived community and spousal support (including more housework by men and other material resources provided to women), though women’s power in intra-household bargaining remains unchanged. Notably, effects on economic activities and aspirations were statistically similar across T1 and T2, suggesting that the marginal contribution of the male allyship training on economic outcomes was limited in the short run.

However, the first endline was administered shortly after the trainings concluded—too soon for early-stage changes in aspirations and economic activities to translate into measurable gains in income, assets, or savings. Further, at the time of registering our PAP, we did not anticipate being able to fund a second follow-up survey. The core motivation for this follow-up amendment is that some of the most meaningful economic effects of the intervention may take time to materialize. Women who began saving toward a business, tried a new crop varietal, talked to others about how to raise their incomes or take out loans, or began saving for a business in the short run and had higher aspirations may only realize higher income,

from the trainings. These effects appear to operate through increased individual and collective agency, as well as a stronger sense of shared purpose (i.e., group identification) and increased procedural knowledge among women. See <https://osf.io/kxyj9/overview> for the pre-analysis plan for the study on political outcomes.

higher asset accumulation, or greater access to household assets over a longer time horizon.

At the same time, it is also possible that long-run effects on some outcomes that improved in the short-term due to the intervention are negative: women whose aspirations were raised by the training may encounter hard structural constraints—limited access to credit, markets, or land—and grow frustrated when those aspirations cannot be reasonably actualized in a constrained environment. Further, any temporary gains might have been wiped out due to a breakdown in social capital or skills following the cessation of organized trainings.

This amendment pre-registers the analysis we will conduct using the second endline data, which will resurvey the same women interviewed at baseline. Our goals are threefold: to assess whether the early-stage economic gains observed at the first endline have translated into measurable improvements in earned income, assets, and savings approximately two years later; and to examine a set of conceptually related outcomes, as well as child well-being outcomes, on the premise that improvements in women’s economic circumstances may extend to their children.

2. Overview of Follow-Up Design

The second endline (April—May 2026) will be fielded approximately 2.25 years after the first endline (January—February 2024) across the same 450 communities in southwestern Nigeria. The second endline will attempt to resurvey the exact same women interviewed at baseline. Some attrition is expected,² and we will assess whether there is differential attrition by treatment status following the same approach outlined in the original PAP and its first amendment.³ No new data sources are introduced.

The vast majority of economic outcomes analyzed in this amendment—including earned income, assets, savings, and financial inclusion—were not measured at the first endline, and

²Actual attrition at the first endline survey was very low; 93.7% (N = 3,634) of those who attended the burn-in meeting and completed the May–July 2023 baseline survey also completed our January–February 2024 endline survey.

³For the first endline survey, attrition did not differ by treatment condition.

the second endline therefore represents the first time these outcomes are being tested. We explicitly flag where an outcome here was also measured in an earlier survey wave. Treatment assignment remains unchanged from the original randomization. For details on the blocking structure and randomization procedure, we refer the reader to the original PAP and its first amendment.

3. Primary Outcomes and Hypotheses

This section pre-registers the primary outcomes and hypotheses (namely, H1 and H2) for the long-run follow-up study. For the majority of outcomes, the second endline represents the first time these outcomes are being tested, and we therefore estimate ITT effects of T1 and T2 at the second endline only. Both T1 versus control and T2 versus T1 are pre-registered as primary comparisons throughout.

3.1 H1: Women’s Personal Income and Assets

We test whether early-stage gains in women’s economic engagement detected in the short run leads to higher personal income and asset accumulation approximately two years later. H1 focuses specifically on income and assets that women personally own or control, rather than household-level resources more broadly. The long-run effects on women’s personal income and assets could be positive, negative, or null. A positive effect would be consistent with the hypothesis that the observed gains in economic engagement and aspirations in the short-term translated into higher income and asset accumulation in the long-term. A negative effect would be consistent with the hypothesis that women whose economic activities and aspirations were raised in the short run encountered external roadblocks and structural constraints—limited access to credit, markets, or land—and were unable to actualize those aspirations, leading to frustration that resulted in reduced long-term economic engagement relative to the control group, with negative implications for income. A null effect might suggest that women’s increased economic engagement was met with structural constraints or

other external roadblocks, and did not result in meaningful shifts in how women work.

H1a: Women's Personal Income

We will first test this hypothesis using a measure for a woman's total personal income and revenue generated from a household business with which she engages.

- *Total personal income*: women's total personal income will be measured as the sum of primary employment earnings (Q2), secondary employment earnings (Q4), and other income (Q6) (i.e., total income = Q2 + Q4 + Q6; this will be a missing value only if Q2, Q4, and Q6 are all missing), where exact question wording is provided below:

- *Primary employment personal earnings*: income earned from primary occupation in the past 30 days (the integer response for Q2 below; coded as 0 if there is no primary employment (i.e., responses of “You do not work,” “No other work,” or “Don't know” in Q1)).

* Q1: What is your current primary employment?

1. Farming
2. Herding
3. Casual / day labor
4. Processing, marketing, and trading agricultural products
5. Working for yourself / off-farm business
6. Salaried job
7. Other (specify): _____
8. You do not work
9. No other work

8888. [UNREAD] Nudged, but Don't Know

9999. [UNREAD] Nudged, but Refused to Answer

- * Q2 [Asked if response is 1-7 above]: How much did you earn from this line of work in the past 30 days? [Integer response]

- *Secondary employment personal earnings*: income earned from secondary occupation in the past 30 days (the integer response for Q4 below; coded as 0 if there is no secondary employment (i.e., responses of “You do not work,” “No other work,” or “Don’t know” in Q3)).
 - * Q3: What is your current secondary employment?
 1. Farming
 2. Herding
 3. Casual / day labor
 4. Processing, marketing, and trading agricultural products
 5. Working for yourself / off-farm business
 6. Salaried job
 7. Other (specify): _____
 8. You do not work
 9. No other work
 - 8888. [UNREAD] Nudged, but Don’t Know
 - 9999. [UNREAD] Nudged, but Refused to Answer
 - * Q4 [Asked if response is 1-7 above]: How much did you earn from this line of work in the past 30 days? [Integer response]
- *Other personal income sources*: income earned from any additional source in the past 30 days (the integer response for Q6 below; coded as 0 if there is no additional source of income (i.e., response of “No” or “Don’t know” in Q5)).
 - * Q5: Did you earn income from any other source in the last 30 days? [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
 - * Q6 [If “Yes” to previous question]: How much did you earn from that line of work in the past 30 days? [Integer Response]
- *Business revenue from woman-owned business*: revenue generated from a household-owned business in the past 30 days, where applicable, coded as 0 if the woman does

not report having a sole or joint ownership stake in the household business (the integer response for Q9 below; coded as 0 if there is no household business income (i.e., response of “No” or “Don’t know” in Q7) or if woman does not report any personal stake in the business (i.e., response of “spouse,” “other household members,” “spouse and other household member,” or “Don’t know” to Q8); and missing if “Refused to answer” to Q7, Q8, or Q9.

- Q7: Does your household own and operate a business? [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer]
- Q8 [If “Yes” to Q7]: Who in the household would you consider the owners of the business?
 1. Self
 2. Spouse
 3. Self and spouse jointly
 4. Other household members
 5. Self and other household members
 6. Spouse and other household members
 7. Self, spouse, and other household members
- 8888. *[UNREAD]* Nudged, but Don’t Know
- 9999. *[UNREAD]* Nudged, but Refused to Answer
- Q9 [If “Yes” to Q7]: How much revenue was generated from this business in the last 30 days? [Integer Response]

We will winsorize all personal earnings and business revenue variables at the top 1% of the sample to limit the influence of extreme reported values on treatment effect estimates.

H1b: Women’s Personal Assets

We will also test whether trained women have accumulated more assets over time. We focus on personal savings and livestock as the primary asset measures for women, as these are assets

that women in rural Nigeria are most likely to own. Livestock are an important asset for women because it is often easier for women to obtain livestock than it is for them to purchase land or other physical assets (Kristjanson et al., 2012). Specifically, we focus on chicken, goats, and sheep as they are the livestock types most commonly owned by women in this context. Chickens in particular are closely associated with women’s asset portfolios in rural Nigeria, as they require little capital to acquire, are managed within the household compound, and are widely understood to fall within women’s domain of ownership and decision-making authority.

- *Chickens owned by woman*: total number of chickens in the household (Q10), coded as 0 if the woman does not report sole or joint ownership of household chickens (i.e., response of “My husband or another member of the household owns them” or “Don’t know” to Q11); missing if “refused to answer” to Q10 or Q11.
 - Q10: “How many chickens does your household own today?” [Integer Response]
 - Q11: “Who in the household owns the chickens?”
 1. I own them on my own
 2. I own them jointly with my husband or another member of the household
 3. My husband or another member of the household owns them
- 8888. [UNREAD] Nudged, but Don’t Know
- 9999. [UNREAD] Nudged, but Refused to Answer
- *Goats and sheep owned by woman*: total number of goats and sheep in the household (Q12), coded as 0 if the woman does not report sole or joint ownership of household goats and sheep (i.e., response of “My husband or another member of the household owns them” or “Don’t know” to Q13); missing if “refused to answer” to Q12 or Q13.
 - Q12: “How many goats and sheep does your household own today?” [Integer Response]
 - Q13: “Who in the household owns the goats and sheep?”
 1. I own them on my own

2. I own them jointly with my husband or another member of the household

3. My husband or another member of the household owns them

8888. [UNREAD] Nudged, but Don't Know

9999. [UNREAD] Nudged, but Refused to Answer

- *Personal savings*: total current value of personal savings, including cash and household goods that the woman owns that could be sold quickly (the integer response for Q14 below).
 - Q14: What is the total current value of your own personal savings? Please include cash or household goods that you currently own that you could sell if you needed income quickly (e.g., a sewing machine, a chicken, or jewelry). [Integer response]
 - * As in the original PAP, we will winsorize personal savings data at the top 1% of the sample to limit the influence of extreme reported values on treatment effect estimates.
- *Woman's business ownership*: two dummy variables measuring women's business ownership, where the first is a dummy variable indicating whether women report owning and operating a business solely or jointly with another household member (1 if answered "Self," "Self and spouse jointly," "Self and other household members," or "Self, spouse, and other household members" to Q8 (described above); 0 if answered "No" or "Don't Know" to Q7 (described above) or if answered "Spouse," "Other household members," "Spouse and other household members," or "Don't Know" to Q8); and the second is a dummy variable indicating whether women report sole ownership of a household business (1 if answered "Self" to Q8; 0 if answered "No" or "Don't Know" to Q7 or if answered "Self and spouse jointly," "Self and other household members," "Self, spouse, and other household members," "Spouse," "Other household members," "Spouse and other household members," or "Don't Know" to Q8). Again, refusals to answer both Q7 and Q8 will lead to missing data, since an unwillingness to answer provides no basis for inferring ownership or control.

- *Women’s revenue control*: We additionally asked women about who controls the revenue for any household business (Q9a), as revenue control and revenue generation are conceptually distinct measures of women’s empowerment. We will explore answers to this question in case of potential mismatches between responses on business ownership and revenue control.
 - * Q9a [If “Yes” to Q7]: Who in the household has control of the revenue generated from the business?
 1. Self
 2. Spouse
 3. Self and spouse jointly
 4. Other household members
 5. Self and other household members
 6. Spouse and other household members
 7. Self, spouse, and other household members
 - 8888. [UNREAD] Nudged, but Don’t Know
 - 9999. [UNREAD] Nudged, but Refused to Answer

To aggregate these various measures of women’s personal assets, we will construct a *women’s personal asset index* using the first principal component from a principal component analysis (PCA) of women’s ownership-adjusted chicken count, goat and sheep count, personal savings, business ownership (sole or joint, as well as sole), and business revenue control. We will also report results for each component identified above separately.

3.2 H2: Combined Income and Assets of the Woman-Husband Pair

All women participants were married at baseline. H2 examines whether the trainings improved the combined economic resources of the woman-husband pair. We focus on the dyadic unit of the woman and her husband rather than the full household, as the household may include other wives and income earners whose economic outcomes are not directly linked to the

intervention. The trainings could affect the combined income and assets of the woman-husband pair through several distinct pathways. The most direct pathway is additive: if the trainings increased women’s own income and assets (H1) without affecting husbands’ economic outcomes, combined dyadic income and assets would increase by the same magnitude as women’s personal gains.

A second and more optimistic pathway is complementary: the trainings may have improved coordination and information sharing within the household, enabling women and their husbands to make better joint decisions about labor allocation, investment, and asset management. For example, women who became more confident and assertive through the trainings may have been better able to advocate for household-level investments in productive assets or to identify economic opportunities that benefit both spouses. In T2 communities specifically, where husbands also received allyship training, this complementary dynamic may be stronger: husbands who were trained to support their wives’ economic activities may have actively redirected household resources toward women’s income-generating activities, producing gains in combined dyadic income that exceed women’s personal gains alone. A third pathway, however, is substitutive: if women’s increased economic engagement came at the expense of time previously allocated to supporting their husbands’ economic activities — for example, by reducing unpaid labor on family farms or in household businesses — gains in women’s personal income could be partially or fully offset by reductions in husbands’ income, producing smaller or null effects on combined income even when H1 effects are positive. Distinguishing between these pathways is one motivation for examining H1 and H2 separately: divergent effects across the two hypotheses would provide evidence about whether the trainings expanded the overall economic resource base of the woman-husband pair or primarily redistributed resources and activities within it.

We use the following measures, drawing on both the women’s second endline survey and the men’s second endline survey:

H2a: Combined Income of the Woman-Husband Pair

- *Combined income of husbands and wives*: sum of women’s total personal income and husband’s total personal income, where women’s total personal income is defined in H1a above (Q2, Q4, Q6) and where husbands total personal income is defined as the total income earned by the husband from all employment sources in the past 30 days, constructed from the men’s second endline survey using the same questions and approach as for women’s total personal income above.
- *Household business revenue — women’s report*: revenue generated from a household-owned business in the past 30 days, as reported by the woman in the women’s second endline survey (Q9 above), coded as 0 if ‘Other household members’ (and not the woman or her spouse) have ownership.
- *Household business revenue — husband’s report*: revenue generated from a household-owned business in the past 30 days, as reported by the husband in the men’s second endline survey, coded as 0 if ‘Other household members’ (and not the woman or her spouse) have ownership. We will report treatment effects on both the women’s and husband’s reports of business revenue separately, and will flag cases where the two reports disagree substantially.

H2b: Combined Assets of the Woman-Husband Pair

- *Total household chickens*: total number of chickens owned by the household, as reported by the woman (Q10 above), without the ownership adjustment applied in H1b (i.e., Q11 is irrelevant).
- *Total household goats and sheep*: total number of goats and sheep owned by the household, as reported by the woman (Q10 above), without the ownership adjustment applied in H1b (i.e., Q13 is irrelevant).
- *Household business ownership*: a dummy variable equal to 1 if the household owns and operates a business, as reported by the woman (Q7 above).

We will construct a *dyadic asset index* using the first principal component from a PCA of total household chicken count, total household goat and sheep count, and household business ownership. Higher values indicate greater asset accumulation at the dyadic level. We will also report results for each component separately.

4. Secondary Outcomes

We pre-register a set of secondary outcomes that are conceptually related to the primary outcomes but are not the focus of the core hypotheses. For each secondary outcome, we estimate ITT effects of T1 versus control, T2 versus control, and T2 versus T1 at the second endpoint using the same covariate-adjusted difference-in-means specification as the primary analysis detailed in the original pre-analysis plan and the first amendment (see more details below). We do not pre-register directional hypotheses for these outcomes and treat them as descriptive tests of whether the intervention had broader effects on women’s economic lives beyond the primary outcomes of earned income and assets.

4.1 Financial Inclusion

We will test whether treated women show higher rates of financial inclusion at the second endpoint, using the following measures:

- *Bank account ownership*: indicator variable for the respondent having a bank account or a joint bank account with her spouse, based upon Q15 (1 if answered “I have an account” or “My spouse and I have a joint account”; 0, otherwise (missing if “Refused to answer”).
 - Q15: Do you have a bank account or a joint bank account with your spouse?
 1. I have an account
 2. My spouse has an account
 3. My spouse and I have a joint account
 4. My spouse and I do NOT have a personal or a joint account

8888. [UNREAD] Nudged, but Don't Know

9999. [UNREAD] Nudged, but Refused to Answer

- *Financial activity*: an equally weighted *financial inclusion index* of three binary inclusion measures (Q16-Q18), ranging from 0 to 3, based on the following yes/no question regarding engagement in financial transactions that are indicative of financial inclusion (i.e., accessing bank accounts, credit, mobile money accounts) to create a series of binary measures (1 if answered “Yes,”; 0, otherwise (missing if “Refused to answer”)).
 - Q16: In the past 12 months, have you ever: Used any mobile money account to make a payment, buy something, or send money to someone? [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
 - Q17: In the past 12 months, have you ever: Deposited or withdrawn money into a bank account? [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
 - Q18: In the last 2 years, did you take out a loan personally or jointly with someone else? This could be a formal loan from a bank or an informal loan from an individual. [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]

4.2 Access to Government and NGO Benefits

We will test whether treated women show greater success in accessing benefits from government or NGO programs at the second endline, using the following measures:

- *Benefit-seeking behavior*: indicator measure for anyone in the household trying to gain access to a benefit from the government or an NGO in the last two years (1 if answered 2, 3, 4, 5, or 97 to Q19; 0 if answered 1 or “Don’t Know” to Q19).
 - Q19: Over the last 2 years, has anyone in your household tried to gain access to a benefit from the government or NGO, even if they did not succeed in doing so? Benefits could include a cash transfer, educational support, health support, employment assistance, and support for agricultural activities.

1. Our household did not apply for any benefits
 2. A cash transfer aimed at low income families
 3. A health insurance scheme
 4. An educational support program
 5. An economic support program (e.g., fertilizer voucher, subsidized seeds)
 97. Other (specify)
 8888. *[UNREAD]* Nudged, but Don't Know
 9999. *[UNREAD]* Nudged, but Refused to Answer
- *Benefit receipt*: indicator measure for the household obtaining any benefits they applied for (1 if answered 2, 3, 4, 5, or 97 to Q20; 0 if answered 1 or "Don't Know" or did not select any response to Q20)
 - Q20: Was your household successful in getting any of the benefits they applied for?
 1. Our household did not apply for any benefits
 2. A cash transfer aimed at low income families
 3. A health insurance scheme
 4. An educational support program
 5. An economic support program (e.g., fertilizer voucher, subsidized seeds)
 97. Other (specify)
 8888. *[UNREAD]* Nudged, but Don't Know
 9999. *[UNREAD]* Nudged, but Refused to Answer

We will report results for each measure separately rather than combining them into an index, as benefit-seeking behavior and benefit receipt capture conceptually distinct dimensions of engagement with government and NGO programs.

4.3 Food Security

We will test whether treated women show higher food security at the second endline, using the following measure:

- *Meal skipping*: indicator measure of whether the respondent or any member of her household had to skip a meal in the past two weeks due to lack of money or other resources (1 if answered “Yes” or “Don’t Know” to Q21; 0, otherwise (missing if “Refused to answer”). Note that we code “Don’t Know” responses as missing, consistent with HFIAS recommendations. The meal skipping measure is consistent with the food security measure used in the first amendment to the original PAP, which analyzed meal skipping as the primary indicator of food security. Higher values indicate lower food security; we will reverse-code it so that higher values reflect better food security outcomes.
 - Q21: In the past 2 weeks, was there a time when, because of lack of money or other resources: You or any member of your household had to skip a meal? [Response Options: “Yes,” “No,” “Don’t Know, “Refused to Answer”]

4.4 Child Well-being

We will also assess whether the women’s training generated positive spillover effects for their children. In these analyses, we will examine the program’s long-term effects on women’s occupational aspirations for their children, child health outcomes, and investments in their children’s education.

- *Occupational aspirations for children*: We will focus, first, on aspirations for respondents’ female children, as the trainings focus on women’s empowerment (Q22). Parents’ aspirations for their daughters will be captured using a variable ranging from 0–2, where a ‘1’ indicates that the preferred occupation is something other than a housewife, and a ‘2’ indicates that the preferred occupation is a teacher, nurse / doctor, or another

salaried job (see [Beaman et al. \(2012\)](#) for a similar approach).⁴ Second, we will focus on any aspirations gaps between respondents’ male (Q23) and female (Q22) children by constructing a similar variable for boys’ attainment and then taking the difference between what respondents aspire for their girl versus boy children.

- Q22: What primary occupation do you hope that your daughter would have in the future? If you do not have a daughter, please think about how you would feel if you had a daughter
- Q23: What primary occupation do you hope that your son would have in the future? If you do not have a son, please think about how you would feel if you had a son.
- *Child health*: Composite index of *daughter health*, combining fever incidence (Q24), diarrhea incidence (Q26), and overall health status (Q28) (this measure is missing for households without a school-aged daughter (i.e., answered “No” to the question: “Do you have any school-aged daughters?”)), as well as an index for *son health* based on Q25, Q27, and Q29 (missing for households without a school-aged son (i.e., answered “No” to the question: “Do you have any school-aged sons?”)). The two composite measures range from 0 to 3, as they each sum three binary measures, where higher values represent better health.
 - Indicator for the youngest school-aged daughter NOT having had a fever in the past 2 weeks (missing for households without a school-aged daughter) based on question Q24: “In the past 2 weeks, has your daughter been sick with a fever at any time? If you have more than one daughter, think about your youngest school-aged daughter.” [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
 - Indicator for the youngest school-aged son NOT having had a fever in the past

⁴Because ‘another salaried job’ could have a wide interpretation, we will probe using an open-ended response what parents have in mind if they select this response category, and determine on a case-by-case basis whether this should be included as among the high educational attainment and/or high income occupations category.

- 2 weeks (missing for households without a school-aged son) based on question Q25: “In the past 2 weeks, has your son been sick with a fever at any time? If you have more than one daughter, think about your youngest school-aged daughter.” [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Indicator for the youngest school-aged daughter NOT having experienced diarrhea in the past 30 days (missing for households without a school-aged daughter) based on question Q26: “Has your daughter suffered from diarrhea in the past 30 days?” [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
 - Indicator for the youngest school-aged son NOT having experienced diarrhea in the past 30 days (missing for households without a school-aged son) based on question Q27: “Has your son suffered from diarrhea in the past 30 days?” [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
 - Indicator for the youngest school-aged daughter having been reported to be “somewhat” or “very” healthy in the past 2 weeks (missing for households without a school-aged daughter or if refused to answer) when asked Q28: “In the last 2 weeks, how healthy has this child been?” [Response Options: “Very,” “Somewhat,” “A little bit,” “Not at all,” “Don’t know,” “Refused to Answer”]
 - Indicator for the youngest school-aged son having been reported to be “somewhat” or “very” healthy in the past 2 weeks (missing for households without a school-aged son or if refused to answer) when asked Q29: “In the last 2 weeks, how healthy has this child been?” [Response Options: “Very,” “Somewhat,” “A little bit,” “Not at all,” “Don’t know,” “Refused to Answer”]
- *Education investment: daughter school investment* is the total amount spent on youngest school-aged daughter (Q30) and *son school investment* is the total amount spent on youngest school-aged son (Q31), based upon the following questions.
 - Q40: Has your household bought any school materials to prepare this child [youngest school-aged daughter] for next school term? If so, how much total

has been spent so far? [Integer Response] (missing for households without a school-aged daughter)

- Q41: Has your household bought any school materials to prepare this child [youngest school-aged son] for next school term? If so, how much total has been spent so far? [Integer Response] (missing for households without a school-aged son)

4.5 Time Allocation.

The trainings may have enabled women to shift time towards income-generating and community participation activities, producing income gains and civic engagement that accumulate over time. Since the 24-hour time budget is fixed, any increase in income-generating or community participation time must come at the expense of other activities including sleep and self-care. We remain agnostic regarding on which activity categories there may be changes in time allocation. We will examine all of the following six time use categories from the stone allocation module described below (Q42) as time allocation outcomes:

1. *Time spent sleeping*
2. *Time spent on self-care and personal activities*
3. *Time spent on unpaid care work and household tasks*
4. *Time spent on farm work, animal husbandry, or fishing*
5. *Time spent on income-generating activities other than farm work*
6. *Time spent on community participation activities*

Specifically, individuals were shown a diagram with 6 boxes and pictures depicting the six activities above. They were then told:

- Q42: “We would now like to understand how women in this community spend their time. I am going to ask you to think about yesterday—from the time you woke up to the time you went to sleep at night. I have 24 stones here. Each stone represents one hour of the day and make up one full day. Please allocate the 24 stones across the 6

boxes on this sheet.

- The first box is for sleeping.
- The second box is for time spent on yourself - not working, not sleeping, just for you. This includes things like washing and dressing in the morning, eating your meals, resting when you are not feeling well, or sitting and talking with friends or family for enjoyment.
- The third box is for unpaid work you do to take care of your family. This includes caring for children, elder members of the family, cooking, cleaning, washing clothes, fetching water, collecting firewood, and doing any other household tasks.
- The fourth box is for work on the farm, with animals, or with fishing - whether for your household to eat or to sell. This includes planting, weeding, watering, harvesting, processing food from your farm such as pounding grain or drying cassava, or tending animals.
- The fifth box is for any work connected to earning money or running a business - other than farm work. This includes selling goods at a market or roadside, running a small business like a food stall or tailoring, or any other work for which you received or expect to receive a wage, payment, or goods in return.
- The sixth box is for the time spent participating in community life - not for pay, and not just for personal enjoyment. This includes attending religious services or prayer rouns, going to community meetings, participating in a women's group, or serving in any village or community role."

We will additionally construct a combined measure of time spent on income-generating activities by summing the farm work and non-farm income-generating categories to measure overall time spent on economic activities. Given that the trainings were primarily focused on building women's confidence and networks for civic participation outside the home, we expect gains to be concentrated in the non-farm income-generating category rather than farm work, and will examine these two categories separately in addition to the combined

measure. Additionally, we will sum time spent sleeping and time spend on self-care and personal activities to assess total “leisure” time in the aggregate. We will also separately look at data reported by women and by their husbands.

4.6 Negative Coping Strategies

We also examine whether the trainings reduced women’s and men’s reliance on harmful economic behaviors when facing resource shortfalls. While the primary outcomes of this PAP focus on income and asset accumulation, negative coping strategies capture a downstream welfare dimension that is closely linked to the economic empowerment story: women and men who have built stronger economic resources, social networks, and advocacy skills through the trainings may be less likely to resort to harmful strategies when facing resource constraints. Each of the eight items described below (Q43a-Q43g) is coded on a 5-point scale, ranging from 0 indicating “never” engaging in the behavior to 4 indicating that they engage in this behavior “frequently (more than once per month).” “Don’t Know” and Refused to Answer” responses are treated as missing.

- Q43: We would now like to ask whether, and how often, members of your household have to engage in any of the following behaviors due to a lack of food or a lack of money to buy food or meet other basic needs in the past 12 months:
 - Q43a: Mortgaged or sold household assets or goods (e.g. radio, furniture, bicycle, jewelry) [Response Options: “Never,” “Once,” “Sometimes (every few months),” “Often (once a month),” “Very frequently (more than once per month)”]
 - Q43b: Mortgaged or sold productive assets (e.g. business equipment, land) or means of transport [Response Options: “Never,” “Once,” “Sometimes (every few months),” “Often (once a month),” “Very frequently (more than once per month)”]
 - Q43c: Spent savings [Response Options: “Never,” “Once,” “Sometimes (every few months),” “Often (once a month),” “Very frequently (more than once per month)”]

- Q43d: Borrowed money to buy food or borrowed food on credit [Response Options: “Never,” “Once,” “Sometimes (every few months),” “Often (once a month),” “Very frequently (more than once per month)”]
- Q43e: Sent one or more household members to live elsewhere [Response Options: “Never,” “Once,” “Sometimes (every few months),” “Often (once a month),” “Very frequently (more than once per month)”]
- Q43f: Decreased expenditures on agriculture, livestock, or fisheries inputs [Response Options: “Never,” “Once,” “Sometimes (every few months),” “Often (once a month),” “Very frequently (more than once per month)”]
- Q43g: Decreased expenditures on health [Response Options: “Never,” “Once,” “Sometimes (every few months),” “Often (once a month),” “Very frequently (more than once per month)”]
- Q43h: Decreased expenditures on education [Response Options: “Never,” “Once,” “Sometimes (every few months),” “Often (once a month),” “Very frequently (more than once per month)”]

We will construct a *negative coping strategies index* using the first principal component from a principal components analysis (PCA) of these eight items, reverse-coded so that higher values reflect lower usage of negative coping strategies. We use PCA rather than an equally-weighted index because the eight items are likely to load differently on the underlying construct of economic stress — for example, selling productive assets or reducing health and education expenditures may represent more severe forms of distress than spending savings or borrowing food on credit, and PCA allows the data to determine the appropriate weighting rather than imposing equal weights across behaviorally heterogeneous items. We will report the factor loadings from the PCA alongside the main results to allow readers to assess which items are driving the index. With the same objective, we will also report results for each component separately. We will also separately look at data reported by women and by their husbands.

4.7 Exposure to Shocks and Shock Recovery

Finally, we examine whether the trainings improved households’ ability to bounce back from adverse events. Women who built stronger economic resources, social networks, and advocacy skills through the trainings should be better positioned to recover from shocks—through access to informal credit and support networks, more effective engagement with local leaders and service providers, and a stronger economic base from which to rebuild. It also provides important context for interpreting the primary income and asset outcomes. Households that experienced severe and unrecovered shocks after the intervention period (i.e., in the nearly 2.5 years preceding the second endline) may show attenuated treatment effects on income and assets, not because the trainings were ineffective, but because adverse shocks eroded the gains that would otherwise have been observed.

Shock exposure and shock recovery are conceptually distinct: shock exposure captures the objective shock environment that households faced, while shock recovery captures households’ capacity to return to normalcy after being affected. Reporting them side by side allows us to distinguish between two possible pathways through which the trainings could improve household welfare: reducing the frequency or severity of shocks that households experience, or improving their capacity to recover from shocks when they do occur. It also provides important context for interpreting the primary income and asset outcomes: households that experienced severe and unrecovered shocks in the two years preceding the second endline may show attenuated treatment effects on income and assets, not because the trainings were ineffective, but because adverse shocks eroded the gains that would otherwise have been observed.

We assess exposure to shocks using a series of indicator variables (Q44) for whether the household answered “yes” (i.e., code “yes” as 1 and “no” as 0) to each of the following adverse events in the past two years, for both women and men:

- Q44: Have you experienced any of the following in the past 2 years?
 - Q44a: Damage to houses / dwellings or any productive assets due to floods, soil

- / water erosion, heavy rains [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Loss of crop due to drought or insufficient rainfall [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Q44b: Crop pest or disease outbreak
- Q44c: Damage to crops, livestock, or farm infrastructure by herders [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Q44d: Loss of livestock due to disease [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Q44e: Loss of livestock due to theft or attack [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Q44f: Death of the household head or primary earner [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Q44g: Large unexpected medical expense [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Q44h: Erosion or land degradation affecting farming area [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Q44i: Water source drying up or becoming contaminated [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Q44j: Kidnapping of a family member [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Q44k: Robbery, banditry, or armed attack on household or farm [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Q44l: Forced displacement from home or farm [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]
- Q44m: Income loss due to conflicts between herders and farmers [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]

We will construct an equally-weighted *shock exposure index* as the share of the 14 shock types that the household experienced in the past two years. For respondents who report not knowing if a shock occurred, we score it as 0 (i.e., as if the shock did not occur); refused to answer responses are treated as missing.

For each of the 14 items above (except for experiencing income loss due to conflicts between herders and farmers, for which we did not ask additional questions), when someone indicated that they experienced the shock, we also asked whether they recovered from the shock.

- Q45: What was the severity of the impact on household welfare? [Response Options: “Mild,” “Moderate,” “Severe,” “Don’t Know,” “Refused to Answer”]
- Q46: Has the household fully recovered? [Response Options: “Yes,” “No,” “Don’t Know,” “Refused to Answer”]

Our primary recovery measure is a *severity-weighted recovery index* that jointly captures both the severity of shocks experienced and the degree to which households recovered from them. For the first 13 of the 14 shock types listed above (Q44), we assign a score as follows: households that did not experience the shock are coded as missing; households that experienced the shock but report not having fully recovered receive a score of 0 regardless of severity; and households that experienced the shock and report having fully recovered receive a score equal to the severity of the shock (mild = 1, moderate = 2, severe = 3). We then compute the index as the equally-weighted average of these scores across all 13 shock types. Higher values indicate stronger recovery capacity, with greater weight given to recovery from more severe shocks. Where a household was unaffected by a shock, that shock does not enter its recovery index, and we average over the remaining shocks that were experienced. Where a household experienced no shocks of any kind, the shock recovery index will be coded as missing, as they have no shock-related recovery to contribute to the measure.

Reporting the shock exposure index and the severity-weighted recovery index side by side allows for a complete picture of the shock environment and recovery dynamics: the shock

exposure index tells us whether treated households faced fewer shocks, while the severity-weighted recovery index tells us whether households that faced shocks were better able to recover from them. A pattern in which the treatment improves recovery but not exposure would be consistent with the theory of change — that the trainings built women’s social capital, self-efficacy, and advocacy skills in ways that helped them bounce back from adversity rather than reducing the adversity itself. Don’t Know” responses on shock occurrence are coded as 0 (no shock); Don’t Know” responses on severity or recovery and all “Refused to Answer” responses are treated as missing at the item level, with both indices computed as the mean of non-missing items.

As a pre-registered robustness check, we additionally estimate treatment effects using an *unweighted recovery index* that treats all shock types and severity levels equally. In other words, for each of the 13 shock types listed above (Q44), we assign a score as follows: households that did not experience the shock are coded as missing; households that experienced the shock but report not having fully recovered receive a score of 0 regardless of severity; and households that experienced the shock and report having fully recovered receive a score of 1. We then compute the index as the equally-weighted average of these scores across all 13 shock types.

5. Analysis Strategy

5.1 Estimation

As in the original PAP and its first amendment, we estimate intent-to-treat (ITT) effects using covariate-adjusted difference-in-means with block fixed effects and standard errors clustered at the community level.

For outcomes that were not measured at the first endline—including earned income, assets and asset control, and access to government benefits—the primary specification is:

$$Y_{ic} = \alpha + \beta_1 T1_c + \beta_2 T2_c + \mathbf{X}'_{ic}\gamma + \delta_b + \varepsilon_{ic}, \quad (1)$$

where Y_{ic} is the outcome for individual i in community c ; $T1_c$ and $T2_c$ are indicators for treatment assignment; $\mathbf{X}'_{ic}$ is the vector of baseline covariates specified in Section 4.2 of the political participation amendment; and δ_b denotes block fixed effects. Standard errors are clustered at the community level. Since these outcomes were not measured at the first endline, no lagged dependent variable is included in the covariate set; identification relies solely on randomization and the baseline covariates from the women's baseline survey. The coefficients of primary interest are β_1 , capturing the ITT effect of T1 relative to control; β_2 , capturing the ITT effect of T1 relative to control; and $\beta_2 - \beta_1$, capturing the marginal effect of T2 relative to T1.

For any measures that we have at both endline and baseline, we will assess for the possibility of change between the first and second endline using a pooled stacked dataset combining both endline surveys. The specification is:

$$Y_{ict} = \alpha + \beta_1 T1_c + \beta_2 T2_c + \beta_3 \text{Wave}_t + \beta_4 (T1_c \times \text{Wave}_t) + \beta_5 (T2_c \times \text{Wave}_t) + \mathbf{X}'_{ict} \gamma + \beta_6 Y_{ic,\text{pre}} + \delta_b + \varepsilon_{ict}, \quad (2)$$

where Y_{ict} is the outcome for individual i in community c at wave t ; Wave_t is an indicator equal to 1 for the second endline and 0 for the first endline; $Y_{ic,\text{pre}}$ is the baseline measure of the outcome included as a covariate (if measured at baseline); and all other terms are as defined above. Standard errors are clustered at the community level and block fixed effects are included throughout. We are interested in $\beta_1 + \beta_4$, as that is the T1 effect in the long-term, as well as $\beta_2 + \beta_5$, as that is the T2 effect in the long-term. We are also interested in β_4 and β_5 , which capture the delta between short-run and long-run effects for T1 and T2, respectively, relative to the control group. We also test the difference in long-run versus short-run effects for T2 relative to T1 directly by testing $H_0: \beta_1 + \beta_4 = \beta_2 + \beta_5$.

5.2 Index Construction and Standardization

Index construction and standardization follow the procedures outlined in the original PAP and its first amendment. All raw index components are standardized by dividing by the

standard deviation in the control group. Where we take averages of index components, we use equally-weighted averages of the standardized components. Indices that are themselves averages of sub-indices are re-standardized using the mean and standard deviation of the control group. This procedure ensures that all treatment effects can be interpreted in standard deviation terms. Multiple testing is addressed through the use of pre-specified indices rather than corrections applied to individual items.

5.3 Treatment of Income and Savings Data

Income and personal savings data will be winsorized at the top 1% of the sample to address the possibility of inaccurately reported values, consistent with the first amendment.

5.4 Missing Data

We follow the missing data procedures outlined in the original PAP and its first amendment. Responses of “refused to answer” are treated as missing. Responses of “don’t know” are treated as missing unless a “don’t know” clearly indicates substantive meaning, such as an incorrect response in a knowledge question or demonstrating lack of asset control (if a respondent does not know who owns an asset, this most plausibly reflects an absence of personal ownership or control). When constructing outcome indices, we follow the strategy outlined in Section 4 of the MetaPAP for handling missing data within indices.

5.5 Attrition

Some attrition is expected between the first and second endlines. We will assess whether attrition between the first and second endline is differential by treatment status by regressing missingness indicators for each outcome on treatment assignment and covariates, following the same specification as our main outcomes. If these tests are non-significant, we will proceed under the assumption that units are either “always-reporters” or “never-reporters” and interpret all effects as ITT effects among always-reporters. If these tests are significant, as

noted in our first amendment, we will report extreme value bounds. We will not use adjustment methods like multiple imputation or weighting by the inverse probability of missingness that rely on an conditional ignorability assumption that conditional on observables, missingness is as-if random.

References

- Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova. 2012. “Female leadership raises aspirations and educational attainment for girls: A policy experiment in India.” *science* 1212382.
- Duflo, Esther. 2003. “Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in South Africa.” *The World Bank Economic Review* 17 (1): 1–25.
- Kristjanson, Patti, Ann Waters-Bayer, Nancy Johnson, Anna Tipilda, Jemimah Njuki, Isabelle Baltenweck, Delia Grace, and Susan MacMillan. 2012. “Livestock and Women’s Livelihoods: A Review of the Recent Evidence.” *ILRI Discussion Paper* 20.