Impact Evaluation of Cash-Based Transfers and Gender in El Salvador

Inception report

November 2021

Key personnel for the evaluation

OFFICE OF EVALUATION

Felipe Dunsch, Evaluation Officer

Jonas Heirman, Evaluation Officer

Andrea Cook, Director of Evaluation

IMPACT EVALUATION TEAM

Erin Kelley, Team Leader

Lelys Dinarte, Team Leader

Sylvia Maria Cesar, Field Coordinator

Miguel Paniagua Melendez, Field Coordinator

Contents

1. Introduction

Gender inequality in economic autonomy is pervasive, particularly in developing countries, and its potential welfare implications are concerning (Jayachandran, 2015). In El Salvador, only 50 percent of women participate in the labour market, in contrast to 80 percent of men (World Bank, 2020). Economic development, gender equality in labor market opportunities, and gender equality in autonomy are all strongly linked, but causality is still unclear. The World Food Programme's (WFP) Office of Evaluation (OEV), Cash-based Transfers (CBT) Division, and Gender Office, partnered with the World Bank's Development Impact Evaluation (DIME) department to create the "Cash-Based Transfers (CBT) and Gender" Impact Evaluation (IE) Window. The CBT and Gender Window aims to understand the impact of CBT interventions targeting women on gender equality and women's empowerment, as well as food and nutrition outcomes.

The first round of impact evaluations selected for this window aim to assess the impacts of increasing women's participation in work outside the household, as a conditionality of cash-based transfers, and directly receiving a wage (the cash-based transfers) on their social and economic empowerment.

WFP El Salvador's capacity in policy and programming is widely recognized. For example, WFP El Salvador designed an innovative model for ensuring vulnerable populations have access to regionally produced nutritious foods using commodity vouchers redeemable at local shops, thus making these foods available to everyone. As women and girls are particularly vulnerable to violence and food insecurity, WFP prioritizes their protection in all its activities according to its regional gender strategy and the El Salvador CO's gender action plan. WFP El Salvador is constantly looking for ways to innovate and learn and is keen to use evidence from this evaluation to inform broader programming priorities, which can also contribute to the upcoming Country Strategic Plan.

The El Salvador impact evaluation aims to understand the impacts of food-assistance for assets (FFA) programming targeting women on both income and asset loss, while also aiming to understand the longer-term implications of the response on gender equality and women's social and economic empowerment. Main direct outcomes of the intervention are increasing women's earnings and supporting her to alter time-use. The theory of change conjects that these then (in the medium term) impact perceptions of gender norms, attitudes, agency, consumption patterns, and well-being (physical, social, and psychological).

2. Evaluation Context

2.1. BACKGROUND AND CONTEXT

El Salvador is Central America's smallest and most densely populated country. Of a population of 6.4 million people, 62 percent live in urban areas and there is a large diaspora. Approximately 1.4 million Salvadorians live in the United States of America.[1] In recent years, El Salvador has made significant progress in reducing food insecurity, chronic malnutrition, poverty, and inequality. However, food insecurity and malnutrition, along with recurring natural disasters, persistent gender inequality, slow economic growth, high public debt, and a rampant homicide rate, are persistent challenges. In 2020, El Salvador's Human Development Index (HDI) ranking is 124th out of 189 (behind Belize (110th), but slightly ahead of Guatemala (127th), Nicaragua (128th), and Honduras (132nd)).

The most recent Integrated Food Security Phase Classification (IPC) analysis estimates that by May 2021, the population classified as experiencing food insecurity "crisis" (IPC Phase 3) and above is expected to increase to 985,000 people – a 44 percent increase from the number estimated at the turn of the year. Of these, 95,000 people are categorized as being in "emergency" conditions. The worst affected populations are those who depend on agricultural and livestock activities, work in the informal sector, or own small businesses. These groups have experienced income losses due to the COVID-19 pandemic, and have had to use more extreme coping strategies to address food gaps, such as the sale of productive assets.

While women and men in El Salvador are equal before the law regardless of their marital status (Grameen Foundation, 2019), there remain many barriers that prevent women from accessing the same opportunities as men. Intrafamilial, social, sexual, and institutional violence against women in the country is among the highest in the world (World Bank, 2020). Moreover, only half of all Salvadorian women participate in the labour market, in contrast with 80 percent of men and 57 percent of women in the rest of Central America (World Bank, 2020). A high percentage of women are considered as economically inactive due to housework, showing a 68 percent gap with respect to men (WFP, 2017b). The gender gap in incomes is 16 percent, reaching nearly 20 percent in rural areas. To achieve sustainability development goal (SDG) 2, a strategic review explicitly suggests that the country should address gender inequalities, which may impede poverty reduction, food security and nutrition (WFP, 2017a).

A Window-level pre-analysis plan outlines the literature regarding women's labor and earned income, and its potential impacts on gender equality and women's empowerment outcomes. However, a few key pieces of literature are particularly relevant to El Salvador. For example, in their study of the impact of a conditional cash transfers on maternal health in El Salvador, de Brauw and Peterman (2020) assert that positive gains in maternal health outcomes are likely partially driven by increases in women's decision-making agency. However, the study was unable to measure this directly, or isolate the empowerment mechanisms from the "income effect", as this evaluation will.

Additionally, a recent qualitative study conducted by the WFP Gender Office on El Salvador's cash-based FFA programming, documented increases in self-esteem and self-worth among both young women and young men beneficiaries, while increasing their food security and nutrition. It also found improvements to gender equality outcomes via a reduction in violence and negative coping mechanisms, and increases in autonomy and equitable resource control (WFP 2019). This evaluation aims to experimentally test the mechanism by which these outcomes occurred, and validate (or challenge) the qualitative study findings with large-scale household-level data.

2.2. PROGRAMME DESCRIPTION

It is in this context, DIME and WFP are collaborating to understand the impact of WFP programming on women's earnings, time use, consumption, agency, attitudes, perception of norms, and well-being. The WFP

---

[1] https://www.migrationpolicy.org/programs/data-hub/charts/immigrant-and-emigrant-populations-country-origin-and-destination?width=1000&height=850&iframe=true

contry office (CO) in El-Salvador joined the CBT & Gender IE Window in August 2019 – motivated by a desire to understand the impact of their FFA programme on gender equality and women's social and economic empowerment. The CO is implementing the FFA intervention as part of Outcome 4 (Activity 9) of its Country Strategic Plan (CSP). One of the CSP's expressed goals is "[through] food assistance for assets, WFP will promote asset creation activities to stimulate early recovery, rebuild livelihoods and reduce long-term vulnerability to food insecurity and malnutrition." The assets that women will develop or contribute to are poultry rearing, cleaning riverbeds/irrigation ditches, flood prevention activities, attending to vegetable gardens, reforestation, road repair, fumigation/pest-control, and communal infrastructure upgrades

To test measurement and randomization strategies before a full-scale evaluation of the FFA programme, the Impact Evaluation Team supported the El Salvador CO to pilot the intervention in November 2019. The pilot phase included joint work with local government and community leadership, baseline data collection, and training sessions with beneficiary communities. The pilot field activities have been carried out in five communities (30–40 households within each community) from the municipality of San Lorenzo, in the Department of San Vicente.

After the successful pilot in 2019/2020, the scale-up FFA intervention in El Salvador in 2021 aims to work with 1,500 households in 75 communities across the country, which present high indicators of food insecurity (as a result of climactic or economic shocks), and will form the basis for the IE (timeline provided in Annex 1 and a more detailed analysis of the evaluation design is provided in Section 8). In the framework of WFP's FFA programming, the goal of the IE is to test whether equalizing opportunities for men and women to work outside the home contributes to closing the gender gap in autonomy and ultimately improves social and economic empowerment.

The map below shows both the pilot activity locations (light blue) and the communities participating in the scale-up (dark blue).
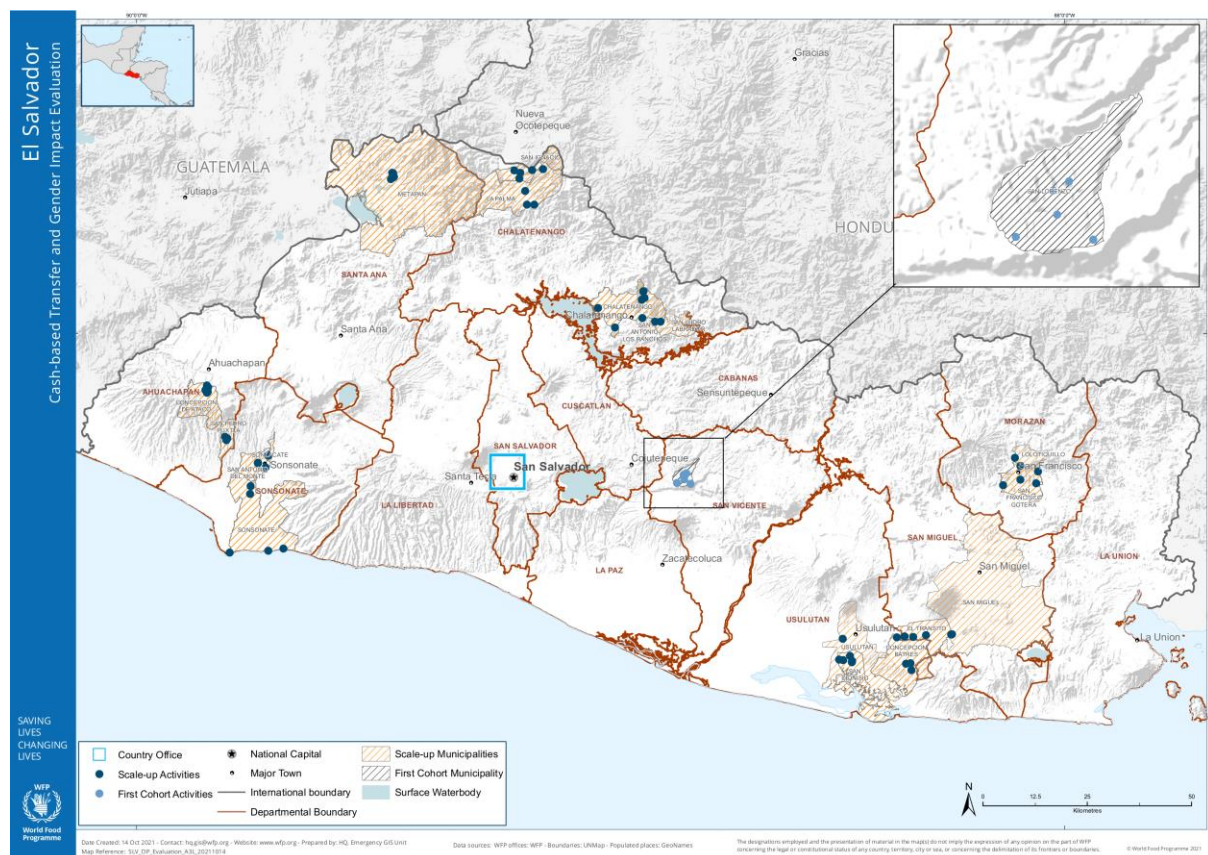
3. Evaluation Approach and Questions

3.1 APPROACH

This impact evaluation will employ a clustered randomized controlled trial (RCT) design, in which 75 communities across the country are randomly assigned into one of three treatment groups containing 25 communities. In each group, 500 households will participate, for a total sample of 1,500.

The evaluation will involve three rounds of data collection in 2021, allowing the team to separately estimate short-term and medium-term impacts (timeline presented in Annex 1). Baseline data collection will take place before programme implementation begins. The FFA programme is expected to last 3 months, with midline data collection taking place 1.5 months into programme implementation (between the 2nd and 3rd transfer). Endline data collection will occur 3 months after final intervention activities.

Evaluation results will feed into the design of upcoming FFA programming in El Salvador and can inform the next Country Strategic Plan (CSP) (2022–2026) which focuses on strengthening institutions and filling gaps in the coverage of government food security and nutrition programmes, including support to drought response. The strategy reaffirms WFP's commitment to facilitating vulnerable households' access to effective, productive, and nutrition-sensitive social protection; and targeting populations and communities in the most food-insecure areas. The strategy also emphasizes WFP's commitment to prioritizing the protection of women in all its activities according to its regional gender strategy and the CO's gender action plan. Special attention will be given to the gender gap in incomes and women's protection needs, especially in urban areas with high levels of violence.

3.2. HYPOTHESES

The impact evaluation is designed to test the hypothesis that cash-based transfer programming targeting women increases gender equality and women's economic empowerment, by increasing women's time spent in paid labour outside the household and thus increasing their earned income.

The first hypothesis is that involving women in work (asset creation through the FFA programme) directly impacts their time use (shifts towards paid work outside the home), as well as their earnings as they are paid directly for their work.

The second (following) hypothesis is that – in the medium run – these combined shifts in time use and earnings will impact women's:

- Perceptions of gender norms
- Attitudes
- Agency
- Consumption patterns
- Well-being (physical, social, and psychological)

Thus, in the longer run, we hypothesize including women in work outside the home can initiate a "virtuous cycle" where a change in women's perceptions of norms, attitudes, and agency further boosts women's participation in paid work outside the home (time use). This then positively impacts their earnings, which could amplify (control over) consumption and well-being, even after the FFA intervention ends. While the programme is targeted at women, it is possible the programming will also impact men's perceptions of gender norms (and those of the wider community) and attitudes in a way that further contributes to improvements in gender equality.

This theory of change is consistent with a body of literature that examines the impacts of providing women opportunities to work outside the household, as summarized by Field et al. (2019): "Female employment

has been shown to delay marriage, increase female work aspirations, improve child health, and reduce the male:female sex ratio (...). In the United States, rapid growth in female labor force participation preceded important changes in norms regarding gender roles in both the economy and the household" (p. 1). Recent experimental work has demonstrated attitudes (Dhar et al., 2018; McKelway, 2019) and norms (Beaman et al., 2009; Bursztyn et al., 2018) shape women's agency and, in turn, women's labor supply. While FFA programs have demonstrated to be an effective tool for economic development through increased earnings (Imbert & Papp, 2015; Gazeaud et al., 2019; Adjognon et al., 2020), there is less evidence on the impacts of participant gender and there is also a lack of evidence on projects with a short duration.

More details on the theory of change are presented in annex 1.

3.3. PRIMARY EVALUATION QUESTIONS

The main objective of the IE is to answer the following evaluation questions:

1. What is the impact of women's participation in a FFA programme (working outside the household and receiving cash in return) on their social and economic empowerment?[2]

2. What is the impact of an unconditional cash transfer to the household on women's social and economic empowerment, as well as on household income and welfare?

Each question will be evaluated using the same outcome indicators (explored further in Section 7):

- Time use
- Earnings
- Perception of norms
- Attitudes
- Agency
- Consumption patterns
- Social, physical, and psychological well-being

These evaluation questions are derived directly from the theory of change, and are intended to isolate the impact of increasing *women's* income and time spent working outside the household on gender equality and women's empowerment, while controlling for the "income effect" of the cash transfer generally (comparing to the second treatment arm), and understanding the overall impact of the WFP programming (comparing to the control).

3.4. SECONDARY EVALUATION QUESTIONS

As secondary questions, the evaluation also asks:

1. Does FFA affect the probability and/or reasons for respondents' migration? Migration behaviors will be measured by the following survey questions:
    - *Have you ever moved to another department or municipality within El Salvador?*
    - *What was the main reason for your move?*

2. Are there heterogenous impacts of the transfer based on respondents' exposure to community violence? Exposure to violence will be measured by the following four survey questions:
    - *During the past 30 days, were there any days you did not leave your home because you felt it would be unsafe for any reason?*
    - *Did you ever not leave your home because of fear of violence in the community (threats, extorsions)?*
    - *Did you ever not leave your home because of fear of the police or army in the streets?*
    - *Outside your home, how many times did you see anyone get attacked?*

3. Does the FFA programme affect key food security outcomes of interest?

---

[2] This can also include negative unintended effects.

The evaluation will compare the two intervention arms across key food security measures commonly used in WFP, including the Food Consumption Score (FCS) and the Livelihoods Coping Strategies Index (LCSI). The analysis will give an estimate of the overall impact of WFP programming on food security.

3.5. OUTCOMES OF INTEREST

From the seven outcome categories, there are 52 outcome indicators of interest. These outcomes were developed in close collaboration with the El Salvador CO to ensure operationally relevant indicators are captured. The outcomes will be collected across all six country IEs in the Window and will support cross-country analysis. The outcomes are selected based on a review of relevant literature and previous studies that aimed to capture similar outcomes.

Inherent in the design of the evaluation is the measurement of progress on gender equality. As both men and women are asked questions on time use, agency, attitudes, perceptions, and well-being separately, the evaluation will be able to identify whether (and how much) inequalities still exist in these areas, and whether the programme contributed to decreasing the gender equality gap.

Table 1: Main Outcomes of Interest

| Outcome type	Outcome name	Definition	Measurement level
Primary	Consumption	Expenditures over reference period on 10 goods	Household
Primary	Earnings	Total earnings from WFP plus total earnings from other paid permanent and temporary work	Household
Primary	Time Use	List of activities from 24-hour recall over past two days; asked separately of men and women	Individual
Primary	Agency	How much the woman's opinion would be considered in a series of decisions	Individual
Primary	Attitudes	The woman's belief of how much time she should spend on productive activities, relative to men	Individual
Primary	Perceptions of Norms	The woman's perception of the time use, agency, and attitudes of women in her community	Individual

Primary	Well-Being	Psychosocial well-being, life satisfaction, mental health, and intimate partner violence	Individual

4. Evaluation Design and Sampling Strategy

To identify the causal impacts of the treatment arms, the IE will employ a clustered randomized control trial (RCT) design. The clustered RCT approach follows from the programme's implementation modality of intervening at the community level, which would not have allowed for a household-level randomization. To start, DIME and the WFP CO selected 13 municipalities.[3] Within each municipality, 5–6 communities were selected for inclusion in the study using the following criteria:

- They do not expect a WFP transfer this year.
- They rank Priority 1 or 2 in the CO's strategy. The priority ranking is devised based on food insecurity, poverty, and vulnerability due to COVID-19 and tropical storms, with rank 1 being the most food insecure.
- There is an even distribution of males and females within the communities.

In a second step, the 75 communities will be randomly assigned into either one of the two treatment groups or the control group (see Figure 2), producing a clustered randomized design.

Figure 2: WFP El Salvador IE Design
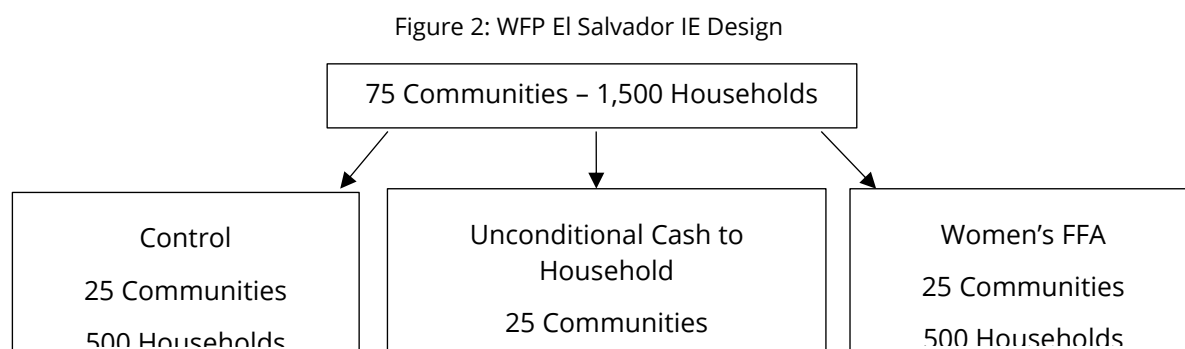
In each community, WFP will work with local community leaders and government officials to identify 20 of the most vulnerable households within each community for a total sample size of 20*75 = 1,500 households (see the next section on power and sample size calculations). A feature of the clustered randomized controlled trial design is that all selected beneficiary households within a community will receive the same treatment to avoid any "spillover" concerns that might arise from a within-community household randomization approach. The household identification process in all 75 communities will be the same regardless of "treatment" assignment to avoid any biases.

4.1. TREATMENT AND CONTROL GROUPS

Details of the two treatment arms and control group are as follows:

Treatment 1: Beneficiaries in this treatment group receive a conditional cash transfer ($300) disbursed over three months, provided they work on an asset – where the primary female decision maker is registered to receive the transfer and work on the asset. The country office estimates that $100 a month are the amount required to fill any existing gaps for food expenditure.

Treatment 2: Beneficiaries in this treatment group receive an unconditional cash transfer ($300) disbursed over three months – where the primary male decision maker is registered to receive an unconditional transfer. Please note that the unconditional cash transfer was provided to men because they work during the

---

[3] These include SAN ANTONIO DEL MONTE, SAN DIONISIO, CONCEPCION BATRES, CONCEPCION DE ATACO, SAN PEDRO PUXTLA, SONZACATE, SONSONATE, METAPAN, SAN IGNACIO, SAN FRANCISCO GOTERA, LA PALMA, EL TRANSITO, CHALATENANGO

day and therefore lower take-up (differential attrition) for any asset related activities was expected (which is a slight deviation from the "window" design where the first treatment arm is a "business as usual" FFA intervention).

<u>Control Group</u>: Beneficiaries in the control group receive a $300 lump sum unconditional cash transfer after the endline surveys are completed.

4.2. SAMPLE SIZE CALCULATIONS

The country office's budget and implementation capacities allow for the impact evaluation to be conducted in 75 communities (with ca. 20 households in each community). For the first power calculations based on these parameters, we use women's preferred consumption as an outcome, as it can be calculated in any household survey. For the second power calculation, we use predicted household consumption (see Appendix 2 for details). We use the El Salvador Income and Expenditure Survey (ENIGH) for these calculations, restricting to rural poor households, consistent with the typical households targeted by WFP FFA programmes. We apply the standard formula for the minimum detectable effect (MDE):

$$MDE = \sigma_e (z_{0.8} + z_{0.975}) + \sqrt{1 + \frac{\rho(m-1)}{NP(1-P)}}$$

Where σ_e is the standard deviation of the outcome, $z_{0.8} + z_{0.975} = 2.80$ is the sum of the two z-scores, ρ is the intra-cluster correlation, m is the number of observations per cluster, N is the number of observations, and P is the share of observations assigned to treatment. We set $\rho = 0.05$ for all calculations.

To calculate our expected effect size for each analysis, we focus on effects during the midline survey. For household consumption as an outcome of pooled treatment, we first apply a marginal propensity to consume from cash transfers of 0.67, estimated based on Haushofer and Shapiro (2016). We then multiply this by the share of households anticipated to take up the intervention, and the monthly transfer size relative to average monthly household consumption. For women's income as an outcome of Women's FFA conditional on being treated, we continue to apply a marginal propensity to consume of 0.67. We then multiply this by take-up, which is now the share of participating households who shift from male to female participants in response to Women's FFA, and the monthly transfer size relative to average monthly household consumption.

We find the following minimum detectable effects (MDE) for El Salvador, which are reasonable (as determined in the literature):

Number of Observations	1,500 Households (20 per community)
Number of Clusters	75 Communities
Transfer Size	300 USD
MDE for Consumption	0.3 SD
MDE for Women's Income	0.25 SD

More details are presented in Annex 2.

5. Data Collection

The timeline of surveys and implementation is presented in Annex 1. All data will be collected using computer assisted personal interviewing (CAPI) techniques, utilizing Android tablets running SurveyCTO data collection software. Surveys are approximately two hours in length.

Baseline surveys will take place just prior to the start of the intervention (February 2021). The WFP programme studied in El Salvador lasts three months. A midline survey will take place during the implementation of cash transfers, and its reference period will lie entirely during the three-month period during which cash transfers are being made. This is necessary so that all questions during the midline, particularly time use and income, can be used to estimate the direct impacts of FFA and Women's FFA. An endline survey will occur just after the end of the intervention – sufficiently after it so that the reference period for the endline survey will exclude the period of the intervention. This is necessary such that all questions during the endline can be used to estimate the persistent indirect impacts of FFA and Women's FFA.

By virtue of IE design, data collected will be disaggregated by gender of the respondent. Importantly, the IE does not consider a "household" to be one unit, but rather individuals are within a household. As such, the survey is repeated to both male and female respondents in the same household for all of the key outcomes described in Section 7. One exception is the module measuring IPV – for ethics and protection, this module is only asked of women responding without the male respondents present or aware of the module.

While the survey is relatively standard across all IEs in the CBT and Gender Impact Evaluation Window, it will be piloted prior to data collection with local communities in El Salvador to ensure questions are relevant to the context. In addition, the consumption module is specifically tailored to context (described below), and the power calculations for the IE use data from the El Salvador Income and Expenditure Survey (ENIGH) as already outlined in Section 8.2.

Finally, as an attempt to capture the range of ways that agency, attitudes, and norms can manifest in everyday decision making, each of these outcomes is measured using multiple questions along three separate productive assets. Additionally, four separate activities are used to understand women's decision-making power and perception of norms. These variations on the key outcomes are described below.

5.1. QUANTITATIVE INSTRUMENTS

The survey includes seven main outcome categories, measured as follows:

Consumption: Expenditures over a standard reference period for up to ten goods are asked. Five goods are selected as the goods that most strongly predict household consumption in a household survey from the same context. Five goods are selected as the goods that most strongly predict women's income, controlling for total household consumption, in a household survey from the same context. Expenditures on education, men's clothing, and women's clothing will be included.

Earnings: Earnings for each household member are collected for the previous six months for the baseline survey, the time since the baseline survey for the midline survey, and the minimum of the previous six months, or two weeks after the intervention was completed, for the endline survey. Earnings are measured as total earnings from WFP plus total earnings from other paid permanent and temporary work.

Time Use: The female respondent is asked for a 24-hour recall of her activities over the past two days, following the approach of American Time Use Survey. When the primary male decision maker in the household is available, he is asked about his activities over the past two days; when he is not, the female respondent is asked about his activities.

Agency: The female respondent is asked, relative to the primary male decision maker in the household, how much her opinion would be considered in a series of decisions. These questions follow the DHS on consumption ("major household purchases", "purchases from the primary male decision maker's income", "purchases from the female respondent's income", "the female respondent's health care"), and include additional questions on decision making over both men's and women's time in three productive activities ("work in self-employment", "work for a salary", "work on household chores").

<u>Attitudes</u>: The female respondent is asked how much time she should spend, relative to the primary male decision maker in the household, on the three productive activities listed above.

<u>Perceptions of norms:</u> The female respondent is asked how much time she believes women, relative to men, in her community spend on three productive activities. Next, the female respondent is asked how much the opinion of women in her community would be considered, relative to primary male decision makers in their households, on the same set of decisions as the Agency questions. Finally, the female respondent is asked about the attitudes of people in her community. These questions mirror the above questions on attitudes towards time use and attitudes towards agency.

<u>Well-being</u>: Modules to measure locus of control, psychosocial well-being, life satisfaction, IPV, and depression (PHQ9) will be administered. Two modules will be used to assess any unintended consequences of the intervention on women. First, the time use module will reveal whether the programme has contributed to a "second shift" for women; as they pick up more work outside the home, this may not be accompanied by reduced domestic labour burdens. The share of time spent on domestic and carework duties between men and women is an important indicator of gender equality in the analysis. Second, questions from the Demographic and Health Survey (DHS) module on IPV are included, with adaptation of questions based on Haushofer et al. (2019). This will allow the IE to identify any unintended consequences of a direct transfer of cash to women on the intra-household dynamics and her experience of IPV.

These outcomes are measured during the baseline data collection, the midline (approximately 1.5 months after the start of the programme), and at endline (after the completion of the three-month programme cycle).

5.2. MANAGEMENT OF DATA QUALITY

The team is taking multiple steps to ensure we collect high-quality data. This begins by hiring a set of 30–40 experienced enumerators. These enumerators have worked with WFP in the past and are hired on short-term contracts for the number of days in question. The hiring process takes place through a third party, a recruitment and human resources management company with whom WFP contracts with. The team then trains the enumerators in best practices, checks incoming data, and communicates any data issues regularly to the enumerators. We describe each one in turn:

<u>Enumerator Training</u>

The training is divided into four stages and will take approximately one week to complete:

Review the survey's content: the team will guide enumerators through each section of the survey, eliciting their feedback about the content and answering any questions they may have about how to administer the questions to respondents. This process ensures that any ambiguities about the questionnaire are resolved ahead of time.

Mock surveys: once the survey has been reviewed, the team will ask the enumerators to pair up and conduct "mock surveys" where they administer the questions to each other. This session is followed by a question and answer period to review any additional concerns or questions, and to provide feedback on individual enumerators' performance.

Review best practices: once the mock surveys are complete, the team comes together to discuss best practices for engaging with respondents and recording their answers into the software. This includes a review of:

- How to record survey responses.

- How to provide alternative phrasing so respondents understand the question.

- How to ensure smooth transition in telephone surveys, especially when the survey will be broken up into several telephone calls.

<u>Ensuring Beneficiary and Enumerator Protection</u>

The survey asks about sensitive topics, including IPV and mental health, that could be distressing for respondents and elicit responses that enumerators may find emotionally difficult to discuss. To address these concerns, the study will follow WFP's guidelines on collecting sensitive data for IEs, and seek support from the gender and/or protection officer to establish the proper protection infrastructure. This includes

mapping referral pathways for communities being surveyed and hiring the services of an NGO specialized in psychosocial support to provide enumerator training. This training will instruct enumerators on how to conduct the more sensitive questionnaire modules, and on when and how to use referral pathways if a beneficiary reports an incident of violence.

<u>Data Quality Protocols</u>

The CAPI survey will ensure the number of logical inconsistencies in the data is reduced to a minimum. Additionally, the team will carry out High Frequency Checks (HFCs) during the entire data collection period. HFCs are a data quality assurance process meant to detect any anomalies in the data we collect. They are run daily so the team can make any necessary adjustments to data collection processes in the field. HFCs look out for the following instances:

- Too many missing observations

- Duplicate observations

- Unusual survey duration (too short or too long)

- Too many respondents stating "no consent"

- Inconsistent patterns in the data

Any anomalies that we detect through this process will be flagged to the data collection team immediately. In addition, the team will also perform a set of back-checks. This refers to drawing a random 10–20 percent sample of households and calling them back to validate some of their answers.

<u>Communication Strategies</u>

The team has developed an innovative data tracking dashboard. Specifically, the team developed code that downloads the raw data from the server and computes the various statistics we use for the HFC checks, as well as completion status of all surveys. This information is then stored in a Google Sheet for different team members to consult. In particular, enumerators can log on to check how many surveys they have completed, and which surveys are still pending. This ensures the team is actively tracking survey progression and data quality.

5.3. QUALITATIVE INSTRUMENTS

Given limited data collection budgets, the team chose to focus data collection on household surveys that capture outcomes at the household and individual level. The data collected is both quantitative and qualitative, with significant opportunity for respondents to elaborate on responses through text fields and for enumerators to record "other" responses. An additional barrier to focus group discussions was the requirement for the IRB to limit "research activities" that increase the risk COVID-19 group-based spread.

In addition to measuring the impact of WFP's programme in El Salvador, the impact evaluation will collect limited qualitative data to examine important process related questions:

1. How did the process of programme implementation contribute to, or hinder, the achievement of measured outcomes? To what extent were programme interventions implemented as planned?

2. How did intended beneficiaries supported by the programme experience participation in selected interventions? And, how do they perceive the positive or negative concequences of any measured outcomes?

If COVID continues to prevent in-person interviews or focus groups, qualiative information will be collected remotely, through phone-based interviews or qualitative surveys.

5.4. IE IMPLEMENTATION MONITORING SYSTEM

WFP and DIME are working together to ensure beneficiaries receive the scheduled WFP programming on time. WFP regularly tracks when transfers are made to programme recipients, and also tracks whether work requirements are met. DIME is complementing these efforts by ensuring that the programme variations we

introduce are properly followed. In other words, DIME is monitoring treatment compliance in the following way:

- Treatment 1: The household's primary male decision maker is registered as the primary beneficiary in WFP's SCOPE database. He will receive cash transfers in a timely fashion.

- Treatment 2: The household's primary female decision maker will be registered as the primary beneficiary in WFP's SCOPE. She will receive cash transfers in a timely fashion. She will also be invited to work on a community asset and will be asked attend any necessary meetings or trainings for this work. Attendance at all meetings will be recorded and digitized.

- Control Group: Households should not receive cash transfers until after endline, nor should they be assigned an asset to work on. They should not attend asset trainings or meetings.

The IE field coordinator will routinely run a code that will flag any discrepancies with the aforementioned treatment compliance indicators. If any discrepancies are flagged, the field coordinator will notify WFP and/or the cooperating partner responsible for implementing field activities

6.   Data Processing and Analysis

6.1. DATA CODING, ENTRY, AND EDITING

All data we use will be collected via tablets. The data will be stored on SurveyCTO servers. As soon as a surveyor marks a filled-out form as "finalized", the form's contents are encrypted. Whenever form data is transmitted via 3G or other Internet network, it is encrypted in transit using SSL as well. Finally, any data that is downloaded from the server will either be encrypted or be purged of any personal identifiers before analysis. A series of back-checks will be performed on the data we collect. Any mistakes that are detected will be recorded and changed. This will avoid missing data systematically across treatments (if there is missing data it will be random across treatments, and therefore does not impact the analysis).

6.2. PROGRAMME-SPECIFIC QUANTITATIVE DATA ANALYSIS

The main objective of the analysis, as per the Window's design, is to estimate the impacts of women's participation in the programme on the main outcomes of interest (Section 7), adjusting for any household-level impacts of increasing income. To do this, we use the fact that the unconditional cash arm shifts income at the household level, while Women's FFA shifts both women's participation in the programme and income at the household level. The model for these estimates is provided in detail in Annex 3 and summarized here. Standard errors will be clustered at the community level, in accordance with the clustered randomization design.

We estimate the following instrumental variables model in each survey wave t. Letting Y_{ht} be outcome Y for household h in survey wave t (0 for baseline, 1 for midline, and 2 for endline), we estimate:

$$Y_{ht} = \beta_{1t} WomanParticipant_{hl} + \beta_{2t} AnyParticipant_{hl} + X_h' \gamma_t^Y + \varepsilon_{ht}^Y (1)$$

$$WomanParticipant_{hl} = n_{1t}^T UnconditionalCash_h + n_{2t}^T CashforWomen'swork_{hc} + X_{h0}' \gamma_t^T + \varepsilon_{ht}^T$$

$$Anyparticipant_{hl} = n_{1t}^I UnconditionalCash_h + n_{2t}^I CashforWomen'swork_h + X_{h0}' \gamma_t^I + \varepsilon_{ht}^I$$

where X_h is a vector of controls which includes the value of the outcome of interest at baseline and any stratifying variables used for randomization (in El Salvador the stratifying variables includes the municipality). The primary coefficient of interest is β_{it} – the estimated impact of shifting all of a household's income from men to women.

For inference, we will control false discovery rate across outcomes, using randomization inference following Anderson (2008).

We will test balance along all outcomes of interest and key demographic variables, including household size.

By virtue of the IE design, outcomes will be analysed by gender to detect inequalities between household members. As mentioned in Section 8, a feature of the clustered randomized controlled trial design is that all selected beneficiary households within a community will receive the same treatment to avoid any "spillover"

concerns that might arise from a within-community household randomization approach. As non-beneficiary households within treated communities are not surveyed, the evaluation will not be able to detect any positive (or negative) spillovers within communities.

For each regression, we will test for differential attrition and, for questions where men respond when present, differential attrition of male respondents. When statistically significant attrition is present for a given outcome, we will estimate Lee Bounds for that outcome and report the average upper bound and average lower bound in robustness.

6.3. PROGRAMME-SPECIFIC QUALITATIVE DATA ANALYSIS

In part due to limitations on gathering and face-to-face interactions due to COVID-19, no qualitative data will be collected on the process of implementation or the experience of programme participants, other than space given within the tool for specifying "other" responses if those listed to not adequately describe the respondent's answer. If a certain "other" response occurs with significant frequency across surveys, this will be coded and included in the analysis.

We will however, be collecting qualitative information relating the implementation process as described in Section 5.3. We will be asking the beneficiaries if in their perception the programme has had a positive or negative impact on outcomes.

7. Ethics and Risks

7.1. ETHICAL CONSIDERATIONS

Evaluations must conform to 2020 (United Nations Evaluation Group (UNEG) ethical guidelines. Accordingly, OEV and DIME are responsible for safeguarding and ensuring ethics at all stages of the evaluation cycle. This includes, but is not limited to, ensuring informed consent, protecting privacy, confidentiality and anonymity of participants, ensuring cultural sensitivity, respecting the autonomy of participants, ensuring fair recruitment of participants (including women and socially excluded groups) and ensuring that the evaluation results in no harm to participants or their communities. During the inception phase, the following ethical issues, related risks, safeguards and measures have been considered:

IRB

The impact evaluation window design, as well as the specifics of the El Salvador study, received ethical approval on 10/03/2020 by Solutions IRB, which is a private commercial AAHRPP fully accredited Institutional Review Board (IRB).

Programme Exclusion

Every IE participant will be a WFP beneficiary, selected through a rigorous process that consults the community to identify the most vulnerable households. All three treatment groups are eligible for cash payments – the control group will just receive this transfer after the IE has concluded.

Informed Consent

Every household enrolled in this IE must consent first to being part of WFP's programme as per WFP guidelines, and then provide informed consent to be surveyed. Refusal to respond to our survey does not preclude participation in the WFP programming. Informed consent will be collected for each survey round separately (baseline, midline, and endline).

Privacy During Interviews

A woman selected as eligible to participate in the FFA programme is the primary respondent for the survey. While most survey questions are addressed to the woman, there are a few questions directed to the primary male decision maker – who is also eligible to participate in the FFA programme. Therefore, the man who is also eligible to participate (referred to as "primary male decision maker") will be surveyed on a reduced set of questions.

Despite the minimal risks, we will take a number of precautions to ensure questions addressed to respondents respect their privacy and comfort. First, interviews will be done at a central point in the village so respondents feel comfortable answering questions about their agency, time use, etc. Second, interviews will be conducted outside of earshot of other participants (including those from the same household) and enumerators. Following a first section of the survey when both female and male respondents might be present, enumerators will request for others to step away as they interview the female or male respondent, with the goal of providing a safe and quiet environment for the survey. Third, in contexts where particularly necessary, enumerators will be female, to ensure the highest degree of comfort for survey respondents. Fourth, we will coordinate with WFP and community leaders to help care for the respondents' children (as necessary) to ensure maximum privacy during the survey. Fourth, all enumerators will go through a training that will last for 1–2 weeks and will be followed by extensive piloting in the field. The goal of the training is to ensure enumerators follow survey best practices in terms of protocols and ethics, but also that questions are asked in a uniform and contextually appropriate manner. Fifth, for the most sensitive questions related to intimate partner violence (IPV), third-party experts will be contracted to train enumerators on how to ask these questions, and handle/refer cases of IPV to the relevant authorities (see below).

These issues will be monitored and managed during the implementation of the evaluation. If any additional ethical issues arise during the implementation of the evaluation, they will be recorded and managed in consultation with OEV and DIME.

7.2. RISKS AND LLIMITATIONS

Limitations and Risks of Evaluation Method

One of the study limitations may be that the results of a single study might not be externally valid. We test the external validity of our results across the six countries where we implement these interventions (see the Window pre-analysis plan for details). As with any in-field RCT, spillover across communities and differential attrition are potential risks for the evaluation. The team will work closely with the implementing partners on the ground to monitor potential spillover risks and design clear implementation protocols. We expect differential attrition to be less common than in other contexts, since the control group is aware they will be receiving the FFA intervention in the second year of the programme.

Short Programme Timeline

The implementation cycles in this programme are short with two transfers spread over three months. Short programmes are not unusual, which is why it is important that we measure impacts. However, the fact that the endline data collection will happen shortly after the conclusion of the final transfer, as the control group will be phased in thereafter, will not allow the team to establish whether any impacts persist over longer timeframes.

Direct Income vs. Work Effect

The impacts of the treatment arm focusing on women's work are a combination of them engaging in work outside the household and receiving a direct cash transfer (as pay for their work). The IE design estimates the combined impact of both features, which makes it hard to disentangle the relative importance of either one feature. However, work outside the household usually entails direct pay, which makes this combination operationally relevant to investigate. There is also already a large body of literature on the impacts of cash transfers to women alone, and the contribution of our study therefore is more focused on the work component.

Risks Due to COVID-19

As a result of COVID-19, the country office has had to implement all of its programmes with third-party NGOs who are now responsible for all field-related activities. This creates additional monitoring challenges as the evaluation team has to make sure the NGOs are complying with the original design (registering dual-headed households, respecting the randomization of communities to treatment arms, and delivering cash and assets on time). The DIME team has developed a strong working relationship with the country office, and is in frequent communication with the country office and the NGOs to monitor these dynamics.

Risks due to instability

A further risk is that a crisis (for example, conflict, political instability, or natural disaster) impedes programme progress or the ability of implementing teams to follow the planned evaluation design. To mitigate the consequences of unforeseen issues, the evaluation team will work with the implementing partners to proactively resolve potential delays ex-ante, including through supporting the planning and implementation of operational activities and timely launch of procurement processes. Furthermore, field coordinators will work closely with implementing partners to ensure programme activities are conducted according to the planned standards and protocols, and to alert the evaluation team in a timely fashion about deviations and other implementation challenges.

8. Organization of the evaluation

8.1. ROLES AND RESPONSABILITIES

Table 2: IE Team and Main Counterparts

Name	Role	Organization/Unit
John Loeser	Principal Investigator, Lead Researcher	DIME
Florence Kondylis	Principal Investigator, Lead Researcher	DIME
Erin Kelley	Principal Investigator, IE Technical Team Leader	DIME
Gregory Lane	Principal Investigator	DIME
Paul Christian	Principal Investigator	DIME
Lelys Dinarte	Principal Investigator, IE Technical Team Leader	DIME
Jonas Heirman	Principal Investigator	WFP OEV
Miguel Paniagua	Field Coordinator	DIME
Sylvia Cesar	Field Coordinator	DIME
Marc-Andrea Fiorina	Research Assistant	DIME
Johanna Constanza	Monitoring and Evaluation	WFP El Salvador
Carlos Martinez	Monitoring and Evaluation	WFP El Salvador
Jaime Hernandez	Programme Policy	WFP El Salvador

Table 3: Evaluation Committee

Name	Role	Organization/Unit
Miguel Paniagua	Field Coordinator	DIME
Sylvia Cesar	Field Coordinator	DIME
Johanna Constanza	Monitoring and Evaluation	WFP El Salvador

Carlos Martinez	Monitoring and Evaluation	WFP El Salvador
Jaime Hernandez	Programme Policy	WFP El Salvador
Jaakko Valli	Deputy Country Director	WFP El Salvador
Elena Ganan	Gender Consultant	WFP Regional Bureau for Latin America and the Caribbean
Michala Assankpon	Evaluation Officer	WFP Regional Bureau for Latin America and the Caribbean
Ana Urgoiti	Evaluation Consultant	WFP Regional Bureau for Latin America and the Caribbean

8.2. WORK PLAN AND DELIVERABLES

Table 4: Milestones, Deliverables, and Estimated Timeline

Milestones	Deliverables	Completion Date
Planning & draft inception report	Draft Inception Report	December 2020
Data collection plan and pilot	TORs Questionnaires	30 January 2021
Data collection (Baseline) completed	Cleaned data Dictionaries	15 April 2021
First data analysis	Presentation Data file Do files Baseline report	15 June 2021
Implementation of intervention aligned to evaluation	Roll-out plan Monitoring reports verifying treatment and control status	May-December 2021
Midline data collection (completed)	Cleaned data	20 July 2021
Follow-up data collection plan	TORs	15 October 2021

	Questionnaire	
Data collection (Follow-up) completed	Cleaned data Dictionaries	15 November 2021
Final report and policy notes	Technical note Policy note Data file Do files	30 March 2022
Dissemination of findings	Presentations	15 December 2022

8.3. QUALITY ASSURANCE AND PEER REVIEW

WFP's Impact Evaluation Quality Assurance System (IEQAS) sets out guidance on definitions, methods, processes and procedures for ensuring that impact evaluation outputs provide robust and credible evidence about impact. The IEQAS consists of process guidance, quality checklists, templates, technical notes and other reference material to guide evaluation teams and partners throughout the evaluation process. Quality assurance will be systematically applied throughout the evaluation phases. These include preparation and selection, design, data collection[1], consistency of programme implementation with the evaluation design, analysis and reporting

Evaluation reports, including inception, baseline and final reports, are prepared by the evaluation team. Drafts are reviewed by the Evaluation Committee (see table 3 above). . Inception reports and endline reports are also reviewed by external quality support peer-reviewers, the Window's Steering Committee (SC) and the Window's technical advisory group (TAG). Reports are revised based on feedback received and reviewed by the Head of Impact Evaluation. WFP's Director of Evaluation finally approves all the reports before being submitted for publication.

In addition, all final evaluation reports will be subjected to a post hoc quality assessment by an independent entity through a process that is managed by OEV. The overall rating category of the reports will be made public alongside the evaluation reports.

---

[1] This includes using high-frequency data quality checks routinely throughout the data collection phases, and ensuring the baseline and endline reports adhere to predesignated standards set by OEV.

9. COMMUNICATION PLAN FOR DISSEMINATION AND USE

DIME and WFP will ensure that regional bureau, and the country offices are full partners in discussing and using the evidence created in the impact evaluations. DIME field coordinators will regularly update country teams on evaluation plans and keep track of any adjustments in field implementation plans to ensure that the evaluation plan remains aligned with field concerns. As data is collected, DIME will be responsible for analysis, which ensures a degree of independence in data analysis, but results of this analysis will be regularly shared and discussed with the country and regional teams to ensure that findings can be used for programme decisions and implementing teams' insights can be incorporated in the data analysis. This analysis will be shared with the relevant teams in the form of baseline and endline reports and accompanying presentations. In addition, the evaluation team will draft an academic paper for submission to a peer-reviewed journal and results from the IE will feed into the broader cross-country analysis being undertaken as part of the partnership.

In addition, DIME and WFP will communicate regularly with the respective national government and other partner agencies to provide them with updates on the IE work and results. This will be done through a series of in-country and virtual seminars (as allowed based on context). As the studies are built into WFP programmes, results will feed into future phases of these programmes. Moreover, knowledge produced by the proposed IE activities will also be more broadly relevant to other actors and governments. Lessons drawn from these IE activities will also inform future policy implementation in other regions. DIME and WFP will support the use of results from these evaluations to inform programme design of other partners by ensuring easy access and promoting awareness for the evidence generated.

Annexes

Annex 1: Window Summary and Theory of Change

1.1. WINDOW SUMMARY

The CBT and Gender Impact Evaluation Window has been developed by WFP's Office of Evaluation (OEV) in partnership with WFP's Cash-Based Transfer (CBT) Division and Gender Office (GEN), as well as the World Bank's Development Impact Evaluation (DIME) department. The Window is part of WFP's Impact Evaluation Strategy (2019-2026) – the Window will coordinate a portfolio of impact evaluations (IEs) to measure the impacts of cash transfers on gender equality and women empowerment (GEWE) outcomes across a series of WFP country programmes. The goal is to increase the predictive power of evidence generated and expand its ability to be generalized across contexts (fostering "external validity"). The hypothesis underlying the CBT and Gender Impact Evaluation Window is that providing women with opportunities to work outside the household will enhance their agency as well as increase their control over financial resources, which in turn leads to expanded social and economic empowerment. The ambition is to learn what works (and what does not) in a way that informs country office programming and contributes to a global evidence base.

A key value of the IE Window more broadly lies in WFP's ability to use evidence in-house to inform future programming and for global engagement. As the window is organized around previously identified evidence gaps, WFP will bring a unique contribution of evidence that would otherwise be missing from the country as well as global dialogue. Throughout the window, and particularly when results from individual evaluations become finalized, the Steering Committee at WFP will develop consistent, targeted policy messages corresponding to the evaluation questions, which can then be used to feed into the upcoming CSP (2022–2026). In 2019, all WFP country offices (COs) with upcoming programmes which have a FFA component were invited to express interest in the IE Window to OEV, which then engaged with COs to assess the feasibility of an IE and inclusion in the Window to provide support. In the first round, the El Salvador, Kenya, Rwanda, and the Syrian Arab Republic COs have been selected.

The first round of programmes selected for the CBT and Gender Window are anchored to a version of WFP's "Food Assistance for Assets" (FFA) intervention modality. FFA is one of WFP's most commonly used interventions aimed at addressing the most food-insecure people's immediate food needs with cash, vouchers, or food-based transfers and improving their long-term food security and resilience. The concept is simple: people receive cash, vouchers, or food-based transfers to address their immediate food needs, while they build or boost assets, such as constructing a road or rehabilitating degraded land, that will improve their livelihoods by creating healthier natural environments, reducing risks and impacts of shocks, increasing food productivity, and strengthening resilience to natural disasters. DIME and OEV have developed a window-wide pre-analysis plan (PAP) that details the overall IE design as a basis to follow across all countries included in the Window, as well as the outcomes to be measured. The specific IE design is adapted to each country context, but should still allow joint analysis across contexts. Within the FFA framework, the IE design for the CBT and Gender Window explicitly focuses on cash payments and aims to include three groups where households or communities are assigned to one of three groups:

- "Standard" FFA: Cash-based programming (households deciding who would be participating, which could be men or women).

- Women's FFA: Naming women participants in the asset creation activities and cash recipients

- Control group: Not benefiting from the FFA programme (in the first cycle – where the programs are usually rolled out in multiple cycles so control group beneficiaries receive the program at a later date).

By including a control group, the impacts of the "standard" FFA programme, which usually targets men, can be measured and compared with the impacts of not participating in FFA. The modified "Women's FFA"

treatment arm additionally allows for comparisons with the "standard" FFA arm, measuring impacts on women's social and economic empowerment when they are directly targeted by the FFA programme.

1.2. THEORY OF CHANGE

Two key elements of the intervention will be evaluated:

    a) the involvement of a household member in asset creation activities, and

    b) the transfer of money to the household (or to a named female recipient).

The impact evaluation's theory of change posits that the Women's FFA treatment arm will result in greater gains for gender equality and women's empowerment outcomes than either the second treatment arm involving an unconditional transfer to the household, or the control.

As a first step, involving women in work (asset creation through the FFA programme) directly impacts their time use (shifts towards paid work outside the home), as well as their earnings as they are paid directly for their work. The hypothesis is that – in the medium run – these combined shifts in time use and earnings will impact women's:

- Perceptions of gender norms
- Attitudes
- Agency
- Consumption patterns
- Well-being (physical, social, and psychological)

This theory of change is consistent with a body of literature that examines the impacts of providing women opportunities to work outside the household, as summarized by Field et al. (2019): "Female employment has been shown to delay marriage, increase female work aspirations, improve child health, and reduce the male:female sex ratio (...). In the United States, rapid growth in female labor force participation preceded important changes in norms regarding gender roles in both the economy and the household" (p. 1). Recent experimental work has demonstrated attitudes (Dhar et al., 2018; McKelway, 2019) and norms (Beaman et al., 2009; Bursztyn et al., 2018) shape women's agency and, in turn, women's labor supply. While FFA programs have demonstrated to be an effective tool for economic development through increased earnings (Imbert & Papp, 2015; Gazeaud et al., 2019; Adjognon et al., 2020), there is less evidence on the impacts of participant gender and there is also a lack of evidence on programmes with a short duration.

In Figure 1a, solid lines trace out the direct impacts of these changes, while dotted lines trace out secondary impacts. For example, suppose we saw that women's participation in public works shifted only agency, time use, earnings, and consumption. We would conclude that impacts on agency were caused by changes in earnings and time use.

In the longer run, we hypothesize including women in work outside the home can initiate a "virtuous cycle" where a change in women's perceptions of norms, attitudes, and agency further boosts women's participation in paid work outside the home (time use). This then positively impacts their earnings, which could amplify consumption and well-being, even after the FFA intervention ends.

In Figure 1b, solid lines trace out the direct impacts of these changes, while dotted lines trace out secondary impacts. For example, suppose we saw long run changes in attitudes and time use. We would conclude that the impacts on time use were driven by persistent changes in attitudes, as opposed to changes in perceptions of norms or agency.

While the programme is targeted at women, it is possible the programming will also impact men's perceptions of gender norms (and those of the wider community) and attitudes in a way that further contributes to improvements in gender equality.
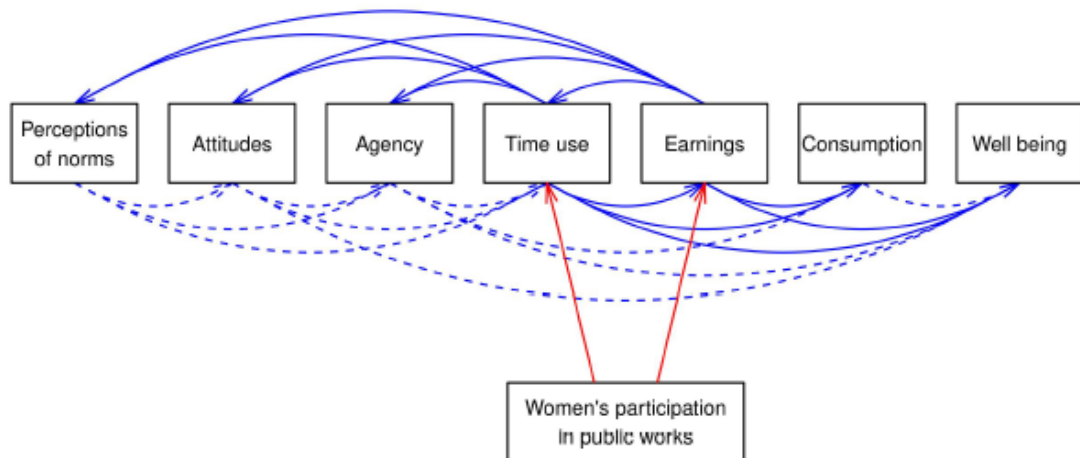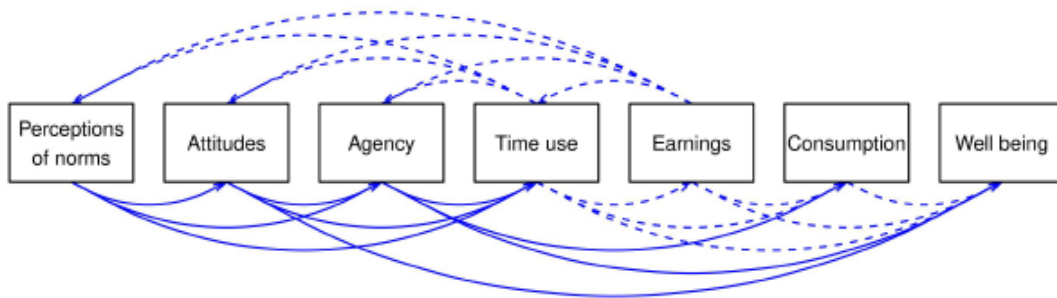
Figure 1a: Theory of Change (Medium Run)

Figure 1b: Theory of Change (Long Run)

Annex 2: Questionnaires

The baseline questionnaire can be found in full at the link here.

Annex 3: Detailed stakeholder analysis

Stakeholders and users of this evaluation are defined as those actors that may influence the evaluation, and those that may be influenced by it. This includes internal, external, national actors and programme beneficiaries. The WFP Country Office in El Salvador is intended to be the primary user of this evaluation, however, the evaluation aims to promote learning and widespread use of the findings generated even beyond the country office.

The various categories of stakeholders include:

- Internal El Salvador based stakeholders: the Country Director and Deputy Director, the Head of Programme, and all technical and management personnel.
- Internal stakeholders outside of El Salvador: OEV, the Regional Bureau of Latin America and the Caribbean (RBP), and the CBT and Gender divisions and Protection unit at Headquarters.
- Population groups in need (affected populations): resident communities and migrants of different sexes and age groups.
- External stakeholders: which includes INGOs, donors, UN agencies and forums in El Salvador.
- National stakeholders: which includes national and sub-national government actors, and NGOs

The main users of the evaluation, i.e. CO management and WFP staff in-country, may be much affected by the evaluation and are actively engaged in its development. Populations in need of WFP assistance will also have a high stake in the results, and will be the primary providers of data for the evaluation.

Stakeholder engagement will vary depending on category, but may include:

- Reviewing and commenting on the TORs and draft Inception report;
- Active monitoring of the evaluation design during programme implementation;
- Participation in the final learning workshop;
- Reviewing and commenting on the draft evaluation report (which will be available in English and Spanish);
- Reading of the final evaluation report and other evaluation communication products.

More detailed information about evaluation users is provided in Table 5 below. This table introduces all categories of stakeholders, the degree to which they have expressed an interest to be included in the evaluation, how they might be engaged and how they are expected to use the evaluation results.

Table 5: Stakeholder Analysis

Who are the stakeholders?	What is their role in the intervention?	What is their interest in the evaluation?	How should they be involved in the evaluation? (be informed, act as key informant, be part of a focus group interview, be part of a reference group, etc.)	At which stage should they be involved?	How important is to involve them in the evaluation? (High, medium, low)
WFP internal stakeholders					
WFP country office	Main implementers of the programme under evaluation	To inform upcoming Country Strategic Plan and relevant programming	The CO is responsible for implementing the programme according to the evaluation design. They actively provide feedback on the tools and outputs of the evaluation.	From the scoping stage	High
WFP Regional Bureau	Governance and technical advisory role	To inform regional programme strategies, to support other COs in evidence generation	As members of the Evaluation Committee; technical advisors on relevant portions of the questionnaire, data collection activities and implementation	From the scoping stage of the evaluation, with regular meetings to provide feedback on tools and ouputs	High
Office of Evaluation	Coordination of IE window and liaisons with CO	As coordinators of the IE and in alignment with the Impact Evaluation Strategy (2019-2026)	The impact evaluation team will be involved in the field coordination meetings and Evaluation Committee meetings as support to the CO and IE team	From the scoping stage	High
External stakeholders					
Affected communities	Affected communities,	Beneficiaries will likely have strong	Beneficiaries and non-beneficiaries alike will provide the	From the targeting and	High

	including men, women, boys, and girls will be the primary participants of the intervention	interest in any changes in targeting, reach, or effectiveness of future programming as a result of the evaluation and recommendations. Women and girls have particular stake in the results meant to shed light on recommendations for improving gender equality	primary source of data on effectiveness	selection stage	
Government at local level	Community development association (ADESCO)	As leading community members of the beneficiary communities, the ADESCO is interested in facilitating an evaluation of the effectiveness	The ADESCO plays a key roll at facilitating access to selected communities and allowing for data collection to take place	At the targeting phase of the intervention	High
Government at regional level	Municipal Government	As the primary source of governmental support for the affected communities, the municipal	The municipal government is kept aware of the activities of WFP in its communities, including the intervention itself and the associated data collection	At the targeting phase of the intervention	Medium

		government is a key partner in the implementation of WFP prorgramming			
Government at central level	General Directorate for Civil Protection (DGPC)	As the primary governmental body for preparedness and emergency response, the results of the evaluation will be useful to inform broader national policy	The DGPC influences which communities are selected for WFP programming and is thus a key stakeholder in setting the geographic focus of the evaluation	At the initial scoping for the intervention	Medium
International NGOs	Plan International; EDUCO	Plan International and EDUCO are both cooperating partners for the implementation of the FFA programe and will be able to use the results for their own future programming.			Low
World Bank	Development Impact Evaluation (DIME) Unit	In line with the OEV-DIME partnership, DIME is interested in producing and disseminating the evaluation results	As the primary investigators and research analysts	At the initial conceptualization of the Window	High

		as part of a broader research portfolio			

Annex 4: Detailed evaluation process

Phase 1 – Preparation		Involved	Estimated Date
	Initial discussion between CO and OEV to assess the feasibility	CO/OEV	November 2020
	MoU between OEV and CO signed	CO/OEV	May 2021
	Set up impact evaluation (IE) team and Evaluation Committee (EC)	OEV/DIME	January 2020
	Agreement on the questions, design, implementation and timelines between CO and IE team	DIME/OEV/CO	December 2020
	Targeting potential intervention sites (including both potential intervention and comparison areas)	CO/DIME	December 2020
Phase 2 - Inception Report			
	Inception Report drafted by IE team, submitted for quality assurance and revisions	DIME	April 2021
	Publication of the Inception Report	OEV	November 2021
	Dissemination of the inception report with country office, regional bureau, evaluation committee, window's reference group, steering committee, online/social media as adequate.	DIME/OEV	November 2021
Phase 3 – Baseline data collection			
	Preparation data collection tools, including survey questionnaire, digital devices, sampling strategy, training material, etc.	DIME	December 2020/January 2021
	Pilot and finalization of data collection tools	DIME/CO	January 2021
	Recruitment enumerators/data collection firm	CO	February 2021
	Enumerators training	DIME/CO	February 2021
	Data collection process and live monitoring data quality checks	DIME/CO	February-April 2021
Phase 4 – Baseline Report			
	Data analysis and baseline report drafted by IE team, submitted for quality assurance and revisions	DIME	September 2021
	Publication of the Baseline Report	OEV	December 2021
	Dissemination of the Baseline Report with survey respondents, country office, regional bureau, evaluation committee (and other evaluation stakeholders), window's reference group, steering committee, online/social media as adequate.	DIME/OEV	December 2021
Phase 5 – Programme implementation			
	Randomization	DIME	January 2021
	Assignment intervention and comparison sites	DIME/CO	January 2021

	Rollout programme activities as per randomization	CO	May 2021
	Monitoring programme activities verifying treatment and control status	CO/DIME	May-December 2021
Phase 6 – Endline data collection			
	Preparation data collection tools, including survey questionnaire, digital devices, sampling strategy, training material, etc.	DIME/CO	October 2021
	Pilot and finalization of data collection tools	DIME	October 2021
	Recruitment enumerators/data collection firm	CO	October 2021
	Enumerators training	CO	November 2021
	Data collection process and live monitoring data quality checks	DIME	Nov-December 2021
	Feedback/ Data sharing mechanisms, as appropriate/possible		
Phase 7 – Final Evaluation Report			
	Data analysis and final evaluation report drafted by IE team, submitted for quality assurance and revisions	DIME	December 2021
	Publication of the Final Evaluation Report	OEV	April 2022
	Dissemination of the Final Evaluation Report with survey respondents, country office, regional bureau, evaluation committee (and other evaluation stakeholders), window's reference group, steering committee, online/social media as adequate.	OEV/DIME/CO	May 2022
	Final Evaluation Report reviewed by post-hoc quality assessment (PHQA)	OEV	May 2022
Phase 8 – Management response			
	Based on findings CO to develop a management response	CO	May 2022
	OEV to review and if needed respond to the management response	OEV	May 2022
	Publication of the management response	OEV	June 2022
Phase 9 – Dissemination and Learning			
	Webinar presenting the findings	OEV/DIME	June 2022
	Blogs, summary briefs, other relevant communication products	OEV/DIME	Ongoing
	Considerations for academic publication	DIME/OEV	July 2022

Annex 5: References

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. Journal of the American statistical Association, 103(484), 1481–1495.

Athey, S., Chetty, R., Imbens, G., & Kang, H. (2016). Estimating treatment effects using multiple surrogates: The role of the surrogate score and the surrogate index. arXiv preprint arXiv:1603.09326.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienťe, W., Shapiro, J., Thuysbaert, B., & Udry, C. (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. Science, 348(6236), 1260799.

de Brauw, A. & Peterman, A. (2020). Can conditional cash transfers improve maternal health care? Evidence from El Salvador's Comunidades Solidarias Rurales program. Health Economics, 29(6), 700–715.

Field, E. M., Pande, R., Rigol, N., Schaner, S. G., & Moore, C. T. (2019). On Her Own Account: How Strengthening Women's Financial Control Affects Labor Supply and Gender Norms (No. w26294). National Bureau of Economic Research.

Haushofer, J., Ringdal, C., Shapiro, J. P., & Wang, X. Y. (2019). Income Changes and Intimate Partner Violence: Evidence from Unconditional Cash Transfers in Kenya. Working Paper 25627, National Bureau of Economic Research.

Haushofer, J. & Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya. The Quarterly Journal of Economics, 131(4), 1973–2042.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. Annual Review of Economics, 7(1), 63–88.

Sakho, S., Caruso, G., & Acuña, M. (2020). Progress without women is not possible. World Bank Blogs: Latin America and Caribbean. https://blogs.worldbank.org/latinamerica/progress-without-women-not-possible

World Food Programme, Gender Office (2019). The potential of cash-based interventions to promote gender equality and women's empowerment: A multi-country study. https://www.wfp.org/publications/gender-and-cash-wfp-study

World Food Programme, Office of Evaluation. (2020). El Salvador: An Evaluation of WFP's Country Strategic Plan (2017–2021).

Acronyms

CAPI	Computer assisted personal interviewing
CBT	Cash-based transfer
CO	Country office
CSP	Country strategic plan
DIME	Development Impact Evaluation Unit (World Bank)
ENIGH	El Salvador Income and Expenditure Survey
FFA	Food-assistance-for-assets
GEN	Gender Office (World Food Programme)
GEWE	Gender equality and women's empowerment
HDI	Human development index
HFC	High-frequency checks
IE	Impact evaluation
IPC	Integrated food security phase classification
IPV	Intimate partner violence
IRB	Institutional review board
OEV	Office of Evaluation (World Food Programme)
PAP	Pre-analysis plan
PHQ9	Patient Health Questionnaire-9
RCT	Randomized controlled trial
WFP	World Food Programme