Impact Evaluation of the Nigeria for Women Project

Baseline Findings

Thomas de Hoop, Adria Molotsky, Christopher Paek, Rosa Castro-Zarzur, Olayinka Adegbite, Iyabo Adeoye, Philomena Panagoulias, Anna Warren, Cody Bock, Yewande Ajayi, Nuraddeen Sambo Umar, Golden Aja Obasi, and Garima Siwach

DECEMBER 2022

AIR®

Advancing Evidence.
Improving Lives.

Impact Evaluation of the Nigeria for Women Project

Baseline Findings

Thomas de Hoop, Adria Molotsky, Christopher Paek, Rosa Castro-Zarzur, Olayinka Adegbite, Iyabo Adeoye, Philomena Panagoulias, Anna Warren, Cody Bock, Yewande Ajayi, Nuraddeen Sambo Umar, Golden Aja Obasi, and Garima Siwach

DECEMBER 2022

Contents

Acronyms

ADP	Agricultural Development Projects
AIR	American Institutes for Research
ATT	Average Treatment Effect on Treated
A-WEIA	Abbreviated Women's Empowerment in Agriculture Index
DHS	Demographic and Health Survey
DID	Difference-in-Differences
EA	Enumeration Area
ECWG	Evidence Consortium on Women's Groups
FPCU	Federal Project Coordinating Unit
FGD	Focus Group Discussion
GBV	Gender-Based Violence
HH	Household
IDI	In-Depth Interview
IGA	Income-Generating Activity
ICBA	Institutional Capacity Building Advisor
ICC	Intra-Class Correlation
ITT	Intent-to-Treat
JEEVIKA	Bihar Rural Livelihoods Project
KII	Key Informant Interview
LATE	Local Average Treatment Effect
LFP	Labor Force Participation
LGA	Local Government Area
M&E	Monitoring and Evaluation
MIS	Management Information System
MOWASD	Ministry of Women and Social Development
NGN	Nigerian Naira
NFWP	Nigeria for Women Project

NRLM   National Rural Livelihoods Mission

OPM   Oxford Policy Management

PLA   Participatory Livelihoods Assessment

RCT   Randomized Controlled Trial

RQA   Rapid Qualitative Assessment

SG   Savings Group

SHG   Self-Help Group

SPCU   State Project Coordinating Unit

ToC   Theory of Change

VSLA   Village Savings and Loan Associations

WAG   Women's Affinity Group

Impact Evaluation of the Nigeria for Women Project: Baseline Report

Acknowledgements

Executive Summary

Background

Women in Nigeria face significant barriers to accessing savings, credit, livelihoods, and markets, constraining their ability to achieve economic empowerment. Traditional gender norms and other social norms, information asymmetries, and limited access to finance all contribute to gender inequality in Nigeria (World Bank, 2018).

Formalized savings groups and other women's groups with economic objectives may have the potential to improve women's economic empowerment in Nigeria (e.g., Desai et al., 2018; de Hoop et al., 2022). Members of savings groups pool small weekly savings into a common fund, which members can then borrow against (Adegbite et al., 2022). "Women's groups" is an umbrella term referring to different models of economic, health, and community groups, including but not limited to savings groups, with a primarily female membership (Anderson et al., 2019). Various impact evaluations show evidence for positive effects of savings groups and other women's groups, including in various settings in sub-Saharan Africa (e.g., Blattman et al., 2016; Brody et al., 2015; Karlan et al., 2017; Desai et al., 2019; de Hoop et al., 2020; de Hoop et al., 2022). While the magnitude of the relationship is small and the analyses do not allow for establishing causality, membership in women's groups and informal savings groups is also positively correlated with women's asset ownership and decision-making power in a nationally representative sample in Nigeria (Meysonnat et al., 2022a; Meysonnat et al., 2022b).

The Government's Nigeria for Women Project (NFWP) aims to achieve improvements in women's economic empowerment by establishing new women's groups and transforming existing informal savings groups and other women's or mixed-gender groups into formal, women-only savings groups, called Women Affinity Groups (WAGs). WAGs are formal savings groups that follow established savings group implementation models (e.g., Village Savings and Loan Associations [VSLAs] and Savings and Internal Lending Committees) to support women's financial inclusion. The NFWP aims to improve women's livelihood opportunities and facilitate their access to economic markets by mobilizing women into WAGs and providing additional trainings and financial support for improving their livelihoods.

The American Institutes for Research (AIR) is conducting quasi-experimental impact evaluations to determine (a) the gendered impacts of the NFWP on financial inclusion, asset ownership, consumption, income, and decision-making power; and (b) the impact of the NFWP on group functioning and inclusiveness. The impact evaluations use a mixed-methods design that

includes a formative assessment, a process evaluation, and an assessment of the cost-effectiveness of the NFWP. This report presents the results of the impact evaluation after 1 year of program implementation. Figure 1 summarizes the evaluation methods and the timeline.

Figure 1. Methodological Approach and Evaluation Timeline
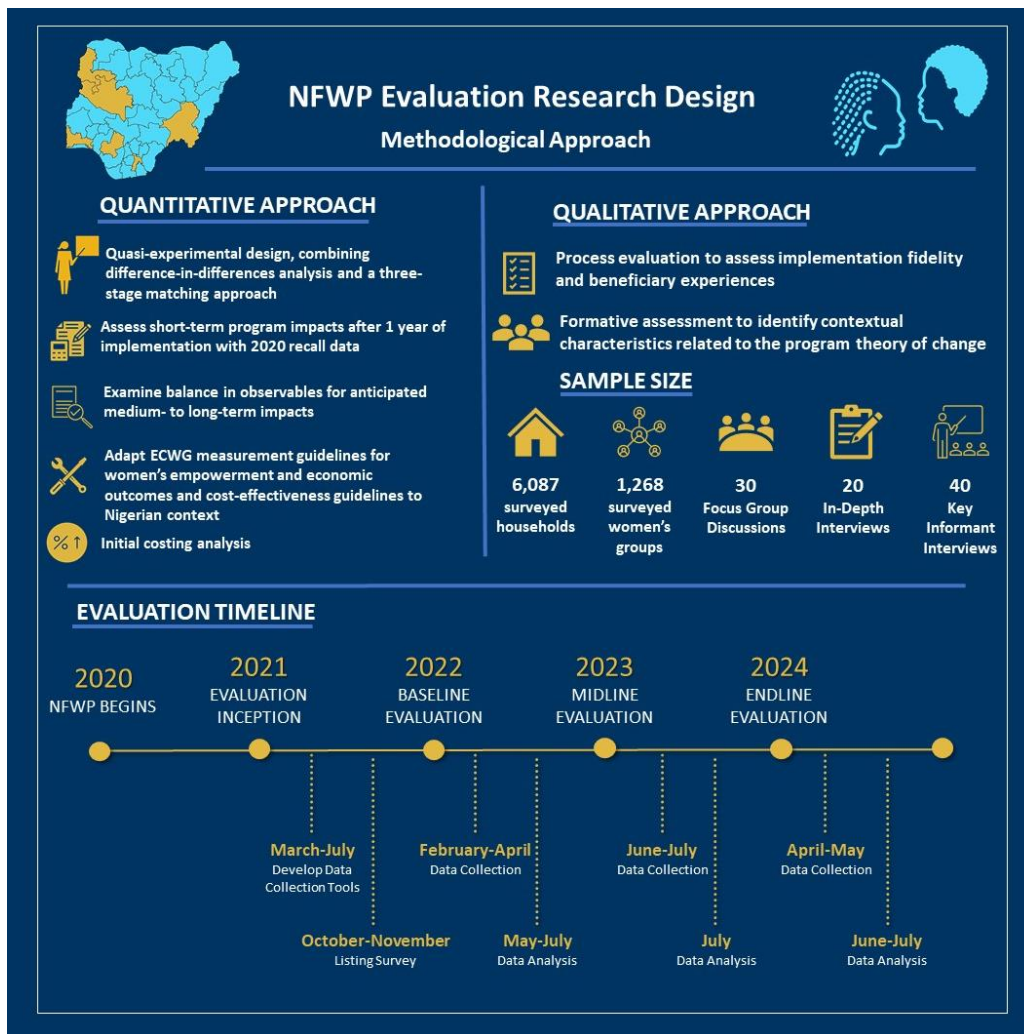
Program Description and Theory of Change

The NFWP uses WAGs as platforms for providing opportunities in four areas: the development of social capital, the building of livelihoods, the creation of partnerships, and messaging about gender and other social norms. The NFWP is being implemented gradually, starting with six states and 19 local government areas (LGAs). During the first phase of the program, the NFWP seeks to expand the participation of women in women's groups by reaching 324,000 women through approximately 21,600 WAGs in Ogun, Taraba, Kebbi, Abia, Niger, and Akwa Ibom states, located in each of the six geopolitical zones of the country. A scale-up operation will

expand the model to additional States based on a set of eligibility criteria over the next five years.

Program Description

The NFWP implementation focuses on the formation and strengthening of WAGs through a five-phase process (see Figure 2).

Figure 2. Formation and Strengthening of WAGs

WAG FORMATION	TRAINING PHASE	INTENSIVE PHASE	DEVELOPMENT PHASE	MATURITY PHASE
• Women organize themseves in new WAGs • Existing groups choose to become WAGs	• Group formation training • Savings training	• Savings and loans ongoing • Financial education training • Business skills training • Preparation of business plans for individual grants for IGAs	• Savings and loans continue • Gender/life skills training • Preparation of business plans continues • Individual grants for IGAs	• Savings and loans continue • Formation of livehoods collectives and grants

Source: Community Operating Manual (COM) – Nigeria For Women Project

Throughout these five phases, groups receive trainings in savings and credit, financial literacy, gender and life skills, and business skills. In addition, the program selects WAG members to receive individual grants to create or expand their income-generating activities, after the development of business plans. The program also aims to create and expand livelihoods collectives to form livelihoods partnerships at the end of the formation process. The livelihoods collectives are eligible to receive grants from the project. Concurrently, the program carries out a series of activities related to behavior change and awareness raising, targeted to gatekeepers as well as all women and men in the community, with the aim of influencing social norms and gender beliefs and behaviors at the community level. At the time of the baseline data collection (after one year of implementation), the program had started preparatory activities and

Impact Evaluation of the Nigeria for Women Project: Baseline Report

commenced with the training phase (with a strong focus on savings and credit) in five States.[1] The timing of the implementation differs slightly by State and LGA, however, with Ogun State moving most rapidly after the start of the NFWP. Some LGAs had already made a start with the trainings linked to the intensive and development phase shortly before the baseline data collection.

Theory of Change

The theory of change presents several pathways through which the NFWP can achieve improvements in women's economic empowerment. Trainings focused on savings can lead to increases in individual and collective savings, which can in turn enable women to gain access to credit. Combining increased access to credit with improved business skills can then help women start or expand their income-generating activities. Livelihoods grants and group credit could enable women to invest in existing or new enterprises (individual or collective), which could generate profits and additional household income. These enterprises could also enable women to accumulate assets, particularly once livelihoods trainings and the preparation of business plans have helped them to make productive investments across diversified agricultural and non-agricultural income-generating activities. The creation and expansion of women's individual and collective businesses could increase women's income and asset ownership, which could stimulate their bargaining power within the household.

Methods

Quasi-Experimental Analysis

We used a quasi-experimental design to determine the impact of the NFWP after one year of implementation. Specifically, we used a difference-in-differences (DID) analysis (which compares the average change over time for the treatment group with the average change over time for the comparison group) and a three-stage matching approach for outcomes for which we had recall data available, and a three-stage matching approach for variables for which we did not have recall data available (with a single-difference comparison between the treatment and comparison group instead of a DID analysis).

State governments, in consultation with the Federal Government and the World Bank selected 15 treatment LGAs to implement the project before the start of the baseline survey. They selected these LGAs based on preferences of State governments.[2] We matched each of the 15

---

[1] Implementation started later in Akwa Ibom. It was not yet certain that Akwa Ibom would join the NFWP at the time of the baseline data collection. However, we collected baseline data in Akwa Ibom in August 2022. We will separately analyze those data because of the different implementation timeline

[2] At the time of launching the evaluation, there were only 5 active States in NFWP, each with 3 LGAs. Presently there are 6 States in the NFWP and the evaluation has been expanded to 18 LGAs.

treatment LGAs to four neighboring comparison LGAs. We then selected four potential neighboring comparison LGAs for each treatment LGA, so that we achieved a sample size of 15 treatment LGAs and 60 comparison LGAs. We selected 60 comparison LGAs that in expectation were most similar to the treatment LGAs because they were neighboring LGAs. In total, we collected data from 1,208 treatment households and 4,879 comparison households, and from 261 treatment women's groups and 1,007 comparison women's groups in 15 treatment and 58 comparison LGAs. We replaced two comparison LGAs because they were too unsafe for data collection.

The key identifying assumption is that the trends in the outcomes in the treatment and comparison LGAs were similar before the start of the NFWP. We believe this assumption is credible because the treatment and comparison LGAs coped with similar shocks and had similar characteristics before the start of the program. We provide further evidence for the validity of the assumption by analyzing recall data before the start of the program.

We used the measurement guide developed by the Evidence Consortium on Women's Groups (ECWG; de Hoop et al., 2020), led by AIR, to aid in designing the survey instruments. The main outcome measures included individual- and household-level savings and credit, income, and asset ownership; household-level consumption; and women's empowerment measures, based on an adaptation of the Women's Empowerment in Agriculture Index. In addition, we included group-level measures related to the composition of groups, the number of group meetings, and group-level savings.

Costing Analysis

For the costing analysis, we estimated the costs of implementing the program for the initial 2 years: 2020 and 2021. We collected budgeted and actual expenditures for each program component from each of the five State Project Coordinating Units (SPCUs) and the Federal Project Coordinating Unit (FPCU), and we collected data on the number of project beneficiaries (in terms of WAG members) in 2020 and 2021 from the project management information systems.

Qualitative Analysis

The qualitative research has two main components: a formative assessment and a process evaluation. We used rapid qualitative assessment (RQA) techniques to probe and gain insight into emerging practices, trends, and dynamics that warrant deeper inquiry. Further, we present *preliminary findings* for the process evaluation; the full process evaluation will follow at midline.

Results

In line with the theory of change, we found a 12 percentage points increase in the likelihood of savings and a 9 percentage points increase in the likelihood of receiving credit, along with some indications of positive impacts of women's ownership of smaller assets. While we found positive impacts on the likelihood of savings, and on savings in women's groups and other savings groups, we did not find positive effects on cumulative savings. We hypothesize that this is because women may have invested some of their initial savings in smaller assets. We found evidence for positive impacts on women's sole ownership of mechanized farm equipment (by 1 percentage point), uncultivated land (by 3 percentage points), and mobile phones (by 5 percentage points). However, we also found some indications for negative effects on women's ownership of livestock (by 1 percentage point).

Overall, the results suggest an increase in the consumption of assets relative to other household consumption, but we did not find positive impacts on an asset index. We found positive impacts on the consumption of large and small consumer durable goods, and especially on mobile phone expenditures. These positive impacts, combined with the evidence of increased sole ownership among women of non-mechanized farm equipment and mobile phones, indicate a substitution of food consumption toward the consumption of smaller assets for women. However, given that we did not find positive impacts on an asset index, the program may only have had some initial short-term effects on asset ownership.

We also did not find short-term impacts on women's income and decision-making power or on total consumption. We found no consistent statistically significant differences between the treatment and comparison groups for these outcomes, which aligns with some evidence from India suggesting that women's groups with economic objectives may not always have positive short-term effects on women's income and decision-making power or on household-level expenditures (Hoffmann et al., 2021; Kochar et al., 2020). Midline results will indicate whether the NFWP is generating positive effects on these outcomes after the introduction of livelihoods grants. We anticipate conducting the midline survey after approximately 2.5 years of program implementation from June-July in 2023.

A triangulation of quantitative and qualitative evidence suggests positive program effects on the likelihood of self-employment. We found statistically significant effects on self-employment among women respondents. Interviews with qualitative respondents revealed that investing savings in small businesses may have contributed to these positive effects. The positive effects on self-employment were exclusively driven by Muslim women. In addition, we found positive effects on the labor force participation of Muslim women, but not on the labor force participation of other women.

Various implementation factors may have contributed to the modest but positive short-term effects of the program on financial inclusion, sole ownership of small assets, and self-employment. While some respondents initially showed a lack of trust in the NFWP, 81 percent of the women in NFWP areas reported membership of a women's or savings group and 65% reported membership of a WAG at the time of the baseline survey. Respondents indicated that sensitization efforts of the NFWP (including the sensitization of local leaders such as village chiefs), as well as endorsements by the Federal government, contributed to building trust among WAG members. WAG members also indicated that the NFWP has had a consistent presence in their communities, leading them to view it as accountable and reliable. In addition, women's social networks encouraged their participation in WAGs. Spousal support was critical because a husband's permission is needed to join a WAG. Elder women also encouraged participation by serving as an important source of information.

WAGs also had different characteristics from women's groups and savings groups in the comparison group. WAGs had a smaller number of group members and were more likely to only include women. We did not find consistent evidence for greater inclusion of marginalized women's group members in WAGs, however. While WAG members were less likely to finish secondary school than members of other women's groups, they were also less likely to live with a disability. In addition, qualitative evidence indicates that women without an income faced challenges participating in WAGs because of the savings requirements, though WAGs were able to include women of the "missing middle" or economically active poor as members.

Per-capita program costs in the initial 2 years of the NFWP ranged from NGN18,441 in Taraba to NGN22,752 in Niger. Initial analyses of program costs indicate that the average costs per program participant were similar across the different states and comparable to the initial costs of program implementation for the JEEVIKA program in Bihar, India. Average costs of savings groups and other women's groups tend to decrease considerably when programs move to scale, however, as shown by a costing analysis of the JEEVIKA program in Bihar, India (Siwach et al., 2022). These cost savings are important because it may not be sustainable to scale the NFWP with the current average costs per program participant.

While program costs are relatively high, various stakeholders provided suggestions to further increase the intensity of programming. For example, respondents recommended increased trainings and a more intensive sensitization effort. These recommendations may have significant cost implications, suggesting that incorporating the suggestions by program staff will likely not enable a sustainable scale-up of the program.

On average, WAG members reported spending 22 hours in the year on group-related trainings. While trainings may lead to additional downstream impacts, they may also impose additional

costs on beneficiaries in terms of time away from income-generating work (although this may be minimal).

Limitations

The study also faces various limitations. First, it is important to recognize that we collected baseline data 1 year after the start of program implementation. As a result, we are not able to say with certainty that the treatment and comparison groups were similar before the start of the program (recall data indicate that the treatment and comparison groups were comparable before the start of program implementation, however). Second, this report only includes impact estimates for 1 year after the start of program implementation. As a result, the report is only able to present short-term effects of the NFWP. We will present longer-term effects during the midline study. Third, we had limited statistical power to detect heterogeneous effects. While our sample size is relatively large, our evaluation design includes a limited number of clusters, constraining our ability to detect heterogeneities in the impact estimates with sufficient precision. Finally, the qualitative interview protocols were designed primarily for the formative assessment rather than the process evaluation, as originally planned. As a result, the process evaluation findings generated at baseline should be considered preliminary. The process evaluation will be the primary focus of the qualitative data collection at midline.

Implications for Policy and Practice

Results after one year of program implementation suggest that the NFWP's sensitization efforts, women's social networks, and endorsements by the Federal government all likely contributed to increased WAG membership, and to the program's modest but positive effects on financial inclusion and women's asset ownership. These factors need continued support to achieve the positive effects of savings group and other women's group programming in Nigeria.

The results also indicate that religion is a key consideration when designing savings group programs such as the NFWP. Gender norms manifest as a result of both cultural and religious influences across Nigeria. When considering the NFWP's intended impact on community views toward gender, religion therefore merits attention, perhaps by including local religious leaders in messaging about gender norms. We also found positive impacts on labor force participation and self-employment for Muslim women, but not for other women.

It is important to assess the program costs of additional sensitization and training activities. Various stakeholders indicated the importance of sensitization and trainings for building trust in the NFWP and increasing WAG membership. Such activities come with additional costs, however, and at this moment it is unclear whether the NFWP can be sustainably scaled up without a reduction in program costs. Previous studies indicate that the scale-up of savings

group and other women's group programs with economic objectives can result in cost savings (Siwach et al., 2022). Nonetheless, in order to ensure that scale-up does not lead to a decline in implementation quality, it is critical to assess which program components are critical for achieving additional impacts of the NFWP without generating a significant increase in costs.

Further, the results suggest that while the NFWP may generate benefits for women who are close to the poverty line and are at risk of falling below the poverty line (or the "missing middle"), it may generate fewer benefits for poorer women without an income source. Women without an income source from wage labor or self-employment may therefore require cash transfers or graduation programs from social protection systems in Nigeria. These findings are aligned with the targeting strategy of the social safety net program in Nigeria and the NFWP. The social safety net program targets the poorest households in Nigeria with cash transfers. The NFWP focuses on a different target group ("the missing middle") than these social safety net programs.

At this moment, it is too early to present definitive conclusions about the impact and cost-effectiveness of the NFWP. We will continue to conduct research to determine the longer-term impacts and cost-effectiveness of the NFWP using the midline and endline surveys.

Impact Evaluation of the Nigeria for Women Project: Baseline Report

Introduction

Women in Nigeria continue to face societal and structural barriers to accessing savings, credit, livelihoods, and markets, constraining their opportunities and well-being. Traditional gender norms and other social norms, information asymmetries, and limited access to finance are among the primary factors that contribute to gender inequality across the country (World Bank, 2018). Data from the Demographic and Health Survey (DHS, 2018) show that in 2018, only 22.1% of women in Nigeria had an account at a financial institution. As a result, women are often unable to make future investments, limiting their ability to respond to shocks (World Bank, 2019). The same data show that in 2018, 29.5% of women between the ages of 15 and 49 reported having experienced intimate partner violence (DHS, 2018). Further, just 44% of married women made decisions, either alone or jointly with their husband, regarding their health care; and just 40% participated in decisions about major household purchases (DHS, 2018).

Savings groups and other women's groups with economic objectives, such as self-help groups and livelihood groups, have emerged as an important means of increasing women's economic empowerment and access to opportunities, including in various settings in sub-Saharan Africa (e.g., Blattman et al., 2016; Brody et al., 2017; Karlan et al., 2017; Desai et al., 2019; de Hoop et al., 2020; de Hoop et al., 2022). Members of savings groups commonly pool small weekly savings into a common fund, which members can then borrow against (Adegbite et al., 2022). "Women's groups" is an umbrella term referring to different models of economic, health, and community groups, including but not limited to savings groups, with a primarily female membership (Anderson et al., 2019). Membership in women's groups and informal savings groups is positively correlated with women's asset ownership and decision-making power in a nationally representative sample in Nigeria. The magnitude of the relationship is small, however, and it is unclear whether the relationship is causal (Meysonnat et al., 2022a; Meysonnat et al., 2022b).

One potential way to improve the effectiveness of savings groups is to transform them into more formalized savings groups with democratic decision-making. Informal savings groups and women's groups often have less democratic decision-making than formalized savings groups, as well as higher dropout rates among members (de Hoop et al., 2022).

The Government's Nigeria for Women Project (NFWP) aims to introduce such processes after establishing new women's groups and transforming existing informal savings groups and other women's or mixed-gender groups into formal, women-only groups, called Women Affinity

Groups (WAGs). WAGs are formal savings groups that follow established savings group implementation models (e.g., Village Savings and Loan Associations [VSLAs] and Savings and Internal Lending Committees) to support women's financial inclusion. The NFWP aims to improve women's livelihood opportunities and facilitate their access to economic markets by mobilizing women into WAGs. The NFWP uses WAGs as platforms for providing opportunities in four areas: the development of social capital, the building of livelihoods, the creation of partnerships, and messaging about gender and other social norms.

The American Institutes for Research (AIR) is conducting quasi-experimental impact evaluations to determine (a) the gendered impacts of the NFWP on financial inclusion, asset ownership, consumption, income, and decision-making power; and (b) the impact of the NFWP on group functioning and inclusiveness. The impact evaluations use a mixed-methods design that includes a formative assessment, a process evaluation, and an assessment of the cost-effectiveness of the NFWP. The evaluations will also include randomized controlled trials (RCTs) to identify impacts of individual program components. These components include messages related to social norms and the layering of interventions to reduce gender-based violence (GBV) and improve maternal health.[3]

This report presents impact evaluation results for the NFWP after 1 year of program implementation. We initially aimed to conduct a baseline survey before the start of the program, but this was delayed to February-March 2022 due to unforeseen circumstances. This report therefore has two objectives. First, we aim to assess program impacts 1 year after the start of program implementation for those outcomes that the NFWP can plausibly affect in the relatively short term. Second, we aim to examine balance in observable characteristics for those variables that the NFWP will not impact in the short term, according to the theory of change and based on results from various other impact evaluations of comparable, government-supported women's and savings group programs in India (e.g., Kochar et al., 2020; Hoffmann et al., 2021). These impact evaluations showed short-term impacts on financial inclusion and mixed evidence for positive effects on asset ownership and women's empowerment (Kochar et al., 2020; Hoffmann et al., 2021), but we do not hypothesize short-term impacts on women's income or consumption. We combine the results with a qualitative evaluation examining contextual characteristics associated with the program theory of change and the fidelity of program implementation. Finally, we provide initial estimates of the costs of the program, both to the implementer and to program participants. Figure 3 summarizes the evaluation methods and timeline.

---

[3] This report does not discuss these RCTs because they have not started yet.

Figure 3: Summary of Evaluation Methods and Timeline
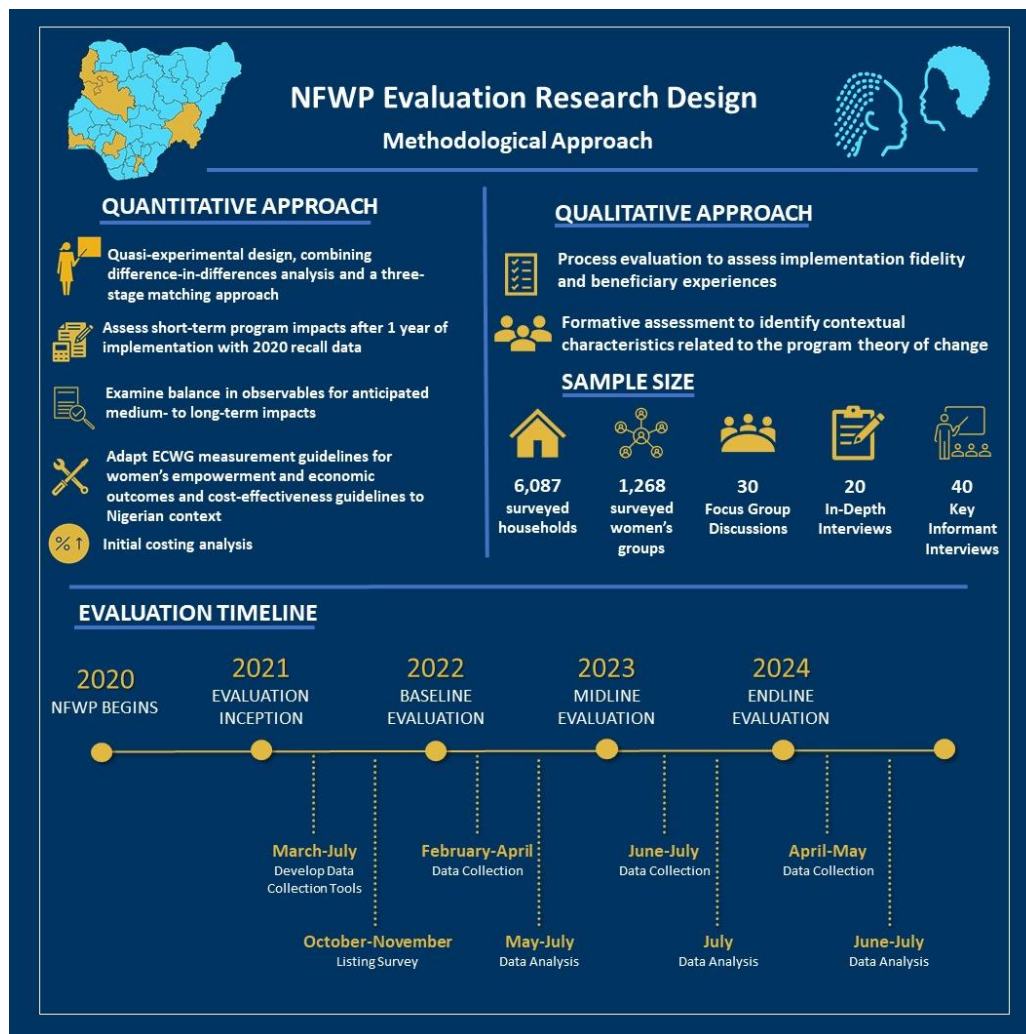
The rest of this report is structured as follows. We start by describing the program and the underlying theory of change. We then present a summary of the impact evaluation methods, followed by a discussion of the results of the quasi-experimental study. Next, we present the results of the qualitative study, followed by the results of the costing analysis. We finish the report by triangulating the results and presenting a conclusion and initial recommendations for policy and practice.

Program Description

The NFWP is a Federal program (supported by a $100 million loan from the World Bank) that is being implemented gradually, starting with six states and 19 local government areas (LGAs). During the first phase of the program, the NFWP seeks to expand the participation of women in

women's groups by reaching 324,000 women through approximately 21,600 WAGs in Ogun, Taraba, Kebbi, Abia, Niger, and Akwa Ibom states, located in each of the six geopolitical zones of the country. Additional scale-up financing will expand the project to other eligible States of Nigeria over the next five to ten years.

The NFWP works with new and existing women's groups for women over the age of 18, and it targets women who are considered part of the "missing middle." These women live close to the poverty line and are therefore vulnerable to shocks, which may bring them below the poverty line in the absence of opportunities to mitigate the negative effects of shocks.

Recent studies on women's groups and COVID-19 indicate that participation in savings groups and other women's groups could limit the negative consequences of shocks, such as COVID-19, because participation could enable women to make use of past savings and access credit to cope with negative shocks (Adegbite et al., 2022; Anderson et al., 2022; Walcott et al., 2021). Economically active women could especially benefit from participation in savings groups and women's groups, because they are able to contribute to savings which can limit their vulnerability to negative shocks. Once women are part of a WAG, they also can access knowledge and resources available within the group, such as individual livelihoods grants and trainings, which can help women start or further expand their economic activities.

It is more challenging for women who are not economically active to participate in savings groups because participation in such groups requires regular savings. Programs such as cash transfers or graduation programs may bring larger benefits for economically inactive women who face challenges contributing to savings regularly.

The first year of NFWP implementation focuses on the formation and strengthening of WAGs through a five-phase process (Figure 4). Throughout these five phases, groups receive trainings in savings and credit, financial literacy, gender and life skills, and business skills. In addition, the program plans to either provide individual grants or revolving loan funds to individual WAG members to create or expand their income-generating activities. The program may also create and expand livelihoods collectives to form livelihoods partnerships at the end of the formation process. These livelihood collectives may receive grants from the project. Concurrently, the program carries out a series of activities related to behavior change and awareness raising, targeted to gatekeepers as well as all women and men in the community, with the aim of influencing social norms and gender beliefs and behaviors at the community level. At the time of baseline data collection from February-March 2022, the program had started preparatory activities and commenced the training phase (with a strong focus on savings and credit). The exact implementation model will likely change during the scale-up process in early 2023.

The program will also introduce health and GBV layering activities to leverage WAGs as platforms for delivering health and GBV programming to many women at once. This approach is in line with the suggestion by Dìaz-Martin and colleagues (2022) that groups could deliver benefits at a lower cost per program participant than programs that focus on delivering health information at the individual level. While the World Bank and the Nigerian Ministry of Women and Social Development (MOWASD) have not yet made decisions about the specific contents of the health programming, the layering activities will likely focus on maternal and child health and may include an existing curriculum developed by the Nigerian Ministry of Health, or may adapt successful programs implemented by nongovernmental organizations.

While the initial phase of the program seeks to provide lessons for potential future expansions, implementation does vary across the six states. The states and LGAs differ significantly and each State government adapts programming to suit its specific context, particularly around social norms. First, entry points for mobilizing women differ by region because of varying social and cultural norms. For example, in the northern regions, women have few opportunities to socialize outside the home other than through informal networks and social ceremonies. The program leverages these informal networks and social ceremonies to find local leaders who can support WAGs. In southern Nigeria, however, the program can approach women through formal networks—for example, through community leaders or existing women's groups. Second, WAGs may set different priorities. While all WAGs engage in saving money and offering credit, savings and credit amounts differ by WAG. Because of differences in familiarity with women's groups and opportunities to participate in such groups in the past, women may also have varying levels of confidence, skills, and understanding of concepts related to banking, savings, credit, organization, and other economic matters (Desai et al., 2018).

Figure 4. Formation and Strengthening of WAGs

WAG FORMATION	TRAINING PHASE	INTENSIVE PHASE	DEVELOPMENT PHASE	MATURITY PHASE
• Women organize themseves in new WAGs • Existing groups choose to become WAGs	• Group formation training • Savings training	• Savings and loans ongoing • Financial education training • Business skills training • Preparation of business plans for individual grants for IGAs	• Savings and loans continue • Gender/life skills training • Preparation of business plans continues • Individual grants for IGAs	• Savings and loans continue • Formation of livehoods collectives and grants

Source: Community Operating Manual (COM) – Nigeria For Women Project

Theory of Change

To inform our study design, we developed a theory of change (ToC) in consultation with the World Bank and the NFWP Implementation Team at the national and State levels focused on the implementation of WAGs and the impact of those groups on group-, household-, individual, and community-level outcomes (see Figure 5 on page 8). The theory of change presents the pathways through which the hypothesize changes could occur. We link these pathways to the NFWP components targeted at the group level (Savings and Livelihoods Trainings, Social Norms Messages and Trainings, Health and GBV Layering) and at the community level (Social Norms Messages). In this section, we present an abbreviated description of the ToC, with a complete description added in Appendix A.

The NFWP can achieve improvements in women's economic empowerment in labor and financial markets through several mechanisms. Savings and financial literacy trainings can lead to greater savings, which can enable women to gain access to individual or group credit. At the same time, trainings related to livelihood development can improve women's business skills.
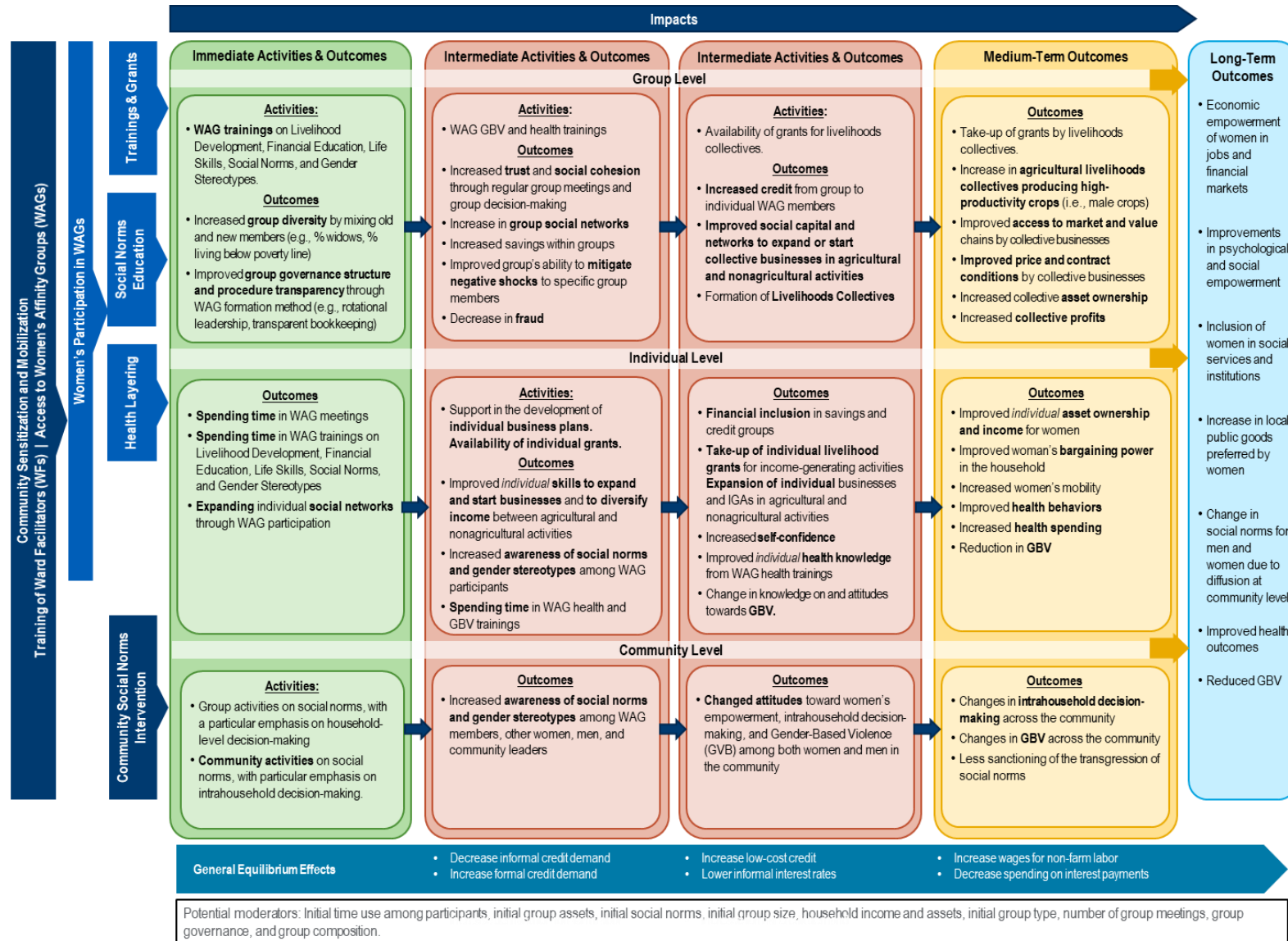
With increased access to credit and improved business skills, as well as additional support through livelihood grants and private sector partnerships, women can expand their income-generating activities, ultimately earning higher incomes and owning more assets, which could stimulate their bargaining power within the household. We hypothesize that these effects are amplified by the formation and expansion of women's livelihoods collectives, including producer organizations, cooperatives, or producer-governed private limited companies.

Social norm messages at the group- and community-level, combined with increased social support within groups can help WAG members reflect on gender norms, change group members' attitudes, and improve women's self-confidence and well-being. These changes in women's attitudes could result in improvements in women's bargaining power within the household. Social norm messages at the community level could lead to additional attitudinal changes and changes in social and gender beliefs and practices held by men and other non-WAG members. In addition, through specific focus on raising awareness about gender stereotypes and gender-based violence, social norms trainings could bring additional improvements in women's social empowerment.

The generation of social capital through group support could improve trust, social networks, and social cohesion by enabling women to collaborate with peers in their community and discuss economic and social issues. Improved social cohesion and networking could have both individual- and group-level effects on empowerment. The sense of confidence, dignity, and self-esteem that comes from women's empowerment could help women achieve "power within" and could positively influence intra-household dynamics—for example, by enabling women to gain bargaining power within the household with regard to financial decision and decisions on education and health.

Based on the program activities and pathways, as well as evaluations of comparable programs in India (e.g., Hoffmann et al., 2021; Kochar et al., 2020), we hypothesize positive impacts on savings, credit, and asset ownership in the shorter-term (after one year of program implementation). In addition, there is some evidence that the program may have short-term impacts on women's decision-making power (Brody et al., 2017). On the other hand, we expect that the program would generate positive impacts on consumption, income, or larger assets after the disbursement of livelihood grants, formation of livelihood collectives, and after women have had a chance to invest their savings and credit into income generating activities. Therefore, these longer-term effects are unlikely to be observed within a year, which is aligned with evidence from evaluations of comparable programs (e.g., Hoffmann et al., 2021; Kochar et al., 2020). Finally, we do not hypothesize positive impacts on social norms in the short-term because these usually require a longer time period to change.

Figure 5. Theory of Change

Impacts

Community Sensitization and Mobilization
Training of Ward Facilitators (WFs) | Access to Women's Affinity Groups (WAGs)

Women's Participation in WAGs

Trainings & Grants | Social Norms Education | Health Layering | Community Social Norms Intervention

Immediate Activities & Outcomes	Intermediate Activities & Outcomes	Intermediate Activities & Outcomes	Medium-Term Outcomes	Long-Term Outcomes

Group Level

Immediate Activities & Outcomes

Activities:
- WAG trainings on Livelihood Development, Financial Education, Life Skills, Social Norms, and Gender Stereotypes.

Outcomes
- Increased **group diversity** by mixing old and new members (e.g., % widows, % living below poverty line)
- Improved **group governance structure and procedure transparency** through WAG formation method (e.g., rotational leadership, transparent bookkeeping)

Intermediate Activities & Outcomes

Activities:
- WAG GBV and health trainings

Outcomes
- Increased **trust** and **social cohesion** through regular group meetings and group decision-making
- Increase in **group social networks**
- Increased savings within groups
- Improved group's ability to **mitigate negative shocks** to specific group members
- Decrease in **fraud**

Intermediate Activities & Outcomes

Activities:
- Availability of grants for livelihoods collectives.

Outcomes
- **Increased credit** from group to individual WAG members
- **Improved social capital and networks to expand or start collective businesses in agricultural and nonagricultural activities**
- Formation of **Livelihoods Collectives**

Medium-Term Outcomes

Outcomes
- Take-up of grants by livelihoods collectives.
- Increase in **agricultural livelihoods collectives producing high-productivity crops** (i.e., male crops)
- Improved **access to market and value chains** by collective businesses
- **Improved price and contract conditions** by collective businesses
- Increased collective **asset ownership**
- Increased collective **profits**

Long-Term Outcomes
- Economic empowerment of women in jobs and financial markets
- Improvements in psychological and social empowerment
- Inclusion of women in social services and institutions
- Increase in local public goods preferred by women
- Change in social norms for men and women due to diffusion at community level
- Improved health outcomes
- Reduced GBV

Individual Level

Outcomes
- **Spending time** in WAG meetings
- **Spending time** in WAG trainings on Livelihood Development, Financial Education, Life Skills, Social Norms, and Gender Stereotypes
- **Expanding** individual **social networks** through WAG participation

Activities:
- Support in the development of **individual business plans. Availability of individual grants.**

Outcomes
- Improved *individual* **skills to expand and start businesses** and **to diversify income** between agricultural and nonagricultural activities
- Increased **awareness of social norms and gender stereotypes** among WAG participants
- **Spending time** in WAG health and GBV trainings

Outcomes
- **Financial inclusion** in savings and credit groups
- **Take-up of individual livelihood grants** for income-generating activities **Expansion of individual** businesses and IGAs in agricultural and nonagricultural activities
- Increased **self-confidence**
- Improved *individual* **health knowledge** from WAG health trainings
- Change in knowledge on and attitudes towards **GBV.**

Outcomes
- Improved *individual* **asset ownership and income** for women
- Improved woman's **bargaining power** in the household
- Increased women's mobility
- Improved **health behaviors**
- Increased **health spending**
- Reduction in **GBV**

Community Level

Activities:
- Group activities on social norms, with a particular emphasis on household-level decision-making
- **Community activities** on social norms, with particular emphasis on intrahousehold decision-making.

Outcomes
- Increased **awareness of social norms and gender stereotypes** among WAG members, other women, men, and community leaders

Outcomes
- **Changed attitudes** toward women's empowerment, intrahousehold decision-making, and Gender-Based Violence (GVB) among both women and men in the community

Outcomes
- Changes in **intrahousehold decision-making** across the community
- Changes in **GBV** across the community
- Less sanctioning of the transgression of social norms

General Equilibrium Effects
- Decrease informal credit demand
- Increase formal credit demand
- Increase low-cost credit
- Lower informal interest rates
- Increase wages for non-farm labor
- Decrease spending on interest payments

Potential moderators: Initial time use among participants, initial group assets, initial social norms, initial group size, household income and assets, initial group type, number of group meetings, group governance, and group composition.

Research Questions

The evaluation of the NFWP is guided by various key research questions, which are based on the theory of change. In this report we focus on the following research questions.

Impact Evaluation Questions

1. What is the impact of the NFWP on group-level outcomes?
   a. What is the impact of the NFWP on collective savings and credit?
   b. What is the impact of the NFWP on group composition?
   c. What is the impact of the NFWP on group meetings and governance?
   d. What is the impact of the NFWP on social cohesion in groups?
2. What is the impact of the NFWP on individual-level and household-level outcomes?
   a. What is the impact of the NFWP on financial inclusion (access to formal credit, and savings)?
   b. What is the impact of the NFWP on economic outcomes (women's and household asset ownership, and income and household consumption)?
   c. What is the impact of the NFWP on women's empowerment and agency (women's decision-making power, women's mobility, women's self-confidence)?

Cost Questions

3. What are the costs of the implementation of WAGs in Nigeria?

Formative and Process Evaluation Questions

4. What are the barriers and facilitators to participating in women's groups and specifically WAGs?
5. How have contextual factors influenced program implementation?
6. What are the perceptions of WAG and other women's group participants and nonparticipants about the benefits and costs of participating in women's groups?
7. How do women's group participants and nonparticipants perceive community gender norms and their relationship with WAGs?
8. How do WAGs interact with nonmembers, including men and Ward facilitators?

The following section details our approach to addressing these research questions using a mixed-methods approach.

Impact Evaluation of the Nigeria for Women Project: Baseline Report

Methods

In the following sections, we describe the details of our mixed-methods evaluation design, starting with the quantitative methods, including the quasi-experimental design and the costing analysis, followed by the qualitative methods, including the formative assessment and the process evaluation.

Quantitative Study Design

We designed the impact evaluation based on the proposed program rollout, as well as methodological principles that enable us to establish a counterfactual. Establishing a counterfactual requires rigorous methodologies to address the following question: What would have happened in the absence of the intervention?

We are using a quasi-experimental design to determine the impact of the NFWP. We first used a three-stage matching approach to select a comparison group. We then combined the matching approach with a difference-in-differences (DID) analysis (which compares the average change over time for the treatment group with the average change over time for the comparison group), and a single-difference analysis for variables for which we did not have recall data available.[4] We assumed that we will not find impacts after 1 year of program implementation for indicators that are not plausibly affected by the program in the short term (i.e., consumption, income, agricultural production, and larger assets). For these indicators, this baseline report presents balance tables comparing the treatment and comparison groups. We assume that these balance tables provide an indication of differences between treatment and comparison groups that are similar to those that existed before the start of the program. We will use DID analysis to determine impacts on these outcomes in the midline and endline reports. We also present tables comparing recall data to validate the assumption that the treatment and comparison group were similar before the start of the program.

To ensure that we compare similar treatment and comparison households or groups, we designed and implemented a three-stage matching process. For our analysis, we define the treatment household as individuals and households residing in LGAs targeted by the NFWP. Within each state, we first selected comparison LGAs neighboring the treatment LGAs. In the second stage, using information from the listing survey, we matched treatment households and groups to households and groups in comparison enumeration areas (EAs) based on background characteristics and variables relevant to the targeting and potential outcomes of the program.

---

[4] We also designed RCTs for components that will be added to the NFWP at later stages, such as the layering of social norm messages and health and GBV interventions. However, this baseline report does not focus on these RCTs because the NFWP has not yet started the layering of social norm messages and health and GBV interventions.

In the third stage, we refined the second-stage matching using data from the full baseline survey.

In the first stage, we matched each of the 15 treatment LGAs to four neighboring comparison LGAs. We worked with the World Bank, the Nigerian government, and Oxford Policy Management (OPM, the firm collecting the baseline data) to select four potential neighboring comparison LGAs for each treatment LGA, so that we achieved a sample size of 15 treatment LGAs and 60 comparison LGAs. Within each LGA, we then implemented a listing survey in 20 enumeration areas to collect data on a small but important number of variables for matching purposes (we conducted a listing survey in a total of 1,500 enumeration areas across the five states). During the listing survey, we dropped two of the 60 comparison LGAs because OPM considered them too insecure for data collection. None of the treatment LGAs were deemed unsafe, suggesting that the matching likely improved because of this selection rule (in addition to the need to ensure safety for the enumerators).

In the second stage, we matched 1,208 treatment households in 265 EAs in 15 treatment LGAs to 4,879 households in 964 EAs in 58 comparison LGAs, based on data from the listing survey. We conducted the matching (without replacement) using a nearest neighbor approach, in which we matched each treatment household to a minimum of one comparison household and a maximum of four comparison households, based on data from the listing survey. We then conducted a baseline survey with these 6,087 households from February-March 2022.

In the third stage, we revisited the household-level matching using a rich set of individual- and household-level data from the baseline survey. We re-matched the 1,208 treatment households in 256 EAs to 3,591 comparison households in 899 EAs. We used individual- and household-level characteristics, as well as recalled values of outcome variables, for the matching. The set of variables we matched on included women's demographics (age, religion, language, level of education, and marital status), women's recalled indicators for having savings and loans, woman's recalled amount of savings and loans, women's recalled labor force participation and income diversification, a household-level asset index based on recall data, and various other individual- and household-level variables. We did not use a matching approach for the impact analysis at the group level because we lacked recalled values of outcomes, and because the survey collected information on just a few exogenous (background) group-level characteristics. Instead, we included these characteristics as control variables in our regressions. We controlled for the age of the group and whether the group was formally registered.

Lastly, we combined DID analysis and matching in two ways as a robustness check. First, we used nearest neighbor matching to identify the comparison group and then compared changes in outcomes over time between NFWP and untreated women and households using a DID

analysis. Second, we used a different matching approach by pursuing a reweighting approach in which we weighted comparison observations to impose the same distribution of covariates for the treated and untreated (Abadie, 2005). Both matching methods aim to correct for any pre-treatment imbalances between the outcomes of treatment and comparison units. The latter method aims to correct for any pre-treatment imbalances between the outcomes of treatment and comparison units. Although we used both methods, the impact estimate tables in this report show only the weighted estimations. The point estimates, standard errors, and levels of statistical significance are very similar between the unweighted and weighted estimations, suggesting that the findings are robust to different matching methods.[5]

Sampling

Since we lacked sufficient roster data on households, women, and groups in the treatment and comparison LGAs, we worked with OPM to conduct a listing exercise in each of the treatment and comparison LGAs, followed by a sampling of groups, households, and women in these communities. The listing survey helped us locate all women's groups and households with women who were at least 18 years old. This approach enabled the research team to identify all women eligible for enrolling in the groups, and allowed for the collection of basic descriptive data on a small number of outcome variables for each of the women's groups and households. We listed 54,215 households and 4,211 women's groups in a total of 1,500 EAs across the 73 LGAs in this study.

As indicated earlier, we next sampled comparison households that were similar in observable characteristics to the treatment households, using a matching approach. Of the 54,215 households listed, we selected a total of 6,087 households (1,208 in the treatment group and 4,879 in the comparison group) to complete the baseline survey, based on the matching.

To select women's groups, we first matched treatment and comparison EAs based on a small number of variables, including average household size, the average age of women, the percentage of the population who were female in different age groups (18–25, 26–35, 36–45, 46–59, and over 60 years old), and the percentage of households with a refrigerator. Next, we randomly selected women's groups within EAs for which the propensity scores overlapped between the treatment and the comparison group (the observations within the common support). Of the 4,211 groups listed, we selected 1,268 groups (261 in the treatment group and 1,007 in the comparison group) and their leaders to complete the group baseline survey, based on the matching. Of the 4,211 groups listed, we selected 1,268 groups (261 in the treatment group and 1,007 in the comparison group) and their leaders to complete the group baseline survey, based on the matching.

---

[5] Tables showing weighted and unweighted impact estimations are available upon request.

Our power calculations suggest that we have an 80% chance of detecting effect sizes between 0.20 and 0.28 standard deviations for household- and individual-level outcomes, and effect sizes between 0.24 and 0.28 standard deviations for group-level outcomes. Appendix E presents details on the power calculations and associated assumptions.

Outcome Measures

We used the measurement guide developed by the Evidence Consortium on Women's Groups (ECWG; de Hoop et al., 2020), led by AIR, to aid in designing the survey instruments. Table 1 presents the main outcome measures we used for this baseline report. These include individual- and household-level savings and credit, income, and asset ownership; household-level consumption; and women's empowerment measures, based on an adaptation of the Women's Empowerment in Agriculture Index. In addition, we included group-level measures related to the composition of groups, the number of group meetings, and the group-level savings.

Table 1. Outcome Measures and Survey Tools

Outcome measure	Survey questions
Women's savings	Social Observatory Questions on savings
Women's access to credit	Social Observatory Questions on credit
Women's and household's asset ownership	Individual- and household-level sole or joint asset ownership based on Women's Empowerment in Agriculture Index
Women's income	Self-reported income from wage employment and self-employment
Household expenditures	Expenditure surveys
Women's economic empowerment	Women's Empowerment in Agriculture Index
Women's psychological empowerment	New General Self-Efficacy Scale
Women's ability to engage in collective action	Social Observatory Questions
Group size	Survey questions based on the ECWG measurement guide
Group composition	Survey questions on the gender, age, demographic characteristics, and economic indicators of group members based on the ECWG measurement guide
Number of group meetings	Survey questions based on the ECWG measurement guide
Collective savings	Survey questions based on the ECWG measurement guide

Analysis

We assessed the impact of the WAG model using a DID regression framework while controlling for time-varying individual and group characteristics. We compared changes in outcomes over time between households, individuals, and groups in treatment LGAs and households, individuals, and groups in comparison LGAs. DID analysis entails calculating the change in outcomes between the recall value of the baseline survey and the current value of the baseline survey for treatment and comparison group units and comparing the magnitude of these changes between the treatment and the comparison groups. We used cluster-robust standard errors to account for a lack of independence across observations due to clustering of households, individuals, and groups because the intervention is allocated at the LGA level. Two key features of the DID estimator are particularly attractive for deriving unbiased program impacts. First, using pre- and post-treatment measures enables us to "difference" out unmeasured fixed (i.e., time-invariant) household and group characteristics that may affect outcomes, such as education level and household composition. The approach also enables us to benchmark the change in the indicator against its value in the absence of treatment. Second, using the change in a comparison group as a counterfactual enables us to account for general trends in the value of the outcome. The key assumption underpinning the DID approach is that there is no systemic, unobserved, time-varying difference between the treatment and comparison groups.

While we cannot examine this assumption in detail, we believe it is credible because treatment LGAs and their corresponding comparison LGAs are geographic neighbors within the same state, and likely coped with similar shocks and had similar characteristics before the start of the program. The Nigerian government and the World Bank jointly selected the 15 treatment LGAs to implement the NFWP during its pilot phase. They selected these LGAs based on preferences of State governments. For this reason, we selected 60 comparison neighboring LGAs that in expectation were most similar to the treatment LGAs. We provide further evidence for the validity of the impact evaluation design by analyzing differences in recall data between the treatment and the comparison group before the start of the program.

The DID design provides us with the intent-to-treat (ITT) effect of the WAG model; in other words, the average treatment effect for those women or groups assigned to a treatment condition regardless of take-up of treatment. We further assessed the treatment effect on the treated; that is, we evaluated the effect of the program for those households and individuals in treatment LGAs that became WAG members.[6]

We also created a limited number of subgroups to assess heterogeneous effects. For example, we examined heterogeneous effects for women in different age categories and women of a

---

[6] This report only presents ITT estimates because the Local Average Treatment Effect (LATE) estimates are very similar to the ITT estimates.

different religion. In addition, we examined heterogeneous effects for various moderators that we identified in the theory of change.

Cost-Effectiveness Analysis

We estimated the costs of implementing the program for the initial 2 years: 2020 and 2021. To compute costs per program participant, we used two data sources: We collected budgeted and actual expenditures for each program component from each of the five State Project Coordinating Units (SPCUs) and the Federal Project Coordinating Unit (FPCU), and we collected data on the number of project beneficiaries (in terms of WAG members) in 2020 and 2021 from the project management information systems.

In addition to program implementation costs, we estimated the costs of participating in WAGs for the women. Women's groups often involve time costs for the women, as well as out-of-pocket expenses for group-related activities. Through the baseline surveys, we collected data on participation in group-related activities including group meetings and trainings, time and frequency of participation, and any other transaction costs borne by the women including expenses related to transportation or meals. We calculated these costs for both NFWP and non-NFWP women to estimate the additive costs of the NFWP. We translated time costs into economic costs using women's hourly income, based on their income in the last 4 weeks, which we collected in the income module of the survey.

Qualitative Study Design

We combined the impact evaluations with rigorous qualitative research, which had two main components: a formative assessment and a process evaluation. While the research design originally intended for the process evaluation to occur at midline, delays in data collection made it possible to collect early, process-related data that may be useful to the implementers at this stage. This report therefore presents *preliminary findings* for process evaluation outcomes; the full process evaluation will follow at midline.

We used rapid qualitative assessment (RQA) techniques to probe and gain insight into emerging practices, trends, and dynamics that warrant deeper inquiry. These qualitative insights informed our development of measurement instruments for the impact evaluation.

Formative Assessment

The objectives of the formative assessment were as follows:

1. Gain a deeper understanding of the local policies, social norms, expectations, and other contextual dynamics that shape women's experiences and perceptions of women's groups, including WAGs.

2. Collect and analyze data that will help inform and refine the impact evaluations' theory of change, measurement instruments, and overall design.

We used three main data collection methods for the formative assessment:

1. Focus group discussions (FGDs) with WAG members, former and non-WAG members, and spouses of women who are in WAGs. The FGDs include a participatory livelihoods assessment (PLA).
2. In-depth interviews (IDIs) with women's group members and non-WAG women in the community, including a financial diaries exercise.
3. Key informant interviews (KIIs) with staff from the World Bank as well as Federal, State, LGA-level, and ward level officials.

Sampling Approach

We sampled respondents from all 5 states where the program was implemented at the time of the data collection—Abia, Kebbi, Niger, Ogun, and Taraba. In each state, we purposively selected 2 LGAs and 1 ward per LGA. Of the two wards sampled in each state, one was an urban environment and one was more rural. In total, we collected qualitative data from **10 wards across the five states** (Table 2). The research team selected the qualitative sample in conjunction with OPM and validated the selection with the World Bank.

Table 2. FGD and IDI Sampling for the Formative Assessment and Process Evaluation

Respondents	Formative research sampling
Ward level (10 selected wards)	
Women's group members	• 10 FGDs (one FGD in each of the 10 wards) – PLA component
	• 10 IDIs (one IDI in each of the 10 wards) – Financial diaries
Women in the community who were former women's group members, or had never been members of a women's group	• 10 FGDs (one FGD in each of the 10 wards) – PLA component
	• 10 IDIs (one IDI in each of the 10 wards) – Financial diaries
Spouses of women's group members	• 10 FGDs (one FGD in each of the 10 wards)
Total number of formative FGDs and IDIs	**30 FGDs and 20 IDIs**

Focus Group Discussions

We conducted 30 FGDs in the 10 selected wards (three FGDs per ward). Focus group research involves guiding a diverse group of participants through a discussion on a particular topic. Focus groups varied in size (generally between six and eight participants) and participants were guided through various discussion topics by a trained facilitator. We conducted 10 FGDs with women's group members, 10 FGDs with women in the community who were former women's group members or had never been members of a women's group (i.e., comparison group participants), and 10 FGDs with spouses of women's group members (Table 2). FGDs with women's group members shed light on members' experiences with women's groups before the start of the program, gendered social norms, and members' livelihoods. FGDs with other women in the community generated broader knowledge about social norms regarding gender and local livelihoods, as well as barriers to entry into and ongoing membership of women's groups. FGDs with the spouses of women's group members shed light on these topics from men's perspectives and explored men's attitudes about women's group participation.

As part of the FGDs, we conducted **participatory livelihood assessments** using a social mapping exercise. Social mapping (Mikkelsen, 2005) is a participatory tool designed to involve community members—the subjects of the research—in the research process as active agents and stakeholders and not just respondents. We used this approach to collect data on local perceptions of gender and poverty, access to services and resources, participants' perceptions of their economic and social situation (as well as that of their community), and available livelihood options.

In-Depth Interviews

We conducted one-on-one IDIs with 20 women participants (Table 2). From the FGD sample, we selected 10 women's group members and 10 women who used to be or have never been women's group members to participate in the IDIs (20 in total). These IDIs explored social norms related to gender, intra-household gender dynamics, and household decision making, and probed for discrete barriers and facilitators related to women's participation in the economic arena. IDIs took place in a private setting where women felt more comfortable discussing these topics.

As part of the IDIs, we conducted **a financial diary exercise**. The evaluation team implemented a streamlined financial diaries approach based on the technique developed by Collins, Morduch, Rutherford, and Ruthven (2010). The financial diaries method uses participants' simplified income statements to help researchers understand household-level income flows and expenditures over time (Table 3).

Table 3. Sample Household Income Statement

Fixed monthly income		Fixed monthly expenditure	
Non-agricultural employment income (men)	₦	Housing (e.g., rent, mortgage)	₦
Non-agricultural employment income (women)	₦	Children's education	₦
Agricultural income (if applicable)	₦	Interest payments on formal loans	₦
State grants	₦	Interest payments on informal loans	₦
Formal loans	₦	Savings	₦
Other (e.g., separate business accounts)	₦	Contributions to women's group	₦
Total:	₦	**Total:**	

Variable weekly income		Variable weekly expenditure	
Informal loans (e.g., borrowed from friends, relatives, money lenders, adashi, esusus, ajo, etc.)	₦	Food/groceries	₦
Gifts	₦	Business-related expenses	₦
Remittances received	₦	Health-related expenses	₦
	₦	Remittances sent	₦
	₦	Airtime/data-related costs	₦
	₦	Gifts	₦
	₦	Transportation	₦
	₦	Entertainment (e.g., dining, alcohol, shows, etc.)	₦
Total:	₦	**Total:**	₦

At baseline, moderators filled out household income statements together with IDI respondents. Respondents were then trained to fill out one income statement per week for the following 3 weeks, for a total duration of 4 weeks. Moderators contacted participants weekly to help them fill out their weekly income statements and provided airtime credits as an incentive and to accommodate follow-up sessions over the telephone. At the conclusion of the 1-month period, moderators collected all weekly income statements in their physical form, if feasible. If not feasible, moderators arranged to receive the income statement inputs over the telephone. This

exercise will be repeated at midline, meaning that a single respondent will be asked to fill out eight weekly balance sheets (four at baseline and four at midline). The research team will analyze and report the data from financial diaries in the midline report.

Key Informant Interviews

We conducted 40 KIIs with stakeholders involved in the design and implementation of the NFWP, as well as local community leaders (Table 4). For our purposes, a key informant is a person who possesses expert knowledge about the NFWP or about a region in which the program is being implemented.

Table 4. Formative Assessment and Process Evaluation KII Sample

Respondents	Formative phase
National level	
World Bank staff	2 KII
Federal Ministry of Women's Affairs officials	1 KIIs
Federal Project Coordinating Units (FPCUs) officials	2 KIIs
State level (five states)	
State Project Coordinating Units (SPCUs) officials	5 KIIs
State-level Ministry of Women's Affairs officials	5 KIIs
Institutional Capacity Building Advisors (ICBAs)	5 KIIs
LGA level (10 LGAs)	
LGA field supervisors	10 KIIs
Ward level (10 wards)	
Ward facilitators	10 KIIs
Total:	40 KIIs

Process Evaluation

To assess implementation fidelity and beneficiary experience thus far, we used a mixed-methods process evaluation. In conjunction with impact evaluations, process evaluations can help ascertain whether a program is ineffective because of its underlying theory or because its delivery was of low quality (Rychetnik et al., 2002). The process evaluation component of this study focused on how the program was implemented, including to what extent program activities were implemented as intended and how beneficiaries experienced the program. The process evaluation also explored external and contextual factors that influenced program implementation. For the qualitative process evaluation, we collected data in the same 10 wards

used for the formative assessment and integrated the process-related questions into the same protocols. We used the following methods for the process evaluation:

1. FGDs with WAG members, former WAG members and women who had never been WAG members, and spouses of women who were in WAGs (FGDs included a participatory livelihood assessment)

2. IDIs with women's group members, former WAG members, and women who had never been WAG members, including a financial diaries exercise

3. KIIs with staff from the World Bank, as well as Federal, State, LGA-level, and ward-level officials

Focus Group Discussions

At baseline, FGDs with WAG members included some process evaluation components that focused on women's experiences of program implementation, perceived challenges in women's group functioning, and perceptions of changes in women's livelihoods or empowerment in relation to their participation in WAGs. FGDs with spouses of WAG members included questions about potential changes in men's attitudes toward women's participation in WAGs and the perceived impacts of the program.

In-Depth Interviews

At baseline, IDIs with WAG members included process evaluation questions about women's experiences of the program, intra-household gender dynamics, household decision making and finances, and perceived changes in women's empowerment. IDIs with former WAG members and women who had never been WAG members investigated their reasons for leaving women's groups, perceived barriers or challenges to retaining WAG members, and consequences for women's livelihoods.

Key Informant Interviews

At baseline, KIIs with NFWP stakeholders included process evaluation questions on how the program was implemented, to what extent it was implemented as intended, and how external or internal factors may have influenced implementation. These KIIs also explored perceived impacts of the program from a stakeholders' perspective. KIIs with community-level project staff investigated perceived program impacts, including perceptions of change in social norms, women's livelihoods, and women's empowerment.

Analysis

In preparing this report, we cleaned all transcripts of KIIs, IDIs, and FGDs and uploaded the raw data into the NVivo qualitative analysis software. We created a preliminary coding structure based on the research questions and the KII, IDI, and FGD protocols. We used this coding

outline to organize and subsequently analyze the information gathered through KIIs, IDIs, and FGDs. To ensure standardization among coders, the research team selected a sample of interviews to double-code. Discrepancies were discussed and resolved before the team coded the rest of the interviews. Afterwards, we analyzed and compared subsections of the data to compile key findings related to the formative and process evaluation questions. We coordinated with the quantitative team consistently to triangulate the qualitative findings with data from the impact evaluation.

Data Collection

Field Staff Training

To ensure data quality, OPM and AIR selected experienced enumerators with the appropriate language skills to carry out the data collection. We then trained field staff to apply the data collection materials and implement quality control procedures for assessments conducted in the field using a train-the-trainers approach. The training for the enumerators took place over two weeks and consisted of one week of GBV-specific training for female enumerators and supervisors and a week for the full team, which included training on best practices for administering surveys, a review of procedures, role play and a pilot of the survey instruments.

Data Collection Oversight

Nigerian team members provided quality control and ensured the technical soundness of the data collection. We generated several protocols to achieve this objective in close collaboration with OPM. For example, data collection supervisors ensured that all surveys and materials were completed correctly before leaving each evaluation site.

For qualitative data, AIR used a systematic and efficient process for organizing and analyzing qualitative data. This process includes audio recording all interviews in the local language, transcribing them in the local language, and then translating them into English. All analysis were performed with deidentified data.

Findings

Implementation Context

Perspectives on the State of Women's Empowerment in Nigeria

According to key interview and focus group respondents, Nigeria has made substantial progress toward gender equality in recent years because of legislation promoting women's rights, increased opportunities to earn an income, and a greater number of women represented in

politics. Most women shared that they generally felt respected in their communities, indicating some level of satisfaction with the status quo. However, restrictive social and gender norms continue to inhibit gender equality, and women expressed a desire to have more agency in their homes and communities. While some program implementers and WAG participants shared optimistic views about the progress of women's empowerment in Nigeria, others shared more pessimistic accounts of women's decision-making power and the barriers and inequalities that women continue to experience.

In the following sections, we discuss qualitative findings regarding the current state of gender equality across the study regions in Nigeria, particularly in relation to barriers to women's economic empowerment, attitudes related to religion and culture, household-level decision-making processes, and GBV.

Barriers to Women's Economic Empowerment

The NFWP places a strong emphasis on women's economic empowerment, yet women across the study regions continue to face barriers to workforce participation. Women face restrictions on their ability to own or inherit land, for example, which influences their financial stability and income-earning ability. Women also have lower levels of education, which resulted in a lack of necessary business skills according to FGD respondents. Most respondents also considered some economic activities such as fishing, harvesting palm fruits, or engaging in large-scale agriculture to be culturally unacceptable for women participants. Furthermore, interview participants indicated that women have less access to financing and other productive assets like machinery or technology, which constrains their productive potential. Gender norms that create these barriers may ultimately create challenges for WAG participants to achieve long-term NFWP goals regarding women's economic empowerment.

Religion

Respondents indicated that religion plays an important role in informing cultural norms about gender and by extension women's empowerment. Although religion is not included in the current NFWP theory of change, the emphasis that program implementers, participants, and community members place on religion indicates that it may be a key factor that moderates the program's influence on community gender views. Several respondents across all regions referenced principles in both Islam and Christianity that seem to support inequitable gender norms. For instance, one spouse of a WAG group member shared, "*our religion tells us that men are the head of women. It even tells us that women have the brain enough for them while the men have the brain of nine people*" (spouse of WAG member, FGD, Niger). However, other respondents claimed that such beliefs are simply religious misinterpretations that the culture perpetuates, or religious justifications for men's perceived superiority. As one female project officer illustrated:

There are some churches that will still not allow their children to speak. Because they would tell you that Jesus is a man, God is a man. I don't know where they got their idea that God is a man. Because the Bible says I was created in the same image. So that means to me […] God too is a woman. Since I was created in his image and likeness (State-level ministry official, KII).

Another project officer understood people's gender biases as vacillating between explaining gender inequality through a lens of culture and through a lens of religion. The staff member explained:

When you push a little, you see that if the [argument] regarding culture and tradition is not working, they'll jump over to religion—what the Bible says, what the Quran says. When you push a little further, they jump into some other things (national-level project staff, KII).

In alignment with the existing scholarship (Beyers, 2017), religion and culture appear to have a mutually reinforcing relationship in the Nigerian context; cultural norms both influence and are influenced by religion. Gender norms therefore manifest as a result of both cultural and religious influences across Nigeria.

Household-Level Decision Making

Respondents shared that households typically abide by traditional patriarchal gender roles within the Nigerian context, both in regard to household responsibilities and decision making. Women are expected to manage household chores, take care of the children, and obey the husband, while men are considered the "head of the household," charged with household leadership and meeting financial needs. Although many people highlighted that some husbands and wives make joint decisions, particularly on issues related to children, most said that husbands have the final say when it comes to household decisions about healthcare, education, and so on. As one woman shared, "if [the husband] is not at home, then your decisions count, but if he is at home, sincerely, he will make all the decisions" (non-WAG member, FGD, Niger).

Respondents also said that women's ability to earn income can lead to more gender equality in the household. They indicated that, in recent years, women have helped to provide financially when the husband cannot, and men have helped with household chores more than they have in the past. Several people, including male spouses, stated that when women have money of their own, they have greater decision-making power; they can make purchases without asking for their husband's permission and have greater input in household decisions. These statements indicate that the NFWP may have the potential to improve women's decision-making power and ability to work outside the home by providing women with access to savings and credit and helping them to increase their incomes.

Gender-Based Violence

GBV is relatively common in both treatment and comparison areas. The quantitative data indicate that approximately a third of women reported their partner has ever physically or verbally abused them, which was slightly more common among NFWP beneficiaries (35%) than comparison group respondents (31%). The qualitative data indicate that instances of physical and verbal abuse of women were somewhat common across states, in both urban and rural communities. Respondents in our qualitative study also provided reports of men neglecting to care for their wives, as well as reports of sexual violence against children. Respondents perceived that poverty, neglect, substance use, and conflicts about money and other domestic issues were at the root of several instances of GBV. Some respondents (both male and female) stated that they believe some women deserve to be beaten by their husbands due to a lack of respect or infidelity. These findings suggest that pathways to reducing GBV in Nigeria may include reducing poverty, changing social norms related to GBV, or increasing women's position of power within the household.

While the qualitative data suggests that some WAG members may have experienced some initial disagreements with their spouses when they first joined WAGs (e.g., shifting expectations around chores and childcare to accommodate group meeting times), most interview respondents generally perceived WAG programming as having helped reduce domestic conflict. These respondents believed that there was less conflict when women had their own sources of income, which the WAGs were helping them to achieve. One person also mentioned that the WAG trained them on conflict resolution, which helped her lead a more peaceful life in the community and at home. We need to exercise some caution in interpreting these results, however, because the evaluation does not allow for establishing causal effects or for assessing whether the NFWP led to increases or decreases in the incidence of GBV.

Profiles of Beneficiaries

The survey data showed few differences between treatment and comparison households related to household size and household infrastructure. The average household had an average of six members, with one household member under the age of 5 and less than one household member (0.42–0.45) of 50 years or older, in both the treatment group and the comparison group. We present these descriptive statistics alongside some details on housing characteristics in Table B1 in Appendix B.

Between treatment and comparison areas, we did not find statistically significant differences related to the age, education level, marital status, or household headship status of women. The average female respondent was 39 years old. Women's ages ranged between 18 and 95, and 84% to 85% of the respondents were married. Only 10% to 11% of the women respondents identified as the household head, and 64% to 65% of the women had received a primary school

education or less (Table B2 in Appendix B). The sample of men (n = 4,763) was also balanced in terms of demographic characteristics: the average male was 43–44 years old, married (83% to 84%), and had received a primary school education or less (50%). Men in treatment households more frequently identified as the household head (65%) relative to their comparison household counterparts (57%) (Table B3 in Appendix B).

NFWP Implementation

This section presents preliminary process evaluation findings based on qualitative data collected during the early stages of the NFWP's implementation. The key factors that respondents perceived as contributing to the success of the NFWP to date were proper training of project staff, engagement and support from key stakeholders, and strong monitoring and evaluation. Major challenges and barriers to implementation included staffing issues, perceptions of delayed and insufficient project funding, transportation and security issues, and administrative bottlenecks.

Key Factors for Implementation Success

Key informants identified several factors that contributed to the successful implementation of the NFWP. Primarily, key informants highlighted the importance of adequate training of project staff—particularly facilitators who interact directly with WAG members—as a key factor for the NFWP's early success. One official from the Ministry of Women's Affairs felt that if the program could not get the training component right, the negative effects would "cascade onto the women"; if the training was conducted properly, however, "it will be easy to affect the women with what [the facilitators] have learned" (State-level ministry official, KII). Another key factor for success was soliciting buy-in and support from relevant stakeholders at all levels of administration (national, state, LGA, and ward), as well as from traditional rulers. As described earlier, buy-in from traditional rulers is critical in terms of building trust for the NFWP. At the same time, key informants warned that while it was important to maintain the support of government officials, bureaucratic processes may "slow the pace of work" and end up "impeding the fast execution and overall success of the project" (State-level project staff, KII, Taraba).

Another key factor that respondents identified as being critical to successful implementation was robust monitoring and evaluation (M&E). Respondents highlighted that M&E was important to not only ensure that program implementers identified and addressed gaps in services in real time, but also that they were able to document and convey successes to other implementers so that these successes could be replicated.

Implementation Challenges and Barriers

Respondents shared how security issues such as banditry, kidnapping, and community-level conflicts were challenges to implementation in certain communities across project states. Some areas were known by locals as "no-go areas" because of violence and the high risk of being targeted by kidnappers, bandits and armed robbers. One ICBA staff explained:

> *You discover that there are people who hide in the bush along the road to attack passers-by and take away their money and belongings, and some kidnap and place [a] ransom on you, a whole lot of that. So, those activities [in] communities are actually tampering with everything (ICBA, KII).*

While we did not observe any regional patterns in what were considered "insecure" areas, respondents specifically identified Takalafiya, Bali, and Takum in Taraba and Wushishi in Niger as highly insecure and thus as challenging from a program implementation perspective. Concerns over security made it difficult to enlist ward facilitators because program staff perceived security provisions as inadequate. One ward facilitator shared that ward facilitators were not given security assurances and warned that they assumed risk themselves if they chose to stay in a community past a certain timeframe:

> *When you go to work, they just said 6 [am] to 6 [pm]. If you stay in the community until after 6pm, then you are on your own because you have been told it is 6 to 6pm…So we are on our own when it comes to insecurities (Ward facilitator, KII).*

It was not clear from the data what prevented program staff in these insecure areas from providing adequate security to ward facilitators (e.g., funding, lack of awareness) nor was it clear how pervasive the lack of security provision was across the program areas.[7]

Another challenge to the NFWP's implementation was related to delays in State approvals of project expenses. One advisor in Niger State shared how financing delays caused by delays in State approvals had adversely affected the project's ability to pay staff, address security issues, and procure goods and services at the current market rate. Another project staff member detailed how such delays affect project activities:

> *You develop a plan. You have a work plan for an activity that would be conducted within a quarter or within a month. Before you get clearance for that work plan or activity plan, it would go beyond the timeline that you've spelled out for that activity (Federal project staff member, KII).*

---

[7] Future data collection can focus on security provisions including what are the barriers to providing adequate security, what kind of security provisions are necessary, and what geographic areas require higher levels of security support.

Key informants pointed to administrative bottlenecks as the cause of delays to project fund disbursement. Specifically, navigating the State bureaucracy to receive the necessary approvals—what one World Bank staff member referred to as "the authorizing environment of the state"—led to implementation delays that delayed fund disbursement.

In addition to these delays, many staff members expressed dissatisfaction with the remuneration amount they receive. As one staff member shared:

> *Let them treat us fine, we are trying so hard. Let them start giving us this performance allowance, let them give us reasonable transport allowance. Let them give us something that after our tour of the communities, we would have something left to buy ourselves a drink. Not dipping our hands into our pockets to complete the processes of touring round the local governments (LGA-level staff, KII).*

The perceived insufficiency of funds appeared to be particularly acute in rural areas of the country, compared with the sampled urban wards. One ward facilitator complained that the stipend amounts were insufficient to cover the costs of transportation to reach the remote parts of the project's coverage area. These low stipend amounts compounded mobility-related challenges such as high transportation costs and poor terrain, which were present in all NFWP-implementing states. One field supervisor in Taraba provided details on the challenges of reaching some of these remote communities:

> *Sometimes, before you get to the community, you will cross three rivers and two mountains … I could not climb the mountain; they ended up carrying me at their back and I have to pay the person that carried me at his back (LGA-level staff, KII).*

Finally, interview respondents claimed that it was a challenge to identify and recruit the right staff to implement the NFWP. One World Bank staff member recounted that, at the Federal level, the project had to replace an entire unit with new staff because the previous staff were not delivering services with acceptable quality. One State-level staff member shared, "in some communities where we work, you find old women, people who cannot read" who are comprise the ward facilitator staff, making it difficult to implement the project with fidelity (State-level project staff, KII, Ogun).

Project Stakeholders' Suggestions for Project Improvements

Key informants made several suggestions on how to improve the NFWP moving forward, based on the challenges and barriers they identified. These suggestions reflect key informants' attitudes and perceptions about what they feel needs to change to improve program implementation moving forward. However, the suggestions may often require additional resources, which may not be feasible when aiming to maximize the cost-effectiveness of the NFWP.

Key informants made the following suggestions:

- Continue investing in sensitization efforts to help the NFWP overcome trust issues regarding the project. Sensitization efforts can target religious and traditional leaders to enlist their help in building trust at the local levels for the project.

- Review and revise the recruitment criteria for ward facilitators to help resolve staffing issues. For example, the education requirements for ward facilitators should be increased to improve the quality of staff. Changes to recruitment approaches should be combined with more training for staff to support capacity building, as well as M&E efforts to continually assess staff performance.

- Increase or reallocate financial support for the project to fund staff pay increases, transportation costs, and improved security.

- Streamline bureaucratic and approval processes to improve implementation. Key informants generally felt that while oversight is important, these processes were redundant and inefficient. Specifically, interview respondents advocated giving more autonomy to local project implementing units to reduce administrative bottlenecks.

We will triangulate these suggestions from key informants with the potential cost implications for the NFWP in the conclusion.

Impact Evaluation

This section presents impact evaluation findings, organized by the potential channels through which the NFWP can achieve its objectives. These channels include (a) group formation and savings training; (b) savings and loans, combined with financial education; (c) business skills training and livelihoods grants; and (d) mature WAGs and livelihood collectives. We hypothesized short-term impacts for some of these channels (group formation and savings training, and savings and loans) and longer term impacts for other channels (business skills training and livelihoods grants, and mature WAGs and livelihood collectives). We made this distinction because it enabled us to clarify where impacts were expected at this stage of the evaluation, and where we should expect no differences between treatment and comparison groups. It is important to make this distinction because evidence from similar programs in other contexts, particularly the National Rural Livelihoods Mission (NRLM) in India, indicates that self-help groups and savings groups such as the NFWP may have different short-term effects (primarily on financial inclusion, with some evidence for positive effects on asset ownership and women's empowerment) and long-term effects (primarily on livelihoods, consumption, and income). For example, the RCT of the JEEVIKA self-help-group program in Bihar showed positive impacts on access to formal credit in the short term (Hoffmann et al., 2021) but only showed impacts on consumption in the long term (Siwach et al., 2022; Kochar et al., 2020). In addition,

there was mixed evidence for positive impacts on women's decision-making power and other women's empowerment indicators (Hoffmann et al., 2021; Brody et al., 2017; Siwach et al., 2022).

Outcomes From Group Formation and Savings Training

Rate of Membership in WAGs

In accordance with the theory of change, there was an increase in WAG membership under the NFWP, but we found a corresponding decrease in membership in other savings and women's groups. Among households in NFWP areas, WAG membership increased by 65 percentage points while other women's and savings group membership simultaneously declined by a smaller margin (31 and 32 percentage points, respectively) (Table 5). We define other women's groups as non-WAGs that a majority of women members, while we define other savings groups as savings groups with a membership of less than 50 percent women. Overall, 65% of women in treatment LGAs reported being WAG members. Membership of any group increased by 16%. The impact of NFWP on WAG group membership was larger for women over 30 years old, and the impact on other women's group membership was larger among women in northern NFWP areas (Table D1 in Appendix D). In the absence of an opportunity to join a WAG, community and religious groups were seemingly among the most common group type: Women in comparison areas who were engaged in other savings and women's groups (67%) were more likely to report membership in community groups (45%), religious groups (27%), friend groups (15%), and livelihood groups (14%).

Table 5: Single-Difference Impact Estimates - NFWP and Savings

	Member of any group	WAG member	Other women's group member	Other savings group member
NFWP	0.158	0.653	-0.313	-0.320
	(0.024)***	(0.035)***	(0.023)***	(0.024)***
Comparison Mean	0.65	0.00	0.62	0.65
N	4789	4789	4789	4789

*p<0.1 ** p<0.05; *** p<0.01*
Estimations control for state fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Barriers to and Perceived Costs of Participation in WAGs

Despite large membership rates in WAGs, qualitative data indicated several key barriers to entry. Most notably, women conveyed that they initially mistrusted the NFWP based on their prior experience with livelihood interventions that they perceived to be similar to WAGs.

Women shared that in the past, they had experienced similar groups arriving, taking money, and providing little to them in the long term. Because of these experiences, the women expressed doubt that the NFWP would fulfill its promises and suspected that the project would exploit women in the community. As one WAG member detailed, members of her community mistrusted the arrival of the NFWP:

> *The community people laughed at us for believing in the project. They said we are not tired of being fooled … they said they are going to pack up our money and not come back.*

> *(WAG member, FGD, Taraba)*

Driven by their mistrust, several women who were not part of WAGs at the time of our interview decided to first observe WAGs before joining. They indicated that they eventually came to see the value of the NFWP after observing its activities in their community, but that they had not seen an opportunity to join after the initial group formation period. Indeed, after initial enrollment, the NFWP halted recruitment of new members in order to assess the capacity of ward facilitators to oversee existing WAGs. This process took upwards of 6 months and did not always result in the formation of additional WAGs.

Among those who did not join WAGs or dropped out of a WAG, economic factors were key barriers to their participation. Notably, qualitative data indicate that WAG members experienced and conceptualized the mandated savings payments differently depending on income level. While some WAG members indicated no issue making payments, others eventually dropped out of the group because the payments strained their household budgets. Indeed, savings payments were conceptualized as a kind of expenditure for both WAG members and non-members. As one non-WAG respondent summarized, "You can't do anything without money. Without money you can't start a business and without [a] business you can't join the association" (non-WAG member, FGD, Ngaski). In contrast, most WAG members shared that they had a business upon joining, leading them to feel comfortable making savings payments to the WAG.

In some wards, miscommunication about the requirements for joining the WAG presented additional "startup" costs for joining the group: an ID and a SIM card. Most WAG members reported that they owned both items before joining, while non-participants in several wards mentioned these items, especially the ID card, as a cost-prohibitive barrier to their enrollment in the group. In fact, IDs and SIMs are not a formal prerequisite for joining WAGs but can facilitate registration for a bank account and communication about group activities, respectively. Nonetheless, project staff and non-participants in two rural wards of Kebbi and Taraba states cited the SIM requirement as a barrier. In five wards across Kebbi, Taraba, and

Ogun states, both WAG participants and women who did not participate cited the ID requirement. Such data indicate sporadic misunderstanding of the actual program requirements among women and, to some extent, project staff (i.e., ward facilitators) in beneficiary communities.

While the NFWP primarily targets the "missing middle" (i.e., women who are economically active but live close to the poverty line), the perceived economic barriers highlight how the project's current design may limit the ability of some economically inactive, poorer women to join WAGs. If women lacked stable income or experienced a household shock (e.g., sickness, building repairs) at the time of WAG enrollment, the barrier to entry was often considered too high. Respondent testimonies underscored how the financial costs of joining WAGs inhibited the participation of some lower income women, though such barriers were unlikely to affect the project's ability to reach most women considered to be in the missing middle.

Additionally, women perceived that joining WAGs came with some time-related and social costs, which were also barriers they had to overcome to continue their participation. Respondents described how the time burden of WAG activities was greater early in the group formation process, before women had adjusted their chore schedules and before the effect of the tardiness fees set in. At that time, WAG meetings could last up to 3 hours, whereas most participants said their meetings now last only 30 minutes. Group members explained that they had learned to adjust their household chore schedule on days when they had WAG meetings, usually by waking up earlier and completing chores before the meeting, or by reallocating chores to a child or spouse. Survey data confirmed that WAG members reported spending more time on group activities and trainings than non-WAG members (we present more details in the costing analyses).

Communities reported a few cases where WAGs instigated spousal conflict, but these, too, occurred early in the formation of the groups as spousal relationships adjusted to the introduction of WAGs. Such conflicts revolved around the use of or communication about the WAG loan. For instance, one respondent reported that a feud ensued after a husband demanded the loan be used for his purposes and not his wife's. In another case, a woman did not notify her husband that she had received the loan and he grew angry with her, believing she was keeping secrets. These conflicts signal how women must navigate gender norms whilst participating in WAGs, and they highlight how the WAG loan can aggravate gendered social dynamics.

Participation in WAGs: Key Facilitators and Motivations
Three key factors helped women overcome the barriers to WAG membership. First, many WAG participants indicated that the sensitization efforts of the NFWP played a crucial role in their

decision to join the project. The participation of local leaders such as the village chief helped to strengthen sensitizations. In addition, sensitizations demonstrated that the project was endorsed by the Federal government, which helped build trust for the NFWP. The sensitizations helped build the credibility of the project, and helped to portray the WAG model as distinct from other savings groups and livelihood interventions that had previously come to communities. In particular, participants pointed to WAGs' low-interest loans and formal regulations around money collection, timeliness, and defaulting as making WAGs distinct from other savings groups.

Second, the community-based savings model and the consistency of the project were key factors that contributed to the perceived transparency and reliability of the project, facilitating women's trust and participation in WAGs. Women who participated in WAGs shared that the NFWP had a consistent presence in their communities, leading them to view it as accountable and reliable. This consistency and visibility reassured participants and their families that they could rely on the project to provide support in the long term and helped address their concerns that the project would disappear without warning. Further, the transparent savings model of WAGs—where several women have keys to a savings box, which stays in the community—helped participants see that their money was safe and would be used for its intended purpose, addressing their concern that their money would be taken away without benefitting them.

Third, the support of women's social networks encouraged their participation in WAGs, particularly support from spouses and elder, well-respected women in the community. Spousal support was critical because a husband's permission is needed to join a WAG. Most WAG members reported that their spouses consented to their participation. In cases where spouses refused to let their wives join WAGs, the women usually convinced their husbands to accept their participation by having other community members intercede on their behalf. The participation of respected women in WAGs also signaled to other women that the group was trustworthy. As one WAG member in Ijebu North explained, "when WAGs came, I asked who was in charge and they told me Mama Adekoya, and my mind was at rest" (WAG member, FGD, Ogun). Beyond joining WAGs, elder women also encouraged participation by serving as important sources of information, explaining the project to women throughout the community.

Finally, the perceived and expected benefits of WAGs motivated women to join. Women perceived that one of the main economic benefits of joining WAGs was their increased savings, access to credit, and support amid household shocks. Respondents noted that these anticipated benefits would make it easier for them to finance household expenditures, invest in their business, and respond to unexpected shocks. In terms of social benefits, respondents were motivated to join WAGs to have a social outlet outside the home, and to build cohesion and a sense of unity among women, enabling them to tackle problems in their community.

Group Composition

After the first year of NFWP implementation, women's groups and savings groups were fairly socially homogenous across treatment and comparison areas, with a majority of group leaders indicating that most or all of their members had the same gender, religion, ethnicity, and language (Table 6). This limited diversity within groups casts doubt on the extent to which ethnic and religious minorities and majorities comingle in these groups.

Despite similarities in group composition by treatment status, group size diverged: On average, groups in NFWP areas were smaller than those in comparison areas, with a higher proportion of groups in comparison areas reporting more than 25 members (47%).

Table 6. Single-Difference Impact Estimates: Group Composition

	Formed before NFWP began	Group size	Has more than 25 members	Age of members	Average member age ≥30 years	Proportion of female members	Have the same religion	Have the same ethnicity	Speak the same language
NFWP	-0.473	-8.321	-0.276	1.876	0.038	0.062	0.072	-0.017	-0.069
	(0.054)***	(3.355)**	(0.053)***	(1.913)	(0.091)	(0.014)***	(0.084)	(0.076)	(0.064)
Comparison Mean	0.78	38.95	0.47	34.47	0.71	0.94	0.77	0.76	0.77
N	1267	1267	1267	1267	1267	1206	1264	1266	1266

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for WG age and registration status. We winsorized group size outliers with the 99th percentile.

In terms of group-member demographics, the data demonstrated systematic differences across treatment and comparison areas, in the age, education levels, and likelihood of a disability of group members. In Table 7, the comparison mean specifically portrays the proportion of sampled women's groups in which a majority of members possessed the same trait (e.g., marital status, highest level of education). As the table illustrates, group members in NFWP areas were less likely to finish secondary school (8%) and to live with a disability (less than 1%) and were more likely to be new group members (i.e., joined in the past 6 months; 21%) and to have income from wage labor or their own business. While this aligns with the focus of NFWP on economically active women, it also suggests that groups in NFWP areas have yet to increase their inclusion of women living with a disability. Despite lacking within-group diversity, the differences in demographic trends between treatment and comparison groups indicate that the NFWP influenced *across*-group diversity, which could potentially strengthen social cohesion within communities and improve women's ability to pool risk and resources in the long term.

Table 7. Single-Difference Impact Estimates: Group Member Characteristics

	Joined in the past 6 months	Joined over a year ago	Are married	Are living with a disability	Finished primary school	Finished secondary school	Have their own business	Do not have income	Receive income from group
NFWP	0.157	-0.233	-0.001	-0.009	-0.005	-0.109	0.059	-0.096	-0.056
	(0.047)***	(0.082)***	(0.019)	(0.004)*	(0.075)	(0.034)***	(0.042)	(0.056)*	(0.101)
Comparison Mean	0.05	0.75	0.96	0.01	0.42	0.19	0.80	0.13	0.27
N	1252	1260	1266	1259	1240	1225	1264	1216	1263

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for WG age and registration status.

We did not find major statistically significant differences in group social cohesion between the treatment and the comparison group. Across both treatment and comparison areas, group leaders reported high levels of social cohesion, affirming their sense of belonging and reliance on group members for support (Table 8). Relative to comparison areas, group leaders in NFWP areas expressed a desire to be with a different women's group less often, indicating a strong sense of belonging in their current group. However, across both treatment and comparison groups, nearly half of the group leaders reported that group members were interested in pursuing their own goals and were worried only about themselves. To the extent that group members possess different interests, pursuing their own self-interest may undermine their group's capacity to pursue joint goals, especially if within-group diversity increases throughout implementation of the NFWP as the theory of change projects.

Table 8. Single-Difference Impact Estimates: Women's Group Social Cohesion

	I feel that I belong to this WG	I would rather be with a different WG	I would prefer this WG over others	Members of this WG are all striving for the same goals	Everyone in our WG wants to pursue their own goals	WG members worry about only themselves	I can count on fellow WG members if I need to borrow money	I can count on fellow WG members to accompany me to the doctor or hospital	I can count on fellow WG members to talk to about my problems	I can count on fellow WG members if I need advice
NFWP	0.006	-0.069	-0.010	0.005	-0.110	-0.020	0.019	-0.021	0.002	-0.015
	(0.005)	(0.041)*	(0.023)	(0.014)	(0.087)	(0.079)	(0.022)	(0.018)	(0.020)	(0.010)
Comparison Mean	1.00	0.15	0.97	0.97	0.52	0.45	0.91	0.97	0.97	1.00
N	1267	1265	1267	1265	1262	1263	1264	1264	1266	1266

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for group age and registration status.

While group leaders reported similar levels of social cohesion across treatment and comparison areas, groups in treatment areas met more frequently, which may increase trust as well as a sense of shared responsibility and community among members in the medium- to long-term

Impact Evaluation of the Nigeria for Women Project: Baseline Report

(Table 9). More specifically, groups in treatment areas were more likely to meet on a weekly basis (91%) relative to groups in comparison areas (62%).

Table 9. Single-Difference Impact Estimates – Group Meeting Frequency

	Meet weekly	Meet bi-weekly	Meet monthly
NFWP	0.287	-0.074	-0.194
	(0.046)***	(0.019)***	(0.031)***
Comparison Mean	0.62	0.09	0.22
N	1265	1265	1265

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for WG age and registration status

Financial Inclusion and Assets

Savings

After the first year of implementation, the NFWP led to a significant increase in the likelihood of saving of 12 percentage points relative to the comparison group (Table 10). This increased savings behavior, however, did not result in statistically significant impacts on cumulative individual- or household-level savings in the short term. We found no statistically significant differences in cumulative individual- or household-level savings between the treatment and comparison groups, possibly because households invested some of their savings in small businesses or assets, as discussed below and shown in the impact estimates on asset ownership and livelihoods.

Table 10. DID Impact Estimates: Individual Savings

	Woman saves money	Total savings (NGN)
ATT	0.122	1,809.381
	(0.022)***	(3,073.513)
Comparison Mean	0.71	21755.48
N	4306	4003

LGA-clustered standard errors in parentheses. * *p*<0.1; ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

In addition to increasing the propensity to save, the NFWP is associated with changes in where women keep their savings, shifting savings away from other savings groups to WAGs. More specifically, as of Year 1, the NFWP led to an increase of NGN5,576 in savings in WAGs and a reduction of NGN1,280 in savings in other savings groups (Table 11). By contrast, respondents in comparison areas were more likely to have savings in other women's groups (48%) as well as other savings groups (35%). While we did not detect heterogeneous effects for individual

savings, we found evidence for some heterogeneities in effects for group-based savings, particularly by age cohort and—to a lesser extent—by religious affiliation; we observe larger positive effects on WAG savings for older NFWP beneficiaries (i.e., over 30 years old). In addition, we observe a larger effect on the likelihood of having savings in other savings groups for Muslim women.

Table 11. Single-Difference Impact Estimates: Group-Based Savings

	Has savings in WAG	Cumulative savings in WAG	Has savings in other WG	Cumulative savings in other WG	Has savings in other SG	Cumulative savings in other SG
NFWP	0.544	5,576.326	-0.239	-1,267.667	-0.200	-12.067
	(0.032)***	(698.005)***	(0.023)***	(263.008)***	(0.018)***	(45.390)
Comparison Mean	0.00	0.00	0.48	2342.12	0.35	80.30
N	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

In FGDs, WAG members indicated that the savings component emerged as a highlight of their experience with the NFWP. For instance, both WAG members and their spouses reported that women increasingly had the necessary savings to cover some household and business expenses on their own, rather than having to ask their husbands for the money. In explaining this trend, WAG members reported that the personal finance trainings taught them concrete skills related to tracking their income, saving money, and reinvesting in their businesses—practices that many women stated they did not previously do. This likely explains the lack of findings on accumulated savings despite the positive impact on the likelihood of saving – while women save at a higher rate, they also use those savings to purchase assets or to invest in their businesses.

Credit

The NFWP had statistically significant impacts on access to credit, increasing the likelihood of borrowing by 9 percentage points (Table 12). This increased borrowing had not yet translated into statistically significant differences in the overall value of outstanding loans. After the first year of implementation, the average value of outstanding loans was slightly higher in NFWP areas (6,564 NGN) than in comparison areas (6,287 NGN) (Table 11). While this difference was not statistically significant, the NFWP was associated with higher average values for women's largest loans (Table 12). The NFWP also had heterogeneous impacts on the value of outstanding loans, for which effects were larger for younger (i.e., 30 years old or less) NFWP beneficiaries (Table D8a in Appendix D).

Impact Evaluation of the Nigeria for Women Project: Baseline Report

Table 12. DID Impact Estimates: Access to Credit

	Has outstanding loans	Value of outstanding loans (NGN) (includes 0)
ATT	0.091	277.201
	(0.017)***	(1,488.257)
Comparison Mean	0.09	6287.11
N	4785	4789

LGA-clustered standard errors in parentheses. * *p*<0.1; ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, a recall value for outstanding loans, and
an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

In addition to increasing credit access, the NFWP was associated with shifts in women's credit sources, with women in NFWP areas more frequently accessing their largest loans from women's groups or other savings groups, relative to comparison areas (Table 13). We observe heterogeneity in the effect sizes by region and religion: The effects on accessing credit through the aforementioned sources was higher in northern NFWP areas and among Muslim beneficiaries (Table D9c in Appendix D). The qualitative data indicate that women perceived WAG loans, in particular, to be more accessible. WAG members explained that they viewed WAG loans as distinct from traditional micro-credit because of their low interest rates and less-severe penalties (e.g., fines) for defaulting.

By facilitating savings and loans, WAGs also helped participants and their families weather unexpected household shocks (e.g., sickness, death) and cover the associated expenses (e.g., hospitalization, funeral). Such uses of WAG loans were common among participants, who also reported that the WAG would sometimes collect funds to donate to a member who experienced a shock, most commonly a death in the family. Just under half of the women in NFWP areas (47%) reported using their largest loan for personal expenses (health, food, and education expenses), while 53% of the respondents using their largest loan for investments (e.g., investments in existing non-agricultural businesses and investments to start an agricultural business). Table D9d in Annex D provides more details on the use of the loans.

Impact Evaluation of the Nigeria for Women Project: Baseline Report

Table 13. Single-Difference Impact Estimates: Loans

	Number of current outstanding loans	Total value of current largest loan (LL)	Value of LL that still needs to be paid (NGN)	WG/SG provided largest loan
NFWP	0.157	2,442.946	1,851.295	0.079
	(0.028)***	(1,921.897)	(1,300.591)	(0.012)***
Comparison Mean	0.11	7356.49	4585.95	0.02
N	4789	4789	4779	4788

LGA-clustered standard errors in parentheses. * *p*<0.1; ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, a recall value for outstanding loans,
and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Asset Ownership

The NFWP led to increased consumption of certain productive and consumption assets but we did not find consistently significant effects on asset ownership. We found positive effects on the household ownership of non-mechanized farm equipment and non-farm business equipment as well as ownership rates for homes, small durable goods, and mobile phones. As mentioned, it is possible that we did not find statistically significant impacts on cumulative savings because women and other household members invested some of their savings in assets. Alternatively, it is possible that some women increased their mobile phone ownership because of the misunderstanding in some communities about phones being a requirement for joining WAGs.

Table 14. DID Impact Estimates: Household Productive Asset Ownership

	HH owns large livestock	HH owns small livestock	HH owns poultry	HH owns non-mechanized farm equipment	HH owns mechanized farm equipment	HH owns non-farm business equipment	HH owns uncultivated land
ATT	-0.001	0.015	-0.005	0.027	0.004	0.012	0.008
	(0.012)	(0.013)	(0.015)	(0.008)***	(0.005)	(0.006)**	(0.009)
Comparison Mean	0.17	0.42	0.43	0.78	0.09	0.18	0.32
N	4763	4763	4763	4763	4763	4763	4763

LGA-clustered standard errors in parentheses. * *p*<0.1; ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, ownership of assets at baseline,
and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Table 15. DID Impact Estimates: Household Consumption Asset Ownership

	HH owns a house	HH owns large durable goods	HH owns small durable goods	HH owns a mobile phone	HH owns a smart-phone	HH owns a means of transportation
NFWP	0.022	0.004	0.026	0.034	0.005	0.010
	(0.007)***	(0.009)	(0.010)**	(0.015)**	(0.011)	(0.010)
Comparison Mean	0.62	0.38	0.70	0.85	0.24	0.42
N	4763	4763	4763	4763	4763	4763

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, ownership of assets at baseline,
and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Increased asset ownership in NFWP areas was seemingly driven by women's sole ownership of these assets. More specifically, NFWP beneficiaries were more likely to solely own mechanized farm equipment (2%), uncultivated land (10%), and a mobile phone (60%); conversely, they were slightly less likely to own large livestock (2%) and large durable goods (15%) relative to their comparison group counterparts (Tables 16 and 17). Nevertheless, despite statistically significant, the magnitude of these impacts was relatively small, ranging between 1 and 5 percentage points. These marginal changes suggest promising yet preliminary progress in this domain.

At the household level, we did not find much variation in the impact on asset ownership across household characteristics. The only exception is that the impact of NFWP on ownership of uncultivated land was less pronounced for households with Christian households. However, we observe stronger heterogeneity in impacts on sole asset ownership, which varied by geographic location and religion of NFWP beneficiaries with larger effects on Muslim women and in Northern Nigeria (Tables D12-D13 in Appendix D).

Table 16. Single-Difference Impact Estimates: Sole Ownership of Productive Assets

	Solely owns large livestock	Solely owns small livestock	Solely owns poultry	Solely owns non-mechanized farm equipment	Solely owns mechanized farm equipment	Solely owns non-farm business equipment	Solely owns land	Solely owns uncultivated land
NFWP	-0.011	-0.009	-0.015	0.002	0.008	-0.005	0.018	0.028
	(0.005)**	(0.026)	(0.019)	(0.027)	(0.003)***	(0.012)	(0.012)	(0.012)**
Comparison Mean	0.03	0.27	0.30	0.30	0.01	0.09	0.14	0.07
N	4789	4789	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 17. Single-Difference Impact Estimates: Sole Ownership of Consumption Assets

	Solely owns a house	Solely owns large durable goods	Solely owns small durable goods	Solely owns a mobile phone	Solely owns a smartphone	Solely owns a means of transport	Total assets solely owned
NFWP	-0.012	-0.031	-0.033	0.052	-0.001	0.003	-0.006
	(0.009)	(0.014)**	(0.021)	(0.025)**	(0.009)	(0.014)	(0.085)
Comparison Mean	0.07	0.18	0.44	0.55	0.07	0.04	2.56
N	4789	4789	4789	4789	4789	4789	4789

Business Skills Training and Livelihoods Grants

According to the theory of change, improved business skills through group-based trainings and the distribution of livelihoods grants to WAG members could impact medium-term outcomes such as labor force participation, income, and income diversification through the expansion of individual and household businesses into agricultural and non-agricultural activities. In the longer term, the theory of change projects changes in consumption and women's empowerment due to the NFWP. At this time, we do not expect meaningful changes in these outcomes because WAG trainings have only started recently and the vast majority of livelihoods grants are yet to be distributed.

In year 1, women in NFWP areas were more likely to receive group-based trainings than comparison households, by 19 percentage points (Table 18). This is expected because NFWP groups were mostly formed within the last year, and the frequency of trainings is usually higher in the beginning stages of group programming. Yet, the proportion of women in NFWP areas who received such trainings was relatively small (28%), suggesting that trainings were still in nascent stages in February and March of 2022. Across treatment and comparison households, the most common training topics were savings, business plans, and social norms (Table C1-C3 in Appendix C). Trainings on savings, business plans, financial literacy, and credit were more common in NFWP areas than in comparison areas. This aligns with the initial focus of the program on financial inclusion.

Table 18. Single-Difference Impact Estimates: Group-Based Training

	Participated in training via WG or SG
NFWP	0.195
	(0.022)***
Comparison Mean	0.08
N	3364

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

In line with the NFWP implementation timeline, most respondents had yet to receive livelihoods grants. As of Year 1, only a small proportion of respondents had received livelihoods grants through their women's group (8%) or savings group (6%). We expect changes in grant disbursement to take place in Year 2, in accordance with implementation plans.

In interviews, respondents shared that the disbursement of livelihoods grants was a much-anticipated element of the NFWP. Project staff, as well as WAG members and their families, all emphasized that implementers had to get grant disbursement right for the NFWP to be successful. Specifically, they shared how grievances among women could arise if the grants were not disbursed equitably. While livelihoods grants had only been disbursed in a few program areas at the time of data collection, this perception highlights the importance of implementing equitable and transparent grant criteria in order to maintain group cohesion and trust.

Women's and Households' Labor Force Participation, Income, and Agricultural Activities

Comparing recall data from pre-program period across treatment and comparison households suggests that there were no statistically significant differences in baseline labor force participation and income diversification—conceptualized as the number of different income-generating activities (IGAs)[8]—between treatment and comparison individuals and households. About 90% of women in our sample participated in the labor market by engaging in IGAs, with approximately 60% being self-employed (Appendix B, Table B6). We also found that the average woman was engaged in 1.2 to 1.3 IGAs, suggesting that there was little income diversification at the individual level. At the household level, Table B7 in Appendix B shows that over 90% of households had at least one member (excluding the female respondent) who

---

[8] We considered the following activities: worked on farm (crop production) outside the household, worked on own farm (crop production) and selling products, livestock tending outside the household, raising own livestock and selling products, casual labor, employed in a company, work in aquaculture/fishing and selling products, and running own business/self-employed.

participated in the labor market before the start of the NFWP; and at baseline, households engaged in just under two different IGAs.

Similarly, and as expected, we did not find any significant differences in agricultural outcomes between treatment and comparison households. Most households in both the treatment and comparison groups participated in agricultural activities. Table B8 in Appendix B shows that between 82% and 85% of households cultivated land, and 36% of household income came from agricultural production. The average household had cultivated three plots during the last 12 months, and the area of its two largest plots corresponded to approximately one hectare.

Although we did not expect changes in labor force participation and income at this point, we still explored DID estimates of the impact of the NFWP on this set of medium- to long-term outcomes after 1 year of implementation. In the short term, the NFWP had not affected households' or women's labor force participation, income diversification, or women's monthly income, on average (Tables 19 and 20). However, we did find a significant (albeit small) increase in women's self-employment. Table 18 shows that women in NFWP areas were 2.6 percentage points (or about 5%) more likely to be self-employed than women in comparison areas. This finding suggests that some WAG members may have started applying their entrepreneurial skills and making productive investments to start or expand their businesses. Indeed, as shown in Table 20, the NFWP is associated with a slight increase in the number of women's businesses. The probability of owning a business increased by 3.1 percentage points (or about 5% relative to the comparison group) for females in NFWP areas. We will revisit these results at midline, when we expect to find that more WAG members have received business skills training and livelihoods grants.

Analyses of heterogeneous effects by women's age, religion, and geographic location revealed that although the labor force participation of Christian women was not affected by the NFWP, Muslim women did experience a significant increase in their workforce participation. Muslim women in NFWP areas were more likely to engage in IGA by about 4.9 percentage points (or about 6%) relative to the mean of the comparison group (Table D14c in Appendix D). This result is robust to exploring impacts by geographic location. Table D14b (in Appendix D) shows that women living in the northern states of Kebbi and Niger, predominantly Muslim, saw a significant increase in their short-term labor force participation due to the NFWP. This effect was not experienced by those living in the southern states (i.e., Abia, Ogun, and Taraba). The differential impact of the NFWP on the labor force participation of Muslim and Christian women may be due to a ceiling effect. Recall data on labor force participation from December 2020 indicate that over 90% of Christian women and a little over 80% of Muslim women participated in the workforce. These numbers suggest that it may be more challenging for the NFWP to increase the high labor force participation among Christian women. By contrast, there is more room for growth in labor force participation among Islamic women.

Table 19. DID Impact Estimates: Women's Labor Force Participation (LFP) and Income Diversification

	Labor force participation	Self-employment	Number of IGAs
ATT	0.013	0.026	-0.007
	(0.014)	(0.014)*	(0.038)
Comparison Mean	0.77	0.55	0.98
N	4789	4789	4789

LGA-clustered standard errors in parentheses. * *p*<0.1; ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table 20. DID Impact Estimates: Household LFP and Income Diversification

	Labor force participation	Number of IGAs
ATT	-0.017	-0.004
	(0.011)	(0.033)
Comparison Mean	0.80	1.14
N	4763	4763

LGA-clustered standard errors in parentheses. * *p*<0.1; ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

The qualitative data also indicate that WAG loans enabled women to scale up their existing businesses or start new businesses. For example, several women reported that they used the WAG loans to buy fertilizer, allowing them to increase the number of plots they cultivated and reap a more productive harvest. Another WAG member explained that she was able to expand her soap business to sell additional hygiene products. Several women reported that, with the help of the WAGs, they no longer had to purchase items for their business on credit and were able to save the money they would have otherwise spent on interest. In a few cases, WAG members had helped non-participating household members start businesses. One WAG member, for instance, shared how she used her loan to help her daughter fund her own business. However, the quantitative findings indicate that starting new businesses had not yet translated into additional income for women, as average business sales remained statistically similar between treatment and comparison groups (Table 21).

Table 21. Single-Difference Impact Estimates: Women's Businesses

	Owns a business	Business sales – last 4 weeks	Number of people employed
NFWP	0.031	-1,710.888	0.090
	(0.017)*	(1,638.925)	(0.069)
Comparison Mean	0.62	21109.56	0.77
N	4789	4563	4789

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

The NFWP's medium- to long-term impact on women's labor force participation, income, and self-employment may depend on the extent to which women's social networks and business linkages expand as a result of participating in WAGs. The qualitative data indicate that even in the short term, participation in WAGs expanded members' connections with key economic actors including value-chain actors, financial institutions, and government entities. At the State level, WAGs that were operating in Abia and Ogun states had established networks with government agencies and agricultural development projects (ADPs). In Ogun, for example, one state-level program officer shared how during a scheduled monitoring trip, staff from the State Ministry of Agriculture interacted directly with WAG members and committed to connecting the women to various development projects. WAGs in Abia connected their members to ADPs as well as the state's Ministry of Rural Development and Poverty Reduction to support women's agriculture-related businesses.

In other states such as Kebbi and Niger, plans were in place to link WAGs with financial institutions so that women's businesses were supported financially. According to one state-level official in Kebbi, program staff also planned to connect WAGs with value-chain actors such as rice and ground nut producers so that women were supported by institutions that would "patronize what they are producing" (state-level project staff, KII, Kebbi).

Respondents shared that LGAs in Kebbi and Ogun states developed strong linkages in both rural and urban areas with financial institutions, service providers, and local buyers. One project staff member shared how they successfully connected WAG members to a value-chain development program:

> … through this program, we've been able to link many farmers to banks and to link good suppliers. Even the program has been able to supply input to farmers at a reduced cost and we're facilitating this especially for our women farmers (LGA-level project staff, KII).

Within our sample, Odeda in Ogun State was the only rural LGA to have established linkages between its WAGs and private partners, possibly because programming started earlier in Ogun than in the coverage areas. While the LGAs in Taraba (Bali and Zing) had not yet developed the same level of linkages as observed in Ogun, interview respondents shared that project staff were providing referrals to WAG members' products.

Expenditures

The NFWP was associated with higher relative consumption of assets in the short term. NFWP households spent a larger proportion of their consumption expenditure on assets and a lower proportion on other expenses, indicating a substitution toward consumption of assets. The impact on asset consumption includes higher consumption of consumer durables, mobile phones, smartphones, and non-mechanized farm equipment (Tables 22 and 23).

Table 22. Single-Difference Impact Estimates: Expenditures (NGN) in the Past Year—Productive Assets

	Large livestock	Small livestock	Poultry	Non-mechanized farm equipment	Mechanized farm equipment	Non-farm business	Uncultivated land
NFWP	-4,764	22,601	32,740	1,369	17,019	1,095	-2,154
	(10,043)	(18,981)	(26,034)	(492)***	(14,093)	(893)	(5,767)
Comparison Mean	29,835	11,427	5,152	4,067	3,072	3,038	18,058
N	4762	4762	4763	4762	4763	4762	4750

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, ownership of assets at baseline,
and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table 23. Single-Difference Impact Estimates: Expenditures (NGN) in the Past Year—Consumption Assets

	Housing	Large consumer durable goods	Small consumer durable goods	Mobile phone	Smartphone	Non-agricultural land	Means of transportation	Other assets
NFWP	337,403	2,523	1,193	1,111	990	449,061	-106,334	15
	(220,109)	(1,304)*	(681)*	(318)***	(529)*	(389,570)	(98,283)	(1,160)
Comparison Mean	64,246	4,442	4,578	3,090	4,659	10,606	196,392	2,733
N	4761	4763	4763	4760	4757	4754	4761	4754

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, ownership of assets at baseline,
and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Impact Evaluation of the Nigeria for Women Project: Baseline Report

Table 24. Single-Difference Impact Estimates: Expenditures (NGN) in the Past Year by Asset Category

	Asset consumption	Asset consumption (excluding land)	Asset consumption (excluding land and housing)	Productive assets consumption
NFWP	754,487	305,054	-33,064	68,105
	(628,057)	(261,762)	(105,109)	(46,423)
Comparison Mean	368,194	357,591	292,797	74,859
N	4722	4722	4722	4747

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, ownership of assets at baseline,
and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Women's Empowerment

The NFWP does not seem to have affected women's decision-making power in the short term. Across a series of adapted A-WEAI indicators, no significant differences emerged by treatment status for women's decision-making power as it pertains to control over resources, namely assets, credit, and income (Table 25). A large majority of NFWP beneficiaries reported decision-making power on the use of income (91%), but only a small share had influence over the use of credit at the household level (2%), which is largely attributed to a lack of credit access among other household members.

Despite differences in ownership of individual assets, differences by treatment status in overall asset ownership were minimal and not statistically significant, regardless of whether the assets were jointly or solely held and whether the assets were large or small. While we earlier saw some indication that sole ownership of assets increased among women in the treatment group, women in NFWP areas whose households owned productive assets were 3 percentage points less likely to own productive assets in comparison with households in the comparison group that owned productive assets.

As expected, the treatment group had higher rates of group membership (81%) relative to women in comparison areas, however. Among NFWP beneficiaries, impacts on group membership were larger for northern regions and lower for non-Christian beneficiaries (Table D15b and D15c in Appendix D), suggesting that the NFWP had larger effects on group membership for Muslims in Northern Nigeria than for Christians.

Impact Evaluation of the Nigeria for Women Project: Baseline Report

Table 25. Single-Difference Impact Estimates: Adapted A-WEIA Domain Indicators

	Productive asset ownership	Sole or joint owner of at least one large asset	Sole or joint owner of at least one asset	Sole owner of at least one large asset	Sole owner of at least one asset	Access to and decisions on credit	Control over use of income	Group membership
NFWP	-0.033	0.020	0.000	0.002	-0.000	0.012	-0.018	0.145
	(0.020)*	(0.043)	(0.021)	(0.023)	(0.021)	(0.010)	(0.017)	(0.035)***
Comparison Mean	0.94	0.75	0.88	0.81	0.83	0.02	0.93	0.66
N	1911	4764	4763	4763	4763	4788	3793	4789

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Estimations control for State fixed effects, women's demographics (age, education, and marital status),

household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),

indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity.

Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Qualitative research respondents did report that the NFWP enabled women to support their families financially, which helped them gain independence and respect in their households. In focus groups, women claimed that their participation in the savings and loan groups provided them additional money to support their families, which enabled them to pay their children's school fees and feed their families. The spouses of WAG members tended to express appreciation for the additional income that women were able to bring. One spouse in Niger State explained, "Before our wives wait for us to do everything even the smallest expenses but now it is not so, they do things now on their own … This has brought about unity and peace in the house" (spouse of WAG member, FGD, Niger). Men also reported feeling like they could talk to their wives more about finances, now that they understand business. Although husbands still made decisions for their wives and families in many ways, all respondents broadly affirmed the belief that women who earned their own money had more agency to make decisions about how to use it.

We did not find statistically significant effects on women's decision-making regarding their mobility. While a vast majority of women reported decision-making input on activities in which they participated, autonomy over their personal mobility was comparatively lower. Less than half of the women reported that they alone decided whether they could go to the market (41% to 45%), visit family or relatives (40% to 44%), visit a neighbor or friend (44% to 48%), attend a training (41% to 44%), or travel outside the community (39% to 43%) (Table 26). These differences were not statistically significant between the treatment and comparison group. The smaller proportion of women who reported decision-making power over their personal mobility perhaps reflects the strength of prevailing gender norms and roles within the household.

Table 26. Single-Difference Impact Estimates: Decision-Making Power—Mobility

	Mobility decision-making power, proportion affirmative, out of 5	Solely decides to go to the market	Solely decides to visit family or relatives	Solely decides to visit friend or neighbor's house	Solely decides to attend a training	Solely decides to go outside of community or village
NFWP	0.036	0.044	0.036	0.036	0.030	0.036
	(0.095)	(0.095)	(0.096)	(0.094)	(0.096)	(0.097)
Comparison Mean	0.41	0.41	0.40	0.44	0.41	0.39
N	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Similar to decision-making power, the NFWP did not seem to impact women's sense of self-efficacy in the short term. Across a set of eight statements, women expressed confidence in their ability to achieve personal goals, and to overcome barriers they may encounter in doing so (Table B12 in Appendix B), but we did not find statistically significant differences between treatment and comparison LGAs.

However, women across regions and across urban and rural settings believed that WAG membership increased their power to influence community decisions. One WAG member in Taraba shared how, during an election, all the different women's associations in the area used their collective power to enact change. As she explained, "The importance of this group and association cannot be overemphasized, because what we say in [the] group is better than what an individual says" (non-WAG member, FGD, Taraba). Qualitative respondents perceived that WAG membership enabled women to make their voices heard in the community. As evidence of this, various respondents mentioned that WAG members had started to tackle community issues such as roads, sanitation, and schooling, and some also reported feelings of improved self-worth, respect, and a sense of independence.

Respondents also suggested that WAGs began to facilitate social empowerment at multiple levels—within the community, among group members, and at the individual level. Qualitative interviews revealed that participation in WAGs led to improved social interactions by creating opportunities for women to interact with and encourage one another. Group membership also fostered mutual cooperation and encouraged women to support other members in need, particularly through the collection of emergency support funds. In some instances, however, WAGs may have unintentionally hindered existing community cooperation by creating a sense of exclusivity. Participants in one focus group of non-WAG members shared that they would now be less likely to help a WAG member in need, because they perceived that WAG members already had each other for support. Nonetheless, most women felt that WAGs had improved social cohesion among women.

Program Costs and Costs to Beneficiaries

We estimated per-capita program costs based on program expenditure statements provided by the SPCUs and the FPCU, and program scale from the project Management Information System. Although project activities started in 2020, mobilization of WAGs primarily started in 2021. To estimate the per-capita costs in the first 2 years, we divided costs in 2020 and in 2021 by WAG members mobilized in 2021. Figure 4 presents the distribution of program scale in 2021 across the five states. As shown here, Taraba and Niger had mobilized the highest number of women into WAGs, followed by Kebbi, Abia, and Ogun. Figure 5 shows per-capita costs over the first 2 years of project implementation in each State (in NGN).
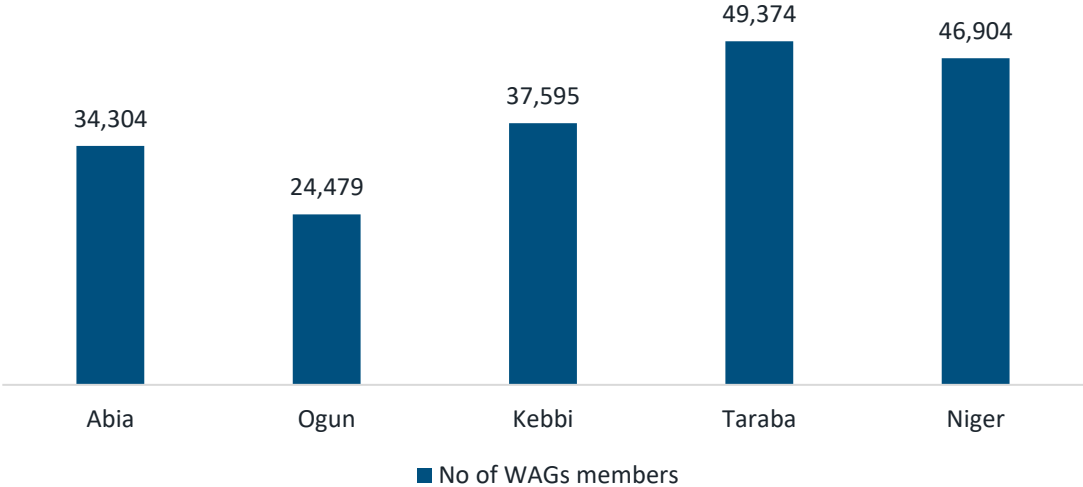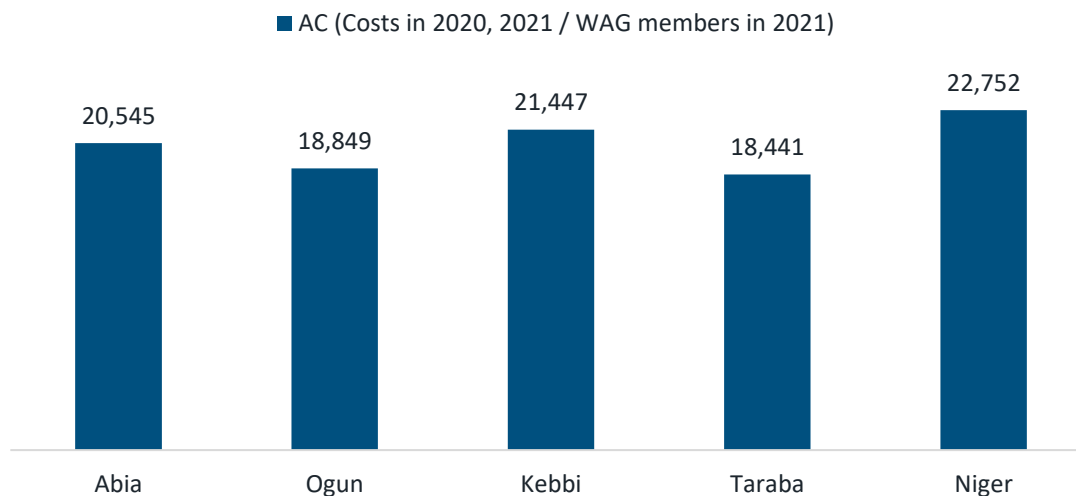
Figure 4. Program Scale in 2021

■ No of WAGs members

Figure 5. Average Costs in First 2 Years (NGN)

■ AC (Costs in 2020, 2021 / WAG members in 2021)

Abia: 20,545
Ogun: 18,849
Kebbi: 21,447
Taraba: 18,441
Niger: 22,752

Our estimates suggest that average costs were relatively comparable across the states. Niger, Kebbi, and Abia had the highest per-capita expenditures (ranging from NGN 20,500 to NGN 22,700), while costs in Ogun and Taraba were lower (at approximately NGN 18,000). Almost all these costs were incurred under the first and fourth project components: building social capital; and project monitoring, evaluation, and learning. While some states had started incurring expenses for livelihood programs (component 2) in the second year, none of the states had yet incurred any expenditures on innovations and partnerships (component 4). The cost estimation relied on two assumptions. First, we divided all costs across program scale achieved in 2021 equally. However, some of these costs are likely fixed investments which will be divided over a larger scale as the program continues mobilization after 2021. Second, we divided costs realized by FPCU across states by the proportion of WAG members in each state. These costs are comparable to the initial-year costs of similar programs in different contexts—specifically, the JEEViKA self-help group (SHG) program in Bihar, India, where average costs in Year 2 were around USD 40 (or around NGN 17,000) (Siwach et al., 2022).

Next, we consider costs to women in terms of time spent participating in group activities and trainings, as well as any real costs incurred in these activities. Table 27 compares time and costs for women who reported being WAG members at any point in the last 12 months, and for women who reported being members of non-WAG groups. These baseline analyses are meant to be exploratory and therefore only offer a descriptive comparison of time spent on group-related work for WAG members compared with other group members. On average, WAG members spent 70 hours in a year on group activities including meetings (compared with 44 hours spent by non-WAG group members), and 22 hours in a year on trainings (compared with 6 hours spent by non-WAG group members). This analysis excluded women who were not

members of any groups, which means that the average difference between all WAG members and non-WAG members will be significantly greater. On the other hand, the surveys were conducted within a year of WAG formation. Time spent on WAG activities and trainings is expected to be high for new groups, but this will likely decline as groups mature. This was indicated in the qualitative data, where women reported longer meetings and more time spent on the project in the initial months. In terms of the total cost difference, WAG members spent around NGN 12,000 more annually, which included both time costs (based on income foregone) and real costs (like expenses related to transportation or meals).

Table 27. Costs to Women

	Mean non-WAG group	Mean WAG	Diff.	Std. error	Obs.
Hours spent on group activities per year	43.58	69.56	25.99***	3.05	4218
Hours spent on group trainings in a year	5.52	21.60	16.08***	2.09	4200
Annual group costs (NGN)	14636.27	26649.58	12013.31***	4234.19	3437

Note. Annual time spent on group activities is extrapolated based on time spent on these activities in the 4 weeks prior to the survey. Annual group costs include the economic value of time spent on group activities and trainings, as well as any real costs incurred while participating in these activities. Women who were not a member of any type of group were excluded from this analysis, because their time spent on any group-related activity was 0.

Conclusion

This report presents impact estimates for the NFWP after 1 year of program implementation using a combination of DID analysis and matching across five states in Nigeria. To determine the short-term impact of the program, we used a quasi-experimental design with 15 LGAs in which the NFWP was implemented and 60 neighboring comparison LGAs. In addition, we conducted a preliminary process evaluation to determine the fidelity of program implementation and a cost analysis to estimate program expenses in the first 2 years.

In line with the theory of change, the results showed positive impacts on the likelihood of savings (by 12 percentage points) and credit (by 9 percentage points), with some indications of positive impacts on women's ownership of smaller assets. While we found positive impacts on the likelihood of savings, and on savings in women's groups and other savings groups, we did not find positive effects on cumulative savings. We hypothesize that this is because women may have invested some of their initial savings in smaller assets, such as mechanized farm

equipment, uncultivated land, and mobile phones, and in small businesses. We found evidence for positive impacts on women's sole ownership of mechanized farm equipment, mobile phones, and uncultivated land as well as on the likelihood of self-employment. However, we also found some indications of negative effects on women's ownership of livestock.

Overall, the results suggest an increase in the consumption of assets relative to other household consumption, but we did not find positive impacts on an asset index. We found positive impacts on the consumption of large and small consumer durable goods, and especially on mobile phone expenditures. These positive impacts, combined with the evidence of increased sole ownership by women of mechanized farm equipment, mobile phones, and uncultivated land indicate a substitution of food and other consumption toward the consumption of smaller assets for women. However, given that we did not find positive impacts on an asset index, the program may only have had some initial effects on asset ownership.

We also did not find short-term impacts on women's income and decision-making power, or on total consumption. We found no statistically significant differences between the treatment and comparison groups for these outcomes, which aligns with some evidence from India suggesting that women's groups with economic objectives, such as savings groups and self-help groups, may not always have positive short-term effects on women's income and decision-making power or on household-level expenditures (Hoffmann et al., 2021; Kochar et al., 2020). Midline results will indicate whether the NFWP is generating positive effects on these outcomes after the introduction of livelihoods grants.

A triangulation of quantitative and qualitative evidence suggests positive program effects on the likelihood of self-employment. We found statistically significant effects on self-employment among women respondents. Interviews with qualitative respondents revealed that investing savings in small businesses may have contributed to these positive effects.

Various implementation factors may have contributed to the modest but positive short-term effects of the program on financial inclusion and asset ownership. While some respondents initially showed a lack of trust in the NFWP, a vast majority of the women in NFWP areas who reported women's group or savings group membership belonged to WAGs at baseline. In the treatment group, 81% of the women reported membership of a women's or savings groups and 65% reported membership of a WAG. Respondents indicated that the sensitization efforts of the NFWP—including the sensitization of local leaders such as village chiefs—as well as endorsements by the Federal government contributed to building trust among WAG members. WAG members also indicated that the NFWP had a consistent presence in their communities, leading them to view it as accountable and reliable. In addition, women's social networks encouraged their participation in WAGs. Spousal support was critical because a husband's

permission is needed to join a WAG. Elder women also encouraged participation by serving as important sources of information.

WAGs had different characteristics from women's groups and savings groups in the comparison group. WAGs had a smaller number of group members and were more likely to only include women. We did not find consistent evidence of greater inclusion of marginalized women in WAGs, however. While WAG members were less likely to finish secondary school than members of other women's groups, they were also less likely to live with a disability. In addition, qualitative evidence indicates that women without an income faced challenges participating in WAGs because of the savings requirements, though WAGs were able to include women from the "missing middle" as members.

Per-capita program costs in the initial 2 years of the NFWP ranged from NGN 18,441 in Taraba to NGN 22,752 in Niger. Initial analyses of the program costs indicate that the average costs per program participant were similar across the different states and comparable to the initial costs of implementing the JEEVIKA program in Bihar, India. Average costs of savings groups and other women's groups tend to decrease considerably when programs move to scale, however, as shown by a costing analysis of the JEEVIKA program in Bihar (Siwach et al., 2022). These cost savings are important because it may not be sustainable to scale the NFWP with the current average costs per program participant.

While the program costs are relatively high, various stakeholders made suggestions to further increase the intensity of programming. For example, respondents recommended increased trainings and a more intensive sensitization effort. These recommendations may have significant cost implications, which suggests that incorporating the suggestions will likely not enable a sustainable scale-up of the program. On average, WAG members reported spending 22 hours in the year on group-related trainings. While trainings may lead to additional downstream impacts, they may also impose additional costs on beneficiaries in terms of time away from income-generating work (although this may be minimal).

Limitations

The study faces various limitations. First, it is important to recognize that we collected baseline data 1 year after the start of program implementation. As a result, we are not able to say with certainty that the treatment and comparison groups were similar before the start of the program. (However, recall data indicate that the treatment and comparison groups were comparable before the start of program implementation.) Second, this report only includes impact estimates for 1 year after the start of program implementation. As a result, the report is only able to present short-term effects of the NFWP. Third, we had limited statistical power to detect heterogeneous effects. While our sample size is relatively large, our evaluation design

includes a limited number of clusters, constraining our ability to detect heterogeneities in the impact estimates with sufficient precision. Finally, the qualitative interview protocols were designed primarily for the formative assessment rather than the process evaluation, as originally planned. As a result, the process evaluation findings generated at baseline should be considered preliminary. The process evaluation will be the primary focus of the qualitative data collection at midline.

Implications for Policy and Practice

This section presents implications of the evaluation for policy and practice, though we need to exercise some caution because our evaluation only covers the first year after the start of NFWP implementation. For this reason, we are only able to examine the short-term impacts, costs, and implementation features of the NFWP. This evaluation does not yet include an assessment of key program components, such as the individual livelihoods grants, business plan trainings, grants for collective enterprises, social norms messaging, and health and GBV layering.

Initial results suggest that the NFWP's sensitization efforts, women's social networks, and endorsements by the Federal government all likely contributed to increased WAG membership, and to the program's modest but positive effects on financial inclusion and women's asset ownership. These factors need continued support to achieve the positive effects of savings group and other women's group programming in Nigeria.

The results also indicate that religion is a key consideration when designing savings group programs such as the NFWP. Gender norms manifest as a result of both cultural and religious influences across Nigeria. When considering the NFWP's intended impact on community views toward gender, religion therefore merits attention, perhaps by including local religious leaders in messaging about gender norms. We also found positive impacts on labor force participation for Muslim women, but not for other women. Impacts on self-employment were also exclusively driven by impacts on Muslim women.

It is important to assess the program costs of additional sensitization and training activities. Various stakeholders indicated the importance of sensitization and trainings for building trust in the NFWP and increasing WAG membership. Such activities come with additional costs, however, and at this moment it is unclear whether the NFWP can be sustainably scaled up without a reduction in program costs. Previous studies indicate that the scale-up of savings group and other women's group programs with economic objectives can result in cost savings (Siwach et al., 2022). Nonetheless, in order to ensure that scale-up does not lead to a decline in implementation quality, it is critical to assess which program components are critical for

achieving additional impacts of the NFWP without generating a significant increase in costs. Previous evidence has shown that undertaking greater sensitization during mobilization— including improvising and adapting messages to target women as well as community stakeholders—is key to maintaining program quality at the outset (Majumdar et al., 2017). At this moment, the NFWP does not result in positive impacts for women's income and decision-making power or consumption. These effects may be realized over the longer term as women's self-employment generates higher income, and/or may require additional investments in individual livelihoods grants or support for collective enterprises organized by WAG members. It is critical to assess both the benefits and the costs of these additional program components.

Further, the results suggest that while the NFWP may generate benefits for economically active poor women (or the "missing middle"), it may generate fewer benefits for poor women without an income source. The latter group faces challenges when required to contribute to savings or pay additional costs for participating in WAGs. Women without an income source from wage labor or self-employment may therefore require alternative programs, such as cash transfers or graduation programs, to benefit from social protection systems in Nigeria.

At this moment, it is too early to present definitive conclusions about the impact and cost-effectiveness of the NFWP. We will continue to conduct research to determine the longer term impacts and cost-effectiveness of the NFWP using the midline and endline surveys.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. Review of Economic Studies, 72(1), 1–19.

Adegbite, O., Anderson, L., Chidiac, S., Dirisu, O., Grzeslo, J., Hakspiel, J., Holla, C., Janoch, E., Jafa, K., Jayaram, S., Majara, G., Mulyampiti, T., Namisango, E., Noble, E., Onyisha, B., Panetta, D., Siwach, G., Sulaiman, M., Walcott, R., Desai, S., & de Hoop, T. (2022). Women's groups and COVID-19: An evidence review on savings groups in Africa. Gates Open Research.

Anderson, C.L., de Hoop, T., Desai, S., Siwach, G., Meysonnat, A., Gupta, R., Haroon, N., Howlett, M., Kolla, N., Sidhu, A., Paul, S., Belyakova, Y., & Singh, R.S. (2019). Investing in women's groups: A portfolio evaluation of the Bill & Melinda Gates Foundation's Investments in South Asia and Africa. Retrieved from https://womensgroupevidence.org/sites/default/files/Investing-in-Women-s-Groups-A-Portfolio-Evaluation-of-the-Bill-and-Melinda-Gates-Foundation-s-Investments-in-South-Asia-and-Africa.pdf

Barooah, B., Chinoy, S. L., Dubey, P., Sarkar, R., Bagai, A., & Rathinam, F. (2019). Improving and sustaining livelihoods through group-based interventions: Mapping the evidence (3ie Evidence Gap Map Report 13). New Delhi, India: International Initiative for Impact Evaluation (3ie).

Blattman, C., Green, E. P., Jamison, J., Lehmann, M. C., & Annan, J. (2016). The returns to microenterprise support among the ultrapoor: A field experiment in postwar Uganda. *American Economic Journal: Applied Economics,* 8(2), 35–64.

Brody, C., Hoop, T., Vojtkova, M., Warnock, R., Dunbar, M., Murthy, P., & Dworkin, S. L. (2015). Economic self-help group programs for improving women's empowerment: A systematic review. *Campbell Systematic Reviews*, 11(1), 1–182. https://doi.org/10.4073/csr.2015.19

Chinen, M., de Hoop, T., Alcázar, L., Balarin, M., & Sennett, J. (2017). *Vocational and business training to improve women's labour market outcomes in low- and middle-income countries: a systematic review. Campbell Systematic Reviews*, *13*(1), 1-195. https://doi.org/10.4073/csr.2017.16

Collins, D., Morduch, J., Rutherford, S., & Ruthven, O. (2010). *Portfolios of the poor: How the world's poor live on $2 a day.* Princeton, NJ: Princeton University Press.

De Hoop, T., Peterman, A., & Anderson, L. (2020). *Guide for measuring women's empowerment and economic outcomes in impact evaluations of women's groups.* Retrieved from http:///www.womensgroupevidence.org

De Hoop, T., Mulyampati, T., Namisango, E., Ojambo, K., Otike, C., Muhidini, N., Okiria, E., Kalyango, R., Wangi, R., Ndagire, R., Obuku, E., & White, H. (2020). The evidence base and gaps related to women's groups in Uganda. (Working Paper).

De Hoop, T., Molotsky, A., Paek, C., Siwach, G., Castro-Zarzur, R., Adegbite, O., & Adeoye, I. (2021). *Impact evaluation of the Nigeria for Women Project: Inception report.*

De Hoop, T., Adeyeye, J., Dirisu, O., Akinola, A., Meysonnat, A., Siwach, G., & Anderson, L. (2022). Taking stock of the evidence on women's groups in Nigeria. Retrieved from https://womensgroupevidence.org/sites/default/files/2022-07/StockTakingEvidence_Nigeria_30June2022_Final.pdf

Desai, P., Speed, M., & MacLean, L. (2018). *Nigeria for Women Project, social analysis—final report.* London, UK: Social Development Direct.

Desai, S., de Hoop, T., Anderson, L., Darmstadt, G., & Siwach, G. (2019). Learning agenda on women's groups. https://womensgroupevidence.org/sites/default/files/Learning-Agenda-on-Women's-Groups.pdf

Dìaz-Martin, L., Gopalan, A., Guarnieri, E., & Jayachandran, S. (2022). Greater than the Sum of the Parts? Evidence on Mechanisms Operating in Women's Groups, *World Bank Research Observer*, https://doi.org/10.1093/wbro/lkac001

Hoffmann, V., Rao, V., Surendra, V., & Datta, U. (2021). Relief from usury: Impact of a self-help group lending program in rural India. *Journal of Development Economics*, *148*, 102567.

Karlan, D., Savonitto, B., Thuysbaert, B., & Udry, C. (2017). Impact of savings groups on the lives of the poor. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(12), 3079–3084. https://doi.org/10.1073/pnas.1611520114

Kochar, A., Barooah, B., Jain, C., Closepet, N., Narayanan, R., Sarkar, R., & Shah, R. (2020). Impact Evaluation of the National Rural Livelihoods Programme. International Initiative for Impact Evaluation. *Impact Evaluation Report* (128).

Majumdar, S., Rao, V., & Sanyal, P. (2017). On the frontlines of scaling-up: a qualitative analysis of implementation challenges in a CDD project in rural India. World Bank Policy Research Working Paper, (8039).

Malapit, H., Kovarik, C., Sproule, K., Meinzen-Dick, R.S., & Quisumbing, A.R. (2020). Instructional guide on the abbreviated Women's Empowerment in Agriculture Index (A-WEAI). Washington, D.C.: International Food Policy Research Institute (IFPRI). http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/129719

Meysonnat, A., Wineman, A., Anderson, C.L., Adegbite, O., de Hoop, T., Mbabazi, G., & Desai, S. (2022a). *Exposure to women's groups and empowerment in agricultural households: Evidence from Nigeria and Uganda*. ECWG Working Paper 08.

Meysonnat, A., Wineman, A., Anderson, C.L., Adegbite, O., de Hoop, T., & Desai, S. (2022b). Participation in informal savings groups and women's empowerment in agriculture in Nigeria. ECWG Working Paper 09 Retrieved from https://womensgroupevidence.org/

Mikkelsen, B. (2005). *Methods for development work and research: A new guide for practitioners*. New Delhi, India: SAGE Publications India. doi:10.4135/9788132108566

Namisango, E., de Hoop, T., Holla, C., Siwach, G., Chidiac, S., Jayaram, S., Grzeslo, J., Sulaiman, M., Janoch, E., Majara, G., Adegbite, O., Anderson, L., Walcott, R., Jafa, K., Desai, S., Dirisu, O., Mulyampiti, T., Hakspiel, J., & Panetta, D. (2021). *Women's Groups and COVID-19: Effects on and challenges of savings groups* Retrieved from http:///www.womensgroupevidence.org

National Population Commission (NPC) [Nigeria] and ICF. (2019). Nigeria Demographic and Health Survey 2018 [Dataset]. Abuja, Nigeria, and Rockville, Maryland, USA: NPC and ICF.

Qasim, F. & Asubaro, W. (2019). Gender-Based Violence: An Analysis of the Implications for the Nigeria for Women Project (English). Washington, D.C.: World Bank Group. http://documents.worldbank.org/curated/en/731521555064925144/Gender-Based-Violence-An-Analysis-of-the-Implications-for-the-Nigeria-for-Women-Project

Rychetnik, L., Frommer, M., Hawe, P., & Shiell, A. (2002). Criteria for evaluating evidence on public health interventions. *Journal of Epidemiology and Community Health*, *56*(2), 119–127.

Siwach, G., Paul, S., de Hoop, T. (2022). Economies of scale of large-scale international development interventions: Evidence from self-help groups in India. *World Development, 153*. 105839.

World Bank (2018). Project appraisal document on a proposed credit to the Federal Republic of Nigeria in the amount of 69.6 SDR Million (US$100 Million Equivalent) for a Nigeria for Women Project. Retrieved from https://documents1.worldbank.org/curated/en/207671530329469779/pdf/NIGERIA-PAD-05252018.pdf

Appendix A. Full Description of Theory of Change

The description of the Theory of Change (ToC) focuses on a generic ToC model for Nigeria and does not distinguish between groups that existed before the intervention that were transformed into WAGs or newly formed WAGs, because all groups will receive the full package of trainings regardless of their prior status or type. In this way, the ToC focuses on the pathways through which the intervention components are expected to generate impacts at different points in time, differentiating between the unit of measurement of the changes (i.e., group, individual and household, and community). In addition, we include potential moderators to generate hypotheses about how WAGs with different program components, different membership composition, differences in the group's implementation before the transformation to WAGs, and in other contexts may show different effects along the causal chain of the ToC.

The ToC suggests that the NFWP can achieve improvements in women's economic empowerment through several mechanisms. Savings and livelihoods trainings can lead to greater access to savings mechanisms and increases in collective savings, which can in turn enable women to gain access to individual or group credit. Combining increased access to credit with improved business skills can then help women start or expand their income-generating activities across high-productivity agricultural and non-agricultural sectors. The creation and expansion of women's individual and collective businesses could increase women's income and asset ownership, which could stimulate their bargaining power within the household.

Social norm messages at the group- and community-level, which aim to change discriminatory social and gender beliefs and practices held by both men and women, could lead to changes in women's empowerment and social norms at the community level. Increased social support within groups can help WAG members reflect on gender norms, change group members' attitudes, and improve women's self-confidence and well-being. These changes in women's attitudes could later result in improvements in women's bargaining power within the household, as women could leverage the economic and social opportunities brought to them through the NFWP. Social norm messages at the community level could lead to additional attitudinal changes for men and other community members and could lead to changes in gender and other social norms. However, changes in social norms generally are non-linear and women may have to transgress social norms to change them. Such transgressions could lead to community backlash, which may create dynamics in which women return to their level of bargaining power before the NFWP because of social sanctions (de Hoop et al., 2014; Brody et al., 2015). The NFWP aims to limits such backlash by sending social norms messages at the community level as well as by engaging through community dialogues, which could lead to changes in social and gender beliefs and practices held by men and other non-WAG members.

As a first step, community sensitization and mobilization will include meeting stakeholders such as leaders and women in their safe spaces with the aim of introducing the project to the community and getting the approval to operate. Trained Ward Facilitators will then identify the women targeted by the program who are either not yet part of any women's group or are part of existing women's groups. Community leaders, peers, and existing women's groups will then encourage women to join a WAG in their community.

Once women are part of a group, they could benefit from livelihood support through trainings, the preparation of business plans, livelihood grants, and increased credit from their WAG. The grants and group credit could enable women to invest in existing or new individual or collective enterprises, which could generate profits and additional household income and could enable women to accumulate assets, especially after the preparation of business plans and livelihoods training help them to make productive investments across diversified agricultural and non-agricultural income-generating activities. The intervention also stimulates partnerships with outside actors, such as the private sector, civil society organizations, NGOs, and individuals, to identify innovations for the groups through a development marketplace. These partnerships provide women with the opportunity to strengthen networks and increase their knowledge and skills (e.g., further improve their business skills and complement what they would have learned through group trainings). The partnerships, networks, and skills development can also give the participating women a comparative advantage and can lead to increased access to financial markets, greater economic opportunities, and more income generation. This market access could result in synergies that could lead to additional effects of the livelihood support, business plans, and livelihood grants.

From Year 2, individual WAG members could also benefit from the health-layering component through health trainings and health-related group activities. The specific focus on health may contribute to women's knowledge about topics such as nutrition and maternal or neo-natal health. We hypothesize that this knowledge could result in changes in health behavior, including changes in health care spending, health-seeking behavior, usage of preventative and pre-natal care, and infant and young child feeding practices.

The generation of social capital through group support could improve trust, social networks, and social cohesion by enabling women to collaborate with peers in their community and discuss economic and social issues. We hypothesize that improved social cohesion and networking could have both individual- and group-level effects. For instance, discussions about social norms may affect women's attitudes toward empowerment and their psychological empowerment. The sense of confidence, dignity, and self-esteem that comes from women's empowerment could help women achieve "power within" and could positively influence intra-household dynamics—for example, by enabling women to gain bargaining power within the

household with regard to financial decision and decisions on education and health (Barooah et al., 2019; Dìaz-Martin, Gopalan, Guarnieri, & Jayachandran, 2022).

Social norms trainings can bring additional improvements in different women's empowerment domains because of their specific focus on raising awareness about gender stereotypes and gender-based violence (GBV). In addition, the community-level social norms intervention is designed to further strengthen the program's impacts on intra-household decision-making and GBV, as household and village members (men and women) engage in community activities and receive targeted messages (via various channels, including but not limited to radio programming, household and community dialogues, theater for change etc.), thereby reflecting upon these issues and possibly adopting better attitudes and behaviors.

The WAGs can also generate impacts at the group level. First, the composition of groups may change considerably after encouraging membership of marginalized women. These changes in group composition may lead to increased diversity either within groups or across groups, which may strengthen social cohesion within the community and may provide women with opportunities to pool risks and resources, thereby improving their resilience — particularly the resilience of marginalized women. Importantly, however, increased within-group diversity may limit the ability of women's groups to pursue joint goals; women with different characteristics likely have different objectives and different means to contribute to savings, which could limit the ability of groups to pool savings and risks. In contrast, the number of women's groups may increase if marginalized women form their own groups to pursue joint financial objectives.

The within-group structures of rotating leadership, group-based decision-making, and regularly planned meetings can help build trust, shared responsibility, and a sense of community. These group-level changes can lead to collective action, such as increasing the frequency of planned meetings or of interactions with third parties regarding training and innovation. Ultimately, these changed group behaviors may result in an increase in individual or collective savings, greater access to formal credit, and a reduction in fraud.

With the support of the Ward Facilitators, WAGs could spark the formation of livelihoods collectives, such as cooperatives, farm and non-farm producer organizations, social enterprises, producer-governed private limited companies, especially after groups receive livelihoods training and livelihoods grants. These trainings and grants can help groups set up livelihoods collectives that include all or a subset of WAG members. By leveraging the groups' increased human capital and members' social networks, and by gaining access to grants (for which collectives are eligible), women's livelihoods collectives could break into higher-productivity agricultural and non-agricultural sectors, improve their access to markets and value chains, and negotiate better transaction conditions (e.g., prices and quantities). Successful livelihoods collectives may also lead to improvements in individual and collective asset ownership and profits.

Increases in the number of WAGs can lead to general equilibrium effects and changes in prices and wages. For example, WAGs may increase demand for credit from women's groups, which likely will become available at lower interest rates than in informal settings. The increased demand for credit from WAGs may come at the expense of demand for informal credit, which could in turn result in a reduction in informal interest rates. The formation of collective enterprises may also result in increases in the demand for labor, which could in turn lead to increases in wages.

Achieving impacts through the described pathways depends on several critical factors. First, the successful mobilization of women will depend on the motivation and incentives of community leaders to identify and target marginalized women who do not yet participate in any well-functioning women's group. In addition, women should have sufficient time and a minimum level of resources to enable their participation in meetings, create a network, and contribute to collective savings. Using accumulated savings to make productive investments in the future further depends on women's ability to expand livelihood activities. In addition, women's savings can only increase sustainably if they can avoid having to take out a large portion of the money in the interim. Furthermore, not all changes will occur at the same time, and behavioral and changes in social norms in the community are likely slow and highly dependent on context and on the existing social norms. Finally, group-level impacts depend on the social cohesion among group members, regular meetings and savings, and the return on investment of investments in collective enterprises.

Although improvements in women's bargaining power in the household can reduce GBV, the current evidence on this mechanism is mixed. Some studies suggest that greater bargaining power can lead to backlash and possibly an increase in GBV. Nonetheless, in the long term, any backlash may reduce because of changes in social norms at the community-level (Brody et al., 2015). We hypothesize that women in communities benefiting from the social norm messages at the community-level and the WAGs may therefore have a larger likelihood of seeing reductions in GBV than women in communities with WAGs without social norms messages. This is particularly likely for women benefiting from the layering of additional GBV-interventions.

Different initial group types may lead to different impacts as they may condition the way WAGs evolve after the first WAG cycle ends (i.e., after trainings and Ward Facilitator support end and groups decide whether to continue saving or not). WAGs may have different effects depending on what components they include, and WAGs could evolve as part of their institutional evolution. For example, some groups may continue to focus on savings and credit after the first cycle, whereas other groups may choose to focus on social activities. We hypothesize that the former group type may have larger impacts on women's financial inclusion, which may result in improvements in women's asset ownership and women's bargaining power in the household. Secondary benefits may include improvements in women's agency, and business opportunities.

We hypothesize that the latter group type (i.e., social groups) may have larger effects on women's psychological empowerment if discussions about social issues focus on building women's self-confidence. Secondary benefits of these groups may include improvements in women's bargaining power, which could in turn result in improvements in women's agency, and economic outcomes, such as asset ownership and income.

Different group types also impose different costs, both program-level costs and opportunity costs related to the time women spend in group meetings and other group activities. Adding components will add ingredients to the intervention, which will increase the program costs. Furthermore, adding intervention layers and trainings may increase the time group members have to spend in group meetings and thus the opportunity costs of participation in WAGs. This may lead to changes in the composition of the group, for example by reducing the participation of women who face time constraints or have higher opportunity costs. Adding new topics to a group will either increase the time group members spend in meetings or reduce the time the groups can spend on other topics.

Based on the theory of change and impact evaluations of comparable programs in India (e.g., Hoffmann et al., 2021; Kochar et al., 2020), we hypothesize positive impacts on savings, credit, and asset ownership after one year of program implementation. In addition, there is some evidence that the program may have short-term impacts on women's decision-making power (Brody et al., 2017). However, we do not hypothesize positive impacts on consumption, income or larger assets after one year of program implementation, especially because impact evaluations of comparable programs did not show positive short-term impacts on these outcomes (e.g., Hoffmann et al., 2021; Kochar et al., 2020). We also do not hypothesize positive impacts on social norms because these usually require a longer time period to change.

Appendix B. Balance Tables

Table B1. Sample Composition: Household Characteristics

	Observations		Mean		
	C	**T**	**C**	**T**	**Difference**
Household size (number of members)	3591	1198	5.53	5.74	0.21
Number of household members under 5 years old	3591	1198	0.95	1.0	0.05
Number of household members over 50 years old	3591	1198	0.42	0.45	0.03
Has electricity	3591	1198	53%	54%	0.01
Owns their home	3591	1198	81%	79%	-0.02
Concrete or cement floors	3591	1198	64%	71%	0.07
Iron sheet roof	3591	1198	76%	81%	0.04
Drinking water from borehole	3591	1198	42%	42%	0.00
Drinking water from a river, lake, pond	3591	1198	21%	26%	0.05
Has an uncovered pit latrine	3591	1198	19%	21%	0.02
Has a covered pit latrine	3591	1198	23%	23%	0.00
Has a toilet (i.e., flush to septic tank, or flush to sewage)	3591	1198	18%	17%	-0.01

Note: C stands for "Comparison"; T stands for Treatment. LGA-clustered standard errors in parentheses.
* p<0.1; ** p<0.05; *** p<0.01

Table B2. Sample Composition: Female Respondents

	Observations		Mean		
	C	**T**	**C**	**T**	**Difference**
Average age (in years)	3482	1149	39	39	0.13
Household head	3591	1198	10%	11%	0.01
Married	3591	1198	85%	84%	-0.01
No education	3591	1198	40%	40%	-0.01
Primary education (some or complete)	3591	1198	24%	25%	0.01
Secondary education (some or complete)	3591	1198	25%	25%	-0.00

Impact Evaluation of the Nigeria for Women Project: Baseline Report

	Observations		Mean		
	C	T	C	T	Difference
More than secondary education (some or complete)	3591	1198	10%	11%	0.00

Note: C stands for "Comparison"; T stands for Treatment. LGA-clustered standard errors in parentheses.
* p<0.1; ** p<0.05; *** p<0.01

Table B3. Sample Composition: Male Respondents

	Observations		Mean		
	C	T	C	T	Difference
Man's age (in years)	3578	1185	42.59	43.51	0.92
Man is head of household	3578	1185	57%	65%	0.08**
Man is married	3578	1185	84%	83%	-0.00
Man has no education	3578	1185	28%	26%	-0.01
Man has primary education	3578	1185	22%	24%	0.02
Man has secondary education	3578	1185	33%	33%	0.01
Man has higher education	3578	1185	17%	16%	-0.01
Man is spouse of the woman respondent	3578	1185	47%	51%	0.04
Man is (grand)child of the woman respondent	3578	1185	5%	7%	0.02*

Note: C stands for "Comparison"; T stands for Treatment. LGA-clustered standard errors in parentheses.
* p<0.1; ** p<0.05; *** p<0.01

Table B4. Financial Inclusion: Savings and Access to Credit

	Observations		Mean		
	C	T	C	T	Difference
Received livelihoods grant through SG	1350	341	7%	6%	-0.010
Received livelihoods grant through WG	1555	874	8%	8%	-0.002

Note: C stands for "Comparison"; T stands for Treatment. LGA-clustered standard errors in parentheses.
* p<0.1; ** p<0.05; *** p<0.01

Impact Evaluation of the Nigeria for Women Project: Baseline Report

Table B5. Household's Baseline Asset Ownership (December 2020)

Asset	Observations		Mean		
	C	T	C	T	Difference
Large livestock	3578	1185	21%	19%	-0.019
Small livestock	3578	1185	44%	45%	0.008
Poultry	3578	1185	45%	45%	0.002
Non-mechanized farm equipment	3578	1185	77%	74%	-0.034
Mechanized farm equipment	3578	1185	9%	9%	0.001
Non-farm business equipment	3578	1185	17%	17%	-0.004
Owns house or building	3578	1185	59%	58%	-0.017
Uncultivated land	3578	1185	33%	36%	0.028
Large durable goods	3578	1185	39%	40%	0.011
Small durable goods	3578	1185	69%	69%	-0.000
A mobile phone	3578	1185	78%	77%	-0.015
A smartphone	3578	1185	22%	25%	0.032
A means of transport	3578	1185	42%	49%	0.065*

Note: C stands for "Comparison"; T stands for Treatment. LGA-clustered standard errors in parentheses.
* p<0.1; ** p<0.05; *** p<0.01

Table B6. Women's Labor Force Participation and Income Diversity (December 2020)

	Observations		Mean		
	C	T	C	T	Difference
Labor participation in 2020	3591	1198	87%	88%	0.008
Self-employment 2020	3591	1198	58%	57%	-0.012
# of IGA 2020	3591	1198	1.282	1.31	0.028

Note: C stands for "Comparison"; T stands for Treatment. LGA-clustered standard errors in parentheses.
* p<0.1; ** p<0.05; *** p<0.01

Impact Evaluation of the Nigeria for Women Project: Baseline Report

Table B7. Household Labor Force Participation and Income Diversity (Excluding Female Respondent)

	Observations		Mean		
	C	T	C	T	Difference
Labor participation in 2020	3578	1185	90%	90%	0.004
# of IGA 2020	3578	1185	1.614	1.708	0.094

Note: C stands for "Comparison"; T stands for Treatment. LGA-clustered standard errors in parentheses.
* p<0.1; ** p<0.05; *** p<0.01

Table B8. Agriculture

	Observations		Mean		
	C	T	C	T	Difference
Household cultivates land	3578	1185	82%	85%	0.031
# of plots cultivated by household	3578	1185	3.101	3.236	0.135
Area of 1 & 2 largest plots (square meters)	3578	1185	11000	16000	5566.7
% of harvest consumed by household	3487	1165	44%	42%	-1.745
% of household income from farm	3435	1144	36%	36%	0.468
Sales of farm/livestock (12 months) (NGN)	3432	1143	180000	180000	736.2

Note: C stands for "Comparison"; T stands for Treatment. LGA-clustered standard errors in parentheses.
* p<0.1; ** p<0.05; *** p<0.01

Table B9. Woman Has Input on Decision-Making: Activities

	Observations		Mean		
	C	T	C	T	Difference
Staple grain farming/harvest	1084	410	92%	91%	-0.01
Livestock raising/processing	339	110	96%	94%	-0.02
Poultry or other small animal raising	330	138	95%	89%	-0.06**
Non-farm economic activities	1567	501	96%	96%	-0.00
Wage and salary employment	173	70	91%	89%	-0.02
Routine household purchases	1176	351	95%	95%	0.01
Borrowing for family	247	94	94%	90%	-0.04
Saving for family	641	201	94%	94%	0.00

Impact Evaluation of the Nigeria for Women Project: Baseline Report

	Observations		Mean		
	C	T	C	T	Difference
Gardens or high-value crop farming	169	69	93%	93%	-0.00

Note: C stands for "Comparison"; T stands for Treatment. LGA-clustered standard errors in parentheses.
* p<0.1; ** p<0.05; *** p<0.01

Table B10. Woman Has Input on Decision-Making: Income From Activities

	Observations		Mean		
	C	T	C	T	Difference
Staple grain farming/harvest	1075	407	92%	89%	-0.03
Livestock raising/processing	324	110	93%	93%	-0.00
Poultry or other small animal raising	320	134	93%	90%	-0.04
Non-farm economic activities	1561	499	96%	95%	-0.01
Wage and salary employment	174	70	94%	94%	0.01
Routine household purchases	1127	338	94%	96%	0.02
Borrowing for family	245	95	94%	91%	-0.04
Saving for family	637	202	93%	95%	0.01
Gardens or high-value crop	161	67	91%	90%	-0.02

Note: C stands for "Comparison"; T stands for Treatment. LGA-clustered standard errors in parentheses.
* p<0.1; ** p<0.05; *** p<0.01

Table B11. Women Has Input on Decision-Making: Personal Mobility

	Observations		Mean		
	C	T	C	T	Difference
Solely decides to go to the market	3591	1198	40%	45%	0.05
Solely decides to visit family or relatives	3591	1198	40%	43%	0.04
Solely decides to visit a neighbor's or friend's house	3591	1198	44%	47%	0.04
Solely decides to attend a training from an NGO or other organization	3591	1198	40%	44%	0.04
Solely decides to go outside the community	3591	1198	39%	43%	0.04

Note: C stands for "Comparison"; T stands for Treatment. LGA-clustered standard errors in parentheses.

Impact Evaluation of the Nigeria for Women Project: Baseline Report

Table B12. Women's Self-Efficacy

	Observations		Mean		
	C	**T**	**C**	**T**	**Difference**
Self-Efficacy Index	3591	1198	7.34	7.29	-0.04
I will be able to achieve most of the goals that I have set for myself.	3591	1198	92%	89%	-0.02
When facing difficult tasks, I am certain that I will accomplish them	3591	1198	88%	89%	0.00
In general, I think that I can obtain outcomes that are important to me.	3591	1198	93%	92%	-0.01
I believe I can succeed at most any endeavor to which I set my mind.	3591	1198	93%	93%	-0.00
I will be able to successfully overcome many challenges.	3591	1198	91%	90%	-0.01
I am confident that I can perform effectively on many different tasks.	3591	1198	92%	92%	0.00
Compared to other people, I can do most tasks very well.	3591	1198	92%	93%	0.01
Even when things are tough, I can perform quite well.	3591	1198	92%	91%	-0.01

Note: C stands for "Comparison"; T stands for Treatment. LGA-clustered standard errors in parentheses.
* p<0.1; ** p<0.05; *** p<0.01

Appendix C. Additional Impact Estimates

Table C1. Single-Difference Impact Estimates: Group-Based Training Topics—Financial Literacy and Livelihoods

	Bookkeeping	Savings	Credit	Business	Financial literacy	Improved agricultural processes	Other livelihoods trainings	Microcredit plan
NFWP	0.019	0.118	0.030	0.073	0.040	0.014	0.027	0.001
	(0.005)***	(0.019)***	(0.006)***	(0.014)***	(0.007)***	(0.004)***	(0.007)***	(0.001)
Comparison Mean	0.00	0.02	0.00	0.01	0.00	0.00	0.01	0.00
N	4789	4789	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)

Table C2. Single-Difference Impact Estimates: Group-Based Training Topics—Health, Norms, and Women's Rights

	Health education	Social norms	Women's rights
NFWP	0.005	0.042	0.031
	(0.003)	(0.007)***	(0.008)***
Comparison Mean	0.01	0.02	0.01
N	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)

Table C3. Single-Difference Impact Estimates: Group-Based Training Topics—Women's Group Concepts

	Women's group/WAG concept and management	Women's group credit linkages	Women's group concept and management and leadership	Women's group concept and bookkeeping
NFWP	0.027	0.003	0.006	0.010
	(0.006)***	(0.001)**	(0.002)***	(0.004)**
Comparison Mean	0.00	-0.00	0.00	0.00
N	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)

Table C4. Single-Difference Impact Estimates: Asset Indices

	PCA Productive Asset Index - sole owner	PCA Small Productive Asset Index - sole owner	PCA Productive Asset Index - sole owner	PCA Productive Asset Index - joint owner	PCA Consumption Asset Index - sole owner	Asset Index - sole ownership	Asset Index - joint ownership	Asset Index - sole or joint ownership
NFWP	0.029	0.033	0.019	-0.042	-0.055	-0.019	-0.121	-0.079
	(0.062)	(0.045)	(0.061)	(0.049)	(0.056)	(0.063)	(0.083)	(0.089)
Comparison Mean	0.03	0.00	0.02	0.04	0.06	0.06	0.09	0.06
N	4789	4789	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, ownership of assets at baseline,
and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Table C5. Single-Difference Impact Estimates: LFP and Income Diversification

	Number of days worked – last 4 weeks	Woman's income – last 4 weeks
NFWP	0.145	-2,872.329
	(0.100)	(1,904.618)
Comparison Mean	0.25	19482.83
N	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status),
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above),
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity.
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Appendix D. Heterogenous Effects

Group-Level Outcomes

Table D1. Single-Difference Impact Estimates with Heterogeneous Effects – Group Composition

	Group size	Group has more than 25 members	Average age of members	Average member age ≥30 years	Proportion of female members	Have the same religion	Have the same ethnicity	Speak the same language
NFWP	-11.551	-0.282	0.946	-0.008	0.068	0.001	-0.040	-0.089
	(4.821)**	(0.054)***	(2.574)	(0.115)	(0.017)***	(0.074)	(0.080)	(0.073)
NFWP x pre_nfwp	3.520	0.223	-0.184	0.085	-0.063	-0.036	-0.072	-0.036
	(7.273)	(0.074)***	(2.602)	(0.119)	(0.030)**	(0.093)	(0.105)	(0.096)
Comparison Mean	38.95	0.47	34.47	0.71	0.94	0.77	0.76	0.77
N	1267	1267	1267	1267	1206	1264	1266	1266

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for group age and registration status.

Table D2. Single-Difference Impact Estimates with Heterogeneous Effects – Group Member Characteristics

	Joined in the past 6 months	Joined over a year ago	Are married	Are living with a disability	Finished primary school	Finished secondary school	Have their own business	Do not have income	Receive income from group
NFWP	0.079	-0.128	0.004	-0.006	-0.000	-0.094	0.129	-0.034	0.021
	(0.073)	(0.116)	(0.029)	(0.006)	(0.080)	(0.047)**	(0.058)**	(0.051)	(0.126)
NFWP x pre_nfwp	-0.059	0.089	-0.001	-0.001	0.054	0.013	-0.076	-0.037	-0.123
	(0.071)	(0.143)	(0.038)	(0.006)	(0.094)	(0.080)	(0.064)	(0.063)	(0.126)
Comparison Mean	0.05	0.75	0.96	0.01	0.42	0.19	0.80	0.13	0.27
N	1252	1260	1266	1259	1240	1225	1264	1216	1263

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for group age and registration status

Table D3. Single-Difference Impact Estimates with Heterogeneous Effects – Group Social Cohesion

	I feel that I belong to this WG	I would rather be with a different WG	I would prefer this WG over others	Members of this WG are all striving for the same goals	Everyone in our WG wants to pursue their own goals	WG members worry about only themselves	I can count on fellow WG members if I need to borrow money	I can count on fellow WG members to accompany me to the doctor or hospital	I can count on fellow WG members to talk to about my problems	I can count on fellow WG members if I need advice
NFWP	0.011	-0.081	0.020	0.001	-0.137	-0.047	0.036	-0.001	-0.002	-0.023
	(0.012)	(0.049)*	(0.035)	(0.015)	(0.099)	(0.086)	(0.029)	(0.022)	(0.024)	(0.011)**
NFWP x	-0.007	0.046	-0.037	-0.006	0.017	0.029	0.025	-0.026	0.023	0.027

Impact Evaluation of the Nigeria for Women Project: Baseline Report

pre_nfwp	(0.010)	(0.055)	(0.044)	(0.022)	(0.110)	(0.121)	(0.041)	(0.038)	(0.025)	(0.011)**
Comparison Mean	1.00	0.15	0.97	0.97	0.52	0.45	0.91	0.97	0.97	1.00
N	1267	1265	1267	1265	1262	1263	1264	1264	1266	1266

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for group age and registration status

Table D4. Single-Difference Impact Estimates with Heterogeneous Effects – Group Meeting Frequency

	Meet weekly	Meet bi-weekly	Meet monthly
NFWP	0.340	-0.054	-0.232
	(0.061)***	(0.022)**	(0.045)***
NFWP x pre_nfwp	-0.137	-0.006	0.078
	(0.087)	(0.036)	(0.059)
Comparison Mean	0.62	0.09	0.22
N	1265	1265	1265

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for group age and registration status

Savings

Table D5a. Single-Difference Impact Estimates with Heterogeneous Effects – Group Membership

	WAG member	Other women's group member	Other savings group member
NFWP	0.615	-0.276	-0.291
	(0.043)***	(0.034)***	(0.036)***
NFWP x age_30orless	-0.076	0.038	0.027
	(0.032)**	(0.034)	(0.034)
Comparison Mean	0.00	0.62	0.65
N	4789	4789	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D5b. Single-Difference Impact Estimates with Heterogeneous Effects – Group Membership

	WAG member	Other women's group member	Other savings group member
NFWP	0.615	-0.276	-0.291
	(0.043)***	(0.034)***	(0.036)***
NFWP x north	-0.060	0.080	0.078
	(0.067)	(0.046)*	(0.048)
Comparison Mean	0.00	0.62	0.65
N	4789	4789	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D5c. Single-Difference Impact Estimates with Heterogeneous Effects – Group Membership

	WAG member	Other women's group member	Other savings group member
NFWP	0.615	-0.276	-0.291
	(0.043)***	(0.034)***	(0.036)***
NFWP x christianity	0.068	-0.068	-0.052
	(0.050)	(0.045)	(0.045)
Comparison Mean	0.00	0.62	0.65
N	4789	4789	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D6a. DID Estimations With Heterogeneous Effects – Individual Savings

	Woman saves money	Total savings (NGN)
NFWP	0.139	3,695.001
	(0.020)***	(5,191.028)
NFWP x age_30orless	0.043	-1,154.063
	(0.032)	(4,889.309)
Comparison Mean	0.71	21755.48
N	4306	4003

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Table D6b. DID Estimations With Heterogeneous Effects – Individual Savings

	Woman saves money	Total savings (NGN)
NFWP	0.139	3,695.001
	(0.020)***	(5,191.028)
NFWP x north	0.029	-4,734.427
	(0.041)	(5,127.329)
Comparison Mean	0.71	21755.48
N	4306	4003

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Table D6c. DID Estimations With Heterogeneous Effects – Individual Savings

	Woman saves money	Total savings (NGN)

Impact Evaluation of the Nigeria for Women Project: Baseline Report

NFWP	0.139	3,695.001
	(0.020)***	(5,191.028)
NFWP x christianity	-0.031	-3,626.198
	(0.033)	(5,312.007)
Comparison Mean	0.71	21755.48
N	4306	4003

Table D7a. Single-Difference Impact Estimates with Heterogeneous Effects – Group-Based Savings

	Has savings in WAG	Cumulative savings in WAG	Has savings in other WG	Cumulative savings in other WG	Has savings in other SG	Cumulative savings in other SG
NFWP	0.506	5,891.492	-0.255	-794.852	-0.233	-34.976
	(0.039)***	(745.534)***	(0.030)***	(414.603)*	(0.023)***	(50.917)
NFWP x age_30orless	-0.077	-1,871.591	-0.028	226.760	-0.030	-107.882
	(0.030)**	(939.890)*	(0.030)	(395.894)	(0.030)	(73.076)
	(0.048)	(850.443)	(0.042)	(616.954)	(0.033)*	(102.127)
Comparison Mean	0.00	0.00	0.48	2342.12	0.35	80.30
N	4789	4789	4789	4789	4789	4789

Table D7b. Single-Difference Impact Estimates with Heterogeneous Effects – Group-Based Savings

	Has savings in WAG	Cumulative savings in WAG	Has savings in other WG	Cumulative savings in other WG	Has savings in other SG	Cumulative savings in other SG
NFWP	0.506	5,891.492	-0.255	-794.852	-0.233	-34.976
	(0.039)***	(745.534)***	(0.030)***	(414.603)*	(0.023)***	(50.917)
NFWP x north	-0.079	-1,573.800	-0.008	632.878	-0.040	-51.421
	(0.061)	(1,202.116)	(0.044)	(504.448)	(0.034)	(83.250)
Comparison Mean	0.00	0.00	0.48	2342.12	0.35	80.30
N	4789	4789	4789	4789	4789	4789

Table D7c. Single-Difference Impact Estimates with Heterogeneous Effects – Group-Based Savings

	Has savings in WAG	Cumulative savings in WAG	Has savings in other WG	Cumulative savings in other WG	Has savings in other SG	Cumulative savings in other SG
NFWP	0.506	5,891.492	-0.255	-794.852	-0.233	-34.976

	(0.039)***	(745.534)***	(0.030)***	(414.603)*	(0.023)***	(50.917)
NFWP x christianity	0.069	-579.032	0.029	-868.670	0.061	42.088
	(0.048)	(850.443)	(0.042)	(616.954)	(0.033)*	(102.127)
Comparison Mean	0.00	0.00	0.48	2342.12	0.35	80.30
N	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Credit

Table D8a. DID Estimations With Heterogeneous Effects - Access to Credit

	Has outstanding loans	Value of outstanding loans (NGN) (includes 0)
NFWP	0.112	2,338.725
	(0.015)***	(852.005)***
NFWP x age_30orless	0.034	4,248.580
	(0.040)	(2,427.642)*
Comparison Mean	0.09	6287.11
N	4785	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Table D8b. DID Estimations With Heterogeneous Effects - Access to Credit

	Has outstanding loans	Value of outstanding loans (NGN) (includes 0)
NFWP	0.112	2,338.725
	(0.015)***	(852.005)***
NFWP x north	0.014	1,884.861
	(0.031)	(2,406.491)
Comparison Mean	0.09	6287.11
N	4785	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Table D8c. DID Estimations With Heterogeneous Effects - Access to Credit

	Has outstanding loans	Value of outstanding loans (NGN) (includes 0)
NFWP	0.112	2,338.725
	(0.015)***	(852.005)***

Impact Evaluation of the Nigeria for Women Project: Baseline Report

NFWP x christianity	-0.038	-3,787.496	
	(0.031)	(2,553.382)	
Comparison Mean	0.09	6287.11	
N	4785	4789	

Table D9a. Single-Difference Impact Estimates with Heterogeneous Effects - Loans

	Number of current outstanding loans	Total value of current largest loan	Value of LL that still needs to be paid (NGN)	WG/SG provided largest loan
NFWP	0.183	1,790.313	1,755.353	0.091
	(0.026)***	(1,142.062)	(1,102.225)	(0.013)***
NFWP x age_30orless	0.105	1,257.827	1,464.105	-0.010
	(0.052)**	(3,357.103)	(2,080.280)	(0.017)
Comparison Mean	0.11	7356.49	4585.95	0.02
N	4789	4789	4779	4788

Table D9b. Single-Difference Impact Estimates with Heterogeneous Effects - Loans

	Number of current outstanding loans	Total value of current largest loan	Value of LL that still needs to be paid (NGN)	WG/SG provided largest loan
NFWP	0.183	1,790.313	1,755.353	0.091
	(0.026)***	(1,142.062)	(1,102.225)	(0.013)***
NFWP x north	0.026	-2,000.095	-1,553.192	0.039
	(0.051)	(3,101.849)	(2,112.026)	(0.022)*
Comparison Mean	0.11	7356.49	4585.95	0.02
N	4789	4789	4779	4788

Table D9c. Single-Difference Impact Estimates with Heterogeneous Effects - Loans

	Number of current outstanding loans	Total value of current largest loan	Value of LL that still needs to be paid (NGN)	WG/SG provided largest loan
NFWP	0.183	1,790.313	1,755.353	0.091
	(0.026)***	(1,142.062)	(1,102.225)	(0.013)***
NFWP x christianity	-0.048	1,199.037	176.136	-0.023
	(0.057)	(3,602.282)	(2,558.614)	(0.022)
Comparison Mean	0.11	7356.49	4585.95	0.02
N	4789	4789	4779	4788

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D9d. Use of Largest Loans

	Observations		Mean		
	C	T	C	T	Difference
Health expenses	293	234	17%	21%	0.042
Invest in livestock	293	234	1%	0%	-0.010*
Invest in other assets (household or productive assets)	293	234	4%	3%	-0.011
Start agricultural business	293	234	9%	10%	0.013
Agricultural input credit	293	234	5%	3%	-0.017
Agricultural marketing credit	293	234	1%	1%	0.006
Other investment in agricultural business	293	234	16%	12%	-0.041
Start in non-agricultural business	293	234	14%	15%	0.011
Invest in existing non-agricultural business	293	234	18%	17%	-0.010
Fund own marriage or dowry	293	234	1%	1%	-0.002
Fund marriage or dowry of family member	293	234	3%	5%	0.013
Fund funeral of family member	293	234	3%	1%	-0.018
Fund other shock	293	234	3%	3%	0.003
Food purchases	293	234	17%	22%	0.052
Make home repairs/improvements	293	234	6%	3%	-0.025
Pay for a loan/debt	293	234	1%	1%	0.003
For security in case of future emergencies	293	234	1%	0%	-0.014**

Impact Evaluation of the Nigeria for Women Project: Baseline Report

	Observations		Mean		
	C	T	C	T	Difference
Education expenses	293	234	10%	9%	-0.008
Other	293	234	1%	4%	0.029**

Assets

Table D10a. DID Estimations With Heterogeneous Effects – Household Productive Asset Ownership

	HH owns large livestock	HH owns small livestock	HH owns poultry	HH owns non-mechanized farm equipment	HH owns mechanized farm equipment	HH owns non-farm business equipment	HH owns uncultivated land
NFWP	-0.016	0.002	0.020	0.022	0.002	0.007	0.028
	(0.025)	(0.020)	(0.023)	(0.011)**	(0.010)	(0.009)	(0.010)***
NFWP x age_30orless	-0.028	-0.021	0.029	0.005	0.004	-0.031	0.015
	(0.020)	(0.028)	(0.033)	(0.015)	(0.010)	(0.015)**	(0.017)
Comparison Mean	0.17	0.42	0.43	0.78	0.09	0.18	0.32
N	4763	4763	4763	4763	4763	4763	4763

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Table D10b. DID Estimations With Heterogeneous Effects – Household Productive Asset Ownership

	HH owns large livestock	HH owns small livestock	HH owns poultry	HH owns non-mechanized farm equipment	HH owns mechanized farm equipment	HH owns non-farm business equipment	HH owns uncultivated land
NFWP	-0.016	0.002	0.020	0.022	0.002	0.007	0.028
	(0.025)	(0.020)	(0.023)	(0.011)**	(0.010)	(0.009)	(0.010)***
NFWP x north	-0.034	-0.039	0.044	-0.001	-0.010	-0.002	0.004
	(0.030)	(0.027)	(0.031)	(0.015)	(0.012)	(0.012)	(0.017)
Comparison Mean	0.17	0.42	0.43	0.78	0.09	0.18	0.32
N	4763	4763	4763	4763	4763	4763	4763

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Impact Evaluation of the Nigeria for Women Project: Baseline Report

Table D10c. DID Estimations With Heterogeneous Effects – Household Productive Asset Ownership

	HH owns large livestock	HH owns small livestock	HH owns poultry	HH owns non-mechanized farm equipment	HH owns mechanized farm equipment	HH owns non-farm business equipment	HH owns uncultivated land
NFWP	-0.016	0.002	0.020	0.022	0.002	0.007	0.028
	(0.025)	(0.020)	(0.023)	(0.011)**	(0.010)	(0.009)	(0.010)***
NFWP x christianity	0.028	0.024	-0.046	0.009	0.003	0.009	-0.037
	(0.027)	(0.025)	(0.034)	(0.017)	(0.011)	(0.013)	(0.018)**
Comparison Mean	0.17	0.42	0.43	0.78	0.09	0.18	0.32
N	4763	4763	4763	4763	4763	4763	4763

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Table D11a. DID Estimations With Heterogeneous Effects - Household Consumer Asset Ownership

	HH owns a house	HH owns large durable goods	HH owns small durable goods	HH owns a mobile phone	HH owns a smart phone	HH owns a means of transportation
NFWP	0.025	-0.007	0.036	0.044	0.009	0.015
	(0.013)*	(0.011)	(0.015)**	(0.018)**	(0.019)	(0.014)
NFWP x age_30orless	-0.019	-0.009	-0.027	-0.014	-0.043	0.020
	(0.019)	(0.016)	(0.023)	(0.028)	(0.016)**	(0.025)
Comparison Mean	0.62	0.38	0.70	0.85	0.24	0.42
N	4763	4763	4763	4763	4763	4763

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Table D11b. DID Estimations With Heterogeneous Effects - Household Consumer Asset Ownership

	HH owns a house	HH owns large durable goods	HH owns small durable goods	HH owns a mobile phone	HH owns a smart phone	HH owns a means of transportation
NFWP	0.025	-0.007	0.036	0.044	0.009	0.015
	(0.013)*	(0.011)	(0.015)**	(0.018)**	(0.019)	(0.014)
NFWP x north	0.001	-0.008	0.018	0.037	0.047	0.007
	(0.016)	(0.017)	(0.019)	(0.028)	(0.024)*	(0.019)
Comparison Mean	0.62	0.38	0.70	0.85	0.24	0.42
N	4763	4763	4763	4763	4763	4763

Table D11c. DID Estimations With Heterogeneous Effects - Household Consumer Asset Ownership

	HH owns a house	HH owns large durable goods	HH owns small durable goods	HH owns a mobile phone	HH owns a smart phone	HH owns a means of transportation
NFWP	0.025	-0.007	0.036	0.044	0.009	0.015
	(0.013)*	(0.011)	(0.015)**	(0.018)**	(0.019)	(0.014)
NFWP x Christianity	-0.005	0.020	-0.018	-0.017	-0.008	-0.008
	(0.016)	(0.016)	(0.022)	(0.028)	(0.020)	(0.021)
Comparison Mean	0.62	0.38	0.70	0.85	0.24	0.42
N	4763	4763	4763	4763	4763	4763

Table D12a. Single-Difference Impact Estimates with Heterogeneous Effects - Sole Ownership of Productive Assets

	Solely owns land	Solely owns large livestock	Solely owns small livestock	Solely owns poultry	Solely owns non-mechanized farm equipment	Solely owns mechanized farm equipment	Solely owns nonfarm business equipment	Solely owns uncultivated land
NFWP	0.057	-0.026	-0.027	-0.031	0.023	0.016	0.005	0.052
	(0.016)***	(0.010)**	(0.035)	(0.022)	(0.034)	(0.006)***	(0.018)	(0.015)***
NFWP x age_30orless	0.020	-0.016	-0.089	-0.034	-0.013	0.009	0.017	-0.022
	(0.024)	(0.012)	(0.034)**	(0.025)	(0.039)	(0.007)	(0.021)	(0.019)
Comparison Mean	0.14	0.03	0.27	0.30	0.30	0.01	0.09	0.07
N	4789	4789	4789	4789	4789	4789	4789	4789

Table D12b. Single-Difference Impact Estimates with Heterogeneous Effects - Sole Ownership of Productive Assets

	Solely owns land	Solely owns large livestock	Solely owns small livestock	Solely owns poultry	Solely owns non-mechanized farm equipment	Solely owns mechanized farm equipment	Solely owns nonfarm business equipment	Solely owns uncultivated land
NFWP	0.057	-0.026	-0.027	-0.031	0.023	0.016	0.005	0.052
	(0.016)***	(0.010)**	(0.035)	(0.022)	(0.034)	(0.006)***	(0.018)	(0.015)***
NFWP x north	0.051	-0.027	-0.098	-0.007	0.076	0.002	0.031	0.024
	(0.023)**	(0.012)**	(0.052)*	(0.035)	(0.051)	(0.006)	(0.025)	(0.023)
Comparison Mean	0.14	0.03	0.27	0.30	0.30	0.01	0.09	0.07
N	4789	4789	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D12c. Single-Difference Impact Estimates with Heterogeneous Effects - Sole Ownership of Productive Assets

	Solely owns land	Solely owns large livestock	Solely owns small livestock	Solely owns poultry	Solely owns non-mechanized farm equipment	Solely owns mechanized farm equipment	Solely owns nonfarm business equipment	Solely owns uncultivated land
NFWP	0.057	-0.026	-0.027	-0.031	0.023	0.016	0.005	0.052
	(0.016)***	(0.010)**	(0.035)	(0.022)	(0.034)	(0.006)***	(0.018)	(0.015)***
NFWP x christianity	-0.071	0.027	0.033	0.029	-0.039	-0.015	-0.017	-0.044
	(0.021)***	(0.011)**	(0.046)	(0.031)	(0.040)	(0.007)**	(0.018)	(0.019)**
Comparison Mean	0.14	0.03	0.27	0.30	0.30	0.01	0.09	0.07
N	4789	4789	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D13a. Single-Difference Impact Estimates with Heterogeneous Effects - Sole Ownership of Consumption Assets

	Solely owns a house	Solely owns large durable goods	Solely owns small durable goods	Solely owns a mobile phone	Solely owns a smartphone	Solely owns a means of transport
NFWP	-0.001	-0.032	-0.067	0.050	0.004	-0.006
	(0.008)	(0.026)	(0.027)**	(0.037)	(0.011)	(0.005)

NFWP x age_30orless	0.013 (0.015)	-0.018 (0.031)	-0.044 (0.045)	0.001 (0.041)	0.006 (0.020)	0.005 (0.021)
Comparison Mean	0.07	0.18	0.44	0.55	0.07	0.04
N	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D13b. Single-Difference Impact Estimates with Heterogeneous Effects - Sole Ownership of Consumption Assets

	Solely owns a house	Solely owns large durable goods	Solely owns small durable goods	Solely owns a mobile phone	Solely owns a smartphone	Solely owns a means of transport
NFWP	-0.001 (0.008)	-0.032 (0.026)	-0.067 (0.027)**	0.050 (0.037)	0.004 (0.011)	-0.006 (0.005)
NFWP x north	0.029 (0.016)*	-0.060 (0.030)*	-0.057 (0.040)	0.062 (0.050)	0.011 (0.016)	-0.017 (0.024)
Comparison Mean	0.07	0.18	0.44	0.55	0.07	0.04
N	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D13c. Single-Difference Impact Estimates with Heterogeneous Effects - Sole Ownership of Consumption Assets

	Solely owns a house	Solely owns large durable goods	Solely owns small durable goods	Solely owns a mobile phone	Solely owns a smartphone	Solely owns a means of transport
NFWP	-0.001 (0.008)	-0.032 (0.026)	-0.067 (0.027)**	0.050 (0.037)	0.004 (0.011)	-0.006 (0.005)
NFWP x Christianity	-0.021 (0.017)	0.003 (0.031)	0.062 (0.042)	0.002 (0.044)	-0.010 (0.019)	0.017 (0.027)
Comparison Mean	0.07	0.18	0.44	0.55	0.07	0.04
N	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Income

Table D14a. DID Estimations With Heterogeneous Effects - LFP and Income Diversification

	Labor force participation	Self-employment	Number of IGAs
NFWP	0.049	0.049	-0.021
	(0.015)***	(0.019)**	(0.042)
NFWP x age_30orless	0.027	0.062	0.006
	(0.027)	(0.028)**	(0.058)
Comparison Mean	0.77	0.55	0.98
N	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Table D14b. DID Estimations With Heterogeneous Effects - LFP and Income Diversification

	Labor force participation	Self-employment	Number of IGAs
NFWP	0.049	0.049	-0.021
	(0.015)***	(0.019)**	(0.042)
NFWP x north	0.050	0.012	0.009
	(0.025)**	(0.027)	(0.072)
Comparison Mean	0.77	0.55	0.98
N	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Table D14c. DID Estimations With Heterogeneous Effects - LFP and Income Diversification

	Labor force participation	Self-employment	Number of IGAs
NFWP	0.049	0.049	-0.021
	(0.015)***	(0.019)**	(0.042)
NFWP x christianity	-0.067	-0.042	0.027
	(0.028)**	(0.025)	(0.079)
Comparison Mean	0.77	0.55	0.98
N	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated

Women's Empowerment

Table D15a. Single-Difference Impact Estimates with Heterogeneous Effects - Adapted A-WEIA Domain Indicators

	Productive asset ownership	Sole or joint owner of at least one large asset	Sole or joint owner of at least one asset	Sole owner of at least one large asset	Sole owner of at least one asset	Access to and decisions on credit	Control over use of income	Group member-ship
NFWP	-0.055	0.031	-0.005	-0.009	-0.012	0.015	-0.032	0.200
	(0.024)**	(0.058)	(0.026)	(0.031)	(0.027)	(0.015)	(0.022)	(0.047)***
NFWP x age_30orless	-0.025	-0.045	-0.034	-0.045	-0.041	0.008	0.008	0.031
	(0.038)	(0.039)	(0.021)	(0.030)	(0.030)	(0.010)	(0.023)	(0.048)
Comparison Mean	0.94	0.75	0.88	0.81	0.83	0.02	0.93	0.66
N	1911	4764	4763	4763	4763	4788	3793	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
ndicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D15b. Single-Difference Impact Estimates with Heterogeneous Effects - Adapted A-WEIA Domain Indicators

	Productive asset ownership	Sole or joint owner of at least one large asset	Sole or joint owner of at least one asset	Sole owner of at least one large asset	Sole owner of at least one asset	Access to and decisions on credit	Control over use of income	Group member-ship
NFWP	-0.055	0.031	-0.005	-0.009	-0.012	0.015	-0.032	0.200
	(0.024)**	(0.058)	(0.026)	(0.031)	(0.027)	(0.015)	(0.022)	(0.047)***
NFWP x north	-0.046	0.057	0.021	0.024	0.018	0.004	-0.028	0.123
	(0.035)	(0.080)	(0.041)	(0.047)	(0.041)	(0.020)	(0.034)	(0.066)*
Comparison Mean	0.94	0.75	0.88	0.81	0.83	0.02	0.93	0.66
N	1911	4764	4763	4763	4763	4788	3793	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Impact Evaluation of the Nigeria for Women Project: Baseline Report

Table D15c. Single-Difference Impact Estimates with Heterogeneous Effects - Adapted A-WEIA Domain Indicators

	Productive asset ownership	Sole or joint owner of at least one large asset	Sole or joint owner of at least one asset	Sole owner of at least one large asset	Sole owner of at least one asset	Access to and decisions on credit	Control over use of income	Group membership
NFWP	-0.055	0.031	-0.005	-0.009	-0.012	0.015	-0.032	0.200
	(0.024)**	(0.058)	(0.026)	(0.031)	(0.027)	(0.015)	(0.022)	(0.047)***
NFWP x christianity	0.038	-0.018	0.010	0.020	0.021	-0.006	0.024	-0.099
	(0.044)	(0.072)	(0.035)	(0.040)	(0.035)	(0.018)	(0.031)	(0.056)*
Comparison Mean	0.94	0.75	0.88	0.81	0.83	0.02	0.93	0.66
N	1911	4764	4763	4763	4763	4788	3793	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D16a. Single-Difference Impact Estimates with Heterogeneous Effects – Decision-Making Power—Expenditures

	Participates in making major household purchases	Feels their opinion is listened to at least sometimes for HH purchases	Can decide how to spend small quantities of money	Can spend own earned money without asking permission
NFWP	-0.002	-0.017	0.006	0.030
	(0.044)	(0.026)	(0.036)	(0.036)
NFWP x age_30orless	-0.003	0.007	-0.029	-0.011
	(0.041)	(0.019)	(0.029)	(0.024)
Comparison Mean	0.38	0.97	0.87	0.88
N	4774	1775	4783	4785

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D16b. Single-Difference Impact Estimates with Heterogeneous Effects – Decision-Making Power—Expenditures

	Participates in making major household purchases	Feels their opinion is listened to at least sometimes for HH purchases	Can decide how to spend small quantities of money	Can spend own earned money without asking permission
NFWP	-0.002	-0.017	0.006	0.030
	(0.044)	(0.026)	(0.036)	(0.036)
NFWP x north	0.043	-0.028	0.008	0.014
	(0.069)	(0.028)	(0.053)	(0.048)
Comparison Mean	0.38	0.97	0.87	0.88
N	4774	1775	4783	4785

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D16c. Single-Difference Impact Estimates with Heterogeneous Effects – Decision-Making Power—Expenditures

	Participates in making major household purchases	Feels their opinion is listened to at least sometimes for HH purchases	Can decide how to spend small quantities of money	Can spend own earned money without asking permission
NFWP	-0.002	-0.017	0.006	0.030
	(0.044)	(0.026)	(0.036)	(0.036)
NFWP x christianity	-0.009	0.030	-0.008	-0.018
	(0.070)	(0.027)	(0.051)	(0.047)
Comparison Mean	0.38	0.97	0.87	0.88
N	4774	1775	4783	4785

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D17a. Single-Difference Impact Estimates with Heterogeneous Effects – Decision-Making Power—Mobility

	Mobility decision-making power, proportion affirmative, out of 5	Solely decides to go to the market	Solely decides to visit family or relatives	Solely decides to visit friend or neighbor's house	Solely decides to attend a training	Solely decides to go outside of community or village
NFWP	0.048	0.069	0.043	0.038	0.032	0.058
	(0.053)	(0.053)	(0.055)	(0.058)	(0.052)	(0.054)
NFWP x age_30orless	-0.012	-0.017	-0.002	-0.017	0.004	-0.027
	(0.054)	(0.051)	(0.062)	(0.053)	(0.058)	(0.054)
Comparison Mean	0.41	0.41	0.40	0.44	0.41	0.39
N	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D17b. Single-Difference Impact Estimates with Heterogeneous Effects – Decision-Making Power—Mobility

	Mobility decision-making power, proportion affirmative, out of 5	Solely decides to go to the market	Solely decides to visit family or relatives	Solely decides to visit friend or neighbor's house	Solely decides to attend a training	Solely decides to go outside of community or village
NFWP	0.048	0.069	0.043	0.038	0.032	0.058
	(0.053)	(0.053)	(0.055)	(0.058)	(0.052)	(0.054)
NFWP x north	-0.017	0.002	-0.022	-0.024	-0.025	-0.015
	(0.103)	(0.101)	(0.100)	(0.109)	(0.102)	(0.107)
Comparison Mean	0.41	0.41	0.40	0.44	0.41	0.39
N	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * *p*<0.1 ** *p*<0.05; *** *p*<0.01
Estimations control for State fixed effects, women's demographics (age, education, and marital status)
household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)
indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity
Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Table D17c. Single-Difference Impact Estimates with Heterogeneous Effects – Decision-Making Power—Mobility

	Mobility decision-making power, proportion affirmative, out of 5	Solely decides to go to the market	Solely decides to visit family or relatives	Solely decides to visit friend or neighbor's house	Solely decides to attend a training	Solely decides to go outside of community or village
NFWP	0.048	0.069	0.043	0.038	0.032	0.058
	(0.053)	(0.053)	(0.055)	(0.058)	(0.052)	(0.054)
NFWP x christianity	-0.013	-0.036	-0.005	0.004	0.005	-0.033
	(0.095)	(0.094)	(0.095)	(0.094)	(0.093)	(0.101)
Comparison Mean	0.41	0.41	0.40	0.44	0.41	0.39
N	4789	4789	4789	4789	4789	4789

LGA-clustered standard errors in parentheses. * $p<0.1$ ** $p<0.05$; *** $p<0.01$

Estimations control for State fixed effects, women's demographics (age, education, and marital status)

household demographics (size, number of children under 5 years of age, number of adults who are 50 years old and above)

indicators for household religion and language, a recall asset index, and an indicator for whether the household has access to electricity

Weights applied to comparison observations correspond to p/(1-p) where p corresponds to the likelihood of being treated.

Appendix E. Power to Detect Effects

We revisited our power calculations using information from the baseline survey to assess the statistical power of this study.

Power Calculations to Estimate Gendered Effects of the NFWP at the Individual and Household Levels. Power calculations at the individual level suggest that interviewing 1,170 treatment women and 4,680 comparison women across 15 treatment LGAs and 58 comparison LGAs will be sufficient to detect small but meaningful effects of the NFWP. They would leave an 80% chance of detecting an intent-to-treat (ITT) effect of 0.20 standard deviations when we assume an intra-class correlation (ICC) of 0.08 for individuals and households clustered in LGAs and an R-squared of 0.25. This effect size is aligned with previous systematic reviews on the impact of self-help groups and vocational and business training on women's labor market outcomes (Brody et al., 2015; Chinen et al., 2017). The minimum detectable treatment effect of 0.22 standard deviations is equivalent to an impact of 10 percentage points on women's self-employment (which has a recalled baseline mean of 20%). For these power calculations, the ICC is based on a household asset index we constructed using recalled values of household asset ownership in December 2020 from the baseline survey.

We would be able to detect similar or smaller effect sizes for indicators related to household asset ownership for which the ICC is 0.086 based on the baseline survey. The ICC for women's ownership of small livestock and mobile phones is 0.16, suggesting that with the current sample size we would be able to detect a minimum treatment effect of 0.28 SD.

Power Calculations to Estimate Group-Level Effects. It is more challenging to determine appropriate ICCs for group-level outcomes. The LSMS includes survey questions about membership in savings groups and women's groups, but the questions and associated data are not sufficiently detailed for the estimation of ICCs. For the purpose of the power calculations, we assume an ICC of 0.10. Power calculations at the group level suggest that a sample size of 261 treatment women's groups and 1007 comparison women's groups across 15 treatment LGAs and 58 comparison LGAs will be sufficient to detect program impacts of 0.275 standard deviations with 80% power when we assume an ICC of 0.10 for groups clustered in LGAs. We would be able to detect similar or smaller effect sizes (0.24 SD) for variables such as cumulative savings for which the ICC is 0.078. The ICC for average monthly savings is 0.13, suggesting that with the current sample size we would be able to detect a minimum treatment effect of 0.28 SD.

We aimed to increase statistical power by oversampling women who are more likely to participate in the NFWP, such as women's group members. Specifically, we propose a sample in

which 4 out of 5 respondent households include a woman who is currently a member of a WAG or other women's group. We expect that by oversampling women for whom the take-up rate would be much higher and by involving community leaders in the program, take-up rates in our proposed sample would be no lower than 80%. Under such a scenario, power calculations at the individual level suggest that the minimum detectable program effect size is 0.24 standard deviations.

About the American Institutes for Research

Established in 1946, with headquarters in Arlington, Virginia, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance to solve some of the most urgent challenges in the U.S. and around the world. We advance evidence in the areas of education, health, the workforce, human services, and international development to create a better, more equitable world. The AIR family of organizations now includes IMPAQ, Maher & Maher, and Kimetrica. For more information, visit AIR.ORG.

AIR® Headquarters
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | **AIR.ORG**