



WFP EVALUATION

WFP
World Food Programme

SAVING LIVES
CHANGING LIVES

Impact Evaluation for Resilience Learning in the Sahel

Niger Inception Report

i2i
DIME
TRANSFORM DEVELOPMENT

Contents

Acronyms	1
Key Personnel for the Evaluation	2
OFFICE OF EVALUATION (WORLD FOOD PROGRAMME).....	2
DEVELOPMENT IMPACT EVALUATION TEAM (WORLD BANK)	2
WFP NIGER COUNTRY OFFICE.....	2
1. Introduction.....	3
2. Evaluation context	4
2.1. BACKGROUND AND CONTEXT	4
2.2. PROGRAMME DESCRIPTION	5
2.3. THEORY OF CHANGE AND HYPOTHESES	8
3. Evaluation approach and questions.....	10
4. Evaluation methodology	13
4.1. CLUSTER RANDOMIZED CONTROLLED TRIAL (RCT) DESIGN:.....	13
4.2. HETEROGENEITY ANALYSIS.....	14
4.3. QUALITATIVE ANALYSIS	15
5. Data collection and measurement.....	16
5.1. DATA COLLECTION FOR THE RCTS	16
5.2. DATA USED FOR HETEROGENEITY ANALYSIS.....	20
5.3. OUTCOMES MEASURED	20
5.4. SURVEY IMPLEMENTATION.....	21
5.5. MANAGEMENT OF DATA QUALITY.....	22
5.6. IMPLEMENTATION MONITORING SYSTEM.....	24
6. Data processing and analysis.....	25
6.1. DATA CODING, ENTRY, AND EDITING.....	25
6.2. PROGRAMME-SPECIFIC QUANTITATIVE DATA ANALYSIS.....	25
6.3. PROGRAMME-SPECIFIC QUALITATIVE DATA ANALYSIS.....	26
7. Ethical considerations	27
8. Risks and limitations	28
9. Quality assurance and peer review.....	30
10. Communication Plan	31
11. Organization of the evaluation.....	32
ANNEX 1: WINDOW SUMMARY	34
ANNEX 2: DEFINING RESILIENCE	35
ANNEX 3: MAIN OUTCOMES OF INTEREST.....	39
ANNEX 4: VILLAGE AND HOUSEHOLD LISTING	41
ANNEX 5: POWER CALCULATIONS	43
ANNEX 6: QUESTIONNAIRES	45
ANNEX 7: DETAILED STAKEHOLDER ANALYSIS	46
ANNEX 8: DETAILED EVALUATION PROCESS.....	49
ANNEX 9: REFERENCES	51

Acronyms

CAPI	Computer-assisted personal interviewing
CBT	Cash-based transfer
CO	Country Office
CSP	Country Strategic Plan
DIME	Development Impact Evaluation (DIME) department (World Bank)
FFA	Food Assistance for Assets
IE	Impact evaluation
IRB	Institutional review board
OEV	Office of Evaluation (World Food Programme)
PAP	Pre-analysis plan
RCT	Randomized controlled trial
WFP	World Food Programme

Key Personnel for the Evaluation

OFFICE OF EVALUATION (WORLD FOOD PROGRAMME)

Director of Evaluation, Andrea Cook

Evaluation Officer, Jonas Heirman

Evaluation Officer, Hanna Paulose

Evaluation Analyst, Theresa Schneider

Evaluation Analyst, Ola Eltoukhi

DEVELOPMENT IMPACT EVALUATION TEAM (WORLD BANK)

Principal Investigator, Patrick Premand (Lead Researcher, contact: ppremand@worldbank.org)

Co-Principal Investigator, Guignonan Serge Adjognon

Resilience Window Manager, Marcus Holmlund

Research Officer, Chloë Fernandez

Field Coordinator, Abdoul-Aziz Adama

Research Assistant, Pulkit Bajpai

WFP NIGER COUNTRY OFFICE

Head of Food Security Analysis (VAM) and Monitoring and Evaluation, Anna Law

Head of Programme, Koffi Akakpo

1. Introduction

1. In 2020, 155 million people worldwide faced “crisis or worse” levels of food insecurity. Close to 115 million of them lived in countries affected by conflict or weather extremes.¹ Extended food crises are predicted in 2022, which would necessitate continued large-scale humanitarian assistance. In Niger, the World Food Programme (WFP) estimates that 20 percent of the population cannot meet their food needs because of various social and environmental factors.² Conflict, forced population displacements, and climate change are exacerbating food insecurity.³ However, evidence is lacking on how development outcomes are affected by these shocks, and how WFP’s programmes support populations to effectively respond to these shocks.
2. WFP’s Office of Evaluation (OEV), Asset-Creation, Livelihoods and Resilience Unit, and Climate and Disaster Risk Reduction Unit partnered with the World Bank’s Development Impact Evaluation (DIME) department to create the Climate and Resilience Impact Evaluation (IE) Window. WFP’s Impact Evaluation Strategy (2019–2026) focuses on delivering impact evaluations that contribute to global evidence and organizational learning. Impact evaluation windows help to achieve this strategy by focusing the portfolios of impact evaluations on priority evidence needs identified through literature reviews and extensive consultations.
3. The Climate and Resilience Window is intended to enable understanding of how WFP’s programmes contribute to the resilience of the populations supported. The first round of impact evaluations selected for this window are intended to estimate the effects of integrated packages of resilience activities on households’ capacities to absorb shocks (absorptive capacity), adapt to increasing environmental or economic stressors (adaptive capacity), and improve well-being in the long-term (transformative capacity).
4. The Niger impact evaluation is intended to estimate the effects of WFP’s Integrated Resilience Programme on absorptive, adaptive, and transformative resilience capacities. WFP’s Integrated Resilience Programme in Niger consists of activities including Food Assistance for Assets (FFA), nutrition support, school feeding, Smallholder Agriculture Market Access (SAMS) activities, and lean season support. The intervention is intended to directly increase both household resilience capacities and food security.
5. This Inception Report outlines the strategy for assessing the impact of WFP’s resilience programme in Niger on the dynamics of population well-being and resilience. Through this impact evaluation, WFP and DIME are working together to complement other ongoing efforts and guide future investments and activities related to resilience in the Sahel.
6. This Inception Report also builds on a [pre-analysis plan \(PAP\)](#) registered with the American Economic Association’s registry for randomized controlled trials. The pre-analysis plan includes detailed information on primary outcomes, research design, randomization method, randomization unit, clustering, sample size (total number, number of clusters, and units per treatment arm), and regression specifications. The purpose of the PAP is to outline the set of hypotheses and analyses that will be performed on the data before it is collected.

¹ FSIN. 2021. Global Report on Food Crises: Joint Analysis for Better Decisions. [\[Link\]](#)

² World Food Programme. n.d. Niger. [\[Link\]](#)

³ FSIN. 2021. Global Report on Food Crises: Joint Analysis for Better Decisions. [\[Link\]](#)

2. Evaluation context

2.1. BACKGROUND AND CONTEXT

7. Niger is a land-locked and food-deficit Sahelian country with a population of over 23.2 million people and annual demographic growth of 3.9 percent, the highest in Africa. It is one of the poorest countries in the world, with gross domestic product (GDP) per capita of US\$895 in 2015 (in constant 2011 US\$) and a poverty rate of 51.4 percent in rural areas.⁴ Niger ranks last among the 189 countries on the United Nations Human Development Index.⁵ Food and nutrition insecurity are driven by poverty, environmental degradation, and other recurring shocks. The November 2021 Cadre Harmonisé (CH) estimated that 2.5 million people were food insecure during November–December 2021, and that 3.6 million would be food insecure over the next agricultural lean season (June–August 2022).⁶ In 2021, nearly 20 percent of the population could not meet their food needs because of factors such as inadequate agricultural production, security constraints, and high demographic growth.⁷ This figure rises to nearly 30 percent during periods of poor rainfall. Additionally, because of regional conflicts, primarily the Boko Haram crisis in Nigeria and instability in Mali, Niger currently hosts 280,583 refugees, 264,257 internally displaced persons, and 37,305 Nigerien nationals who have returned from Nigeria.⁸
8. Niger’s rural population is highly exposed to climatic shocks such as drought, and the incidence and severity of these shocks is expected to increase because of climate change.^{9,10} Employment is highly concentrated in subsistence (rain-fed) agriculture, in which more than 80 percent of the labour force is engaged. Cereal crops, such as millet and sorghum, are the most common, and agricultural production is correlated with rainfall variation. Agricultural production is constrained by a short rainy season and limited access to irrigation.
9. Climate-related shocks are particularly frequent in Niger and have detrimental effects on rural households. Annan and Sanoh (2018) estimate that in Niger, household consumption declines by 31–48 percent when households are exposed to extreme shocks – leading to a wide range of costly coping strategies. Weather shocks are also found to negatively affect technology adoption and use of modern inputs,¹¹ and to be associated with large movements in food prices.¹² While weather shocks are not necessarily driven by climate change, households in rural Niger report perceiving rainfall as scarcer, more erratic, more often delayed, and more likely to generate droughts than before 2009.¹³ The high welfare costs of shocks highlighted in these studies indicate that households’ risk-coping and risk-management strategies are imperfect.¹⁴

⁴ World Bank. 2016. Republic of Niger: Priorities for Ending Poverty and Boosting Shared Prosperity Systematic Country Diagnostic (SCD) Concept Note. World Bank: Washington, D.C.

⁵ United Nations Development Programme. 2021. Human Development Report 2020: The Next Frontier-human Development and the Anthropocene. UN.

⁶ World Food Programme. 2022. Niger Country Brief. January 2022 [[Link](#)]

⁷ World Food Programme. 2021. Niger Country Brief. January 2021. [[Link](#)]

⁸ UNHCR Operational Portal. Figures as of 31 January 2022. [[Link](#)]

⁹ World Bank. 2016. Republic of Niger: Priorities for Ending Poverty and Boosting Shared Prosperity Systematic Country Diagnostic (SCD) Concept Note. World Bank: Washington, D.C.

¹⁰ World Bank. 2009. World Development Report 2010: Development and Climate Change. Washington, DC.

¹¹ Asfaw, S., Pallante, G. & Palma, A. 2018. Diversification Strategies and Adaptation Deficit: Evidence from Rural Communities in Niger. World Development, 101: 219–234.

¹² Aker, J. C., Boumnijel, R., McClelland, A. & Tierney, N. 2016. Payment Mechanisms and Antipoverty Programs: Evidence from a Mobile Money Cash Transfer Experiment in Niger. Economic Development and Cultural Change, 65(1): 1–37.

¹³ World Bank. 2016. Republic of Niger: Priorities for Ending Poverty and Boosting Shared Prosperity Systematic Country Diagnostic (SCD) Concept Note. World Bank: Washington, D.C.

¹⁴ Asfaw, S., Pallante, G. & Palma, A. 2018. Diversification Strategies and Adaptation Deficit: Evidence from Rural Communities in Niger. World Development, 101: 219–234.

10. Niger has received substantial humanitarian assistance over the years. For example, in 2020, WFP alone carried out humanitarian operations to a value of US\$161 million.¹⁵ A range of interventions has been implemented in response to shocks and seasonal food insecurity, including cash and food transfers during the lean agricultural season.^{16,17} Over time, policymakers and development stakeholders have made efforts to better design emergency responses to shocks, as well as to support more permanent approaches to strengthen households' ability to protect themselves against future shocks. These efforts include expanding the coverage of a government-led national safety net programme, as well as the layering of behavioural change interventions and multi-faceted economic inclusion or livelihood support interventions to promote economic diversification and the resilience of households participating in the national cash transfer programme.^{18, 19, 20, 21}
11. Consistent with these broader efforts at national level, a key focus of WFP's Country Strategic Plan (2020–2024) in Niger is to implement integrated resilience activities to protect livelihoods and foster long-term recovery. WFP has established a resilience programme that layers a set of integrated interventions on Food Assistance for Assets (FFA) interventions. The main objective of the WFP resilience programme is to strengthen the socioeconomic resilience of smallholder farmers and vulnerable populations. The programme is intended to build the resilience of food systems and livelihoods of targeted communities, while also strengthening community structures to support social cohesion, thus contributing to conflict prevention and prospects for peace.
12. Given this context and policy environment, DIME and WFP, with support from Germany's Federal Ministry for Economic Cooperation and Development (BMZ), are collaborating to generate evidence on how multiple interventions can be combined or sequenced to boost the resilience of poor and vulnerable households in Niger. This evidence agenda contributes to the implementation of WFP's integrated resilience programme in the Sahel, and is also of interest for other initiatives in the region, such as the Sahel Adaptive Social Protection Program managed by the World Bank
13. This impact evaluation for Niger is intended to identify the impact of WFP's programme on resilience, as measured by people's capacity to maintain food security while experiencing shocks and stressors. The evaluation will also use two-monthly surveys to track the food security and coping strategies of households supported by the programme and those who are not part of the programme. This approach will help provide an understanding of when food security peaks in Niger, and which households are vulnerable to becoming food insecure at different phases. Surveys will be complemented by qualitative data and analysis to understand how the programme is implemented, how the support provided is perceived by beneficiaries and opportunities for future improvements, and to generate insights about the patterns observed in the quantitative data.

2.2. PROGRAMME DESCRIPTION

14. WFP's [Country Strategic Plan in Niger 2020 – 2024](#) (CSP) stresses the importance of maintaining WFP's capacity to respond to emergency needs while also increasing its focus on longer-term outcomes (such as diversified livelihoods, agricultural productivity, peace, and social cohesion), to improve households' and communities' capacities to respond to shocks and stressors, and thus to enhance resilience.

¹⁵ World Food Programme Niger. 2021. Annual Country Report 2020. [\[Link\]](#)

¹⁶ Aker, J. C., Boumnijel, R., McClelland, A. & Tierney, N. 2016. Payment Mechanisms and Antipoverty Programs: Evidence from a Mobile Money Cash Transfer Experiment in Niger. *Economic Development and Cultural Change*, 65(1): 1–37.

¹⁷ Hoddinott, J., Sandström, S. & Upton, J. 2018. The Impact of Cash and Food Transfers: Evidence from a Randomized Intervention in Niger. *American Journal of Agricultural Economics*, 100(4): 1032–1049.

¹⁸ Premand, P. & Barry, O. 2020. Behavioural Change Promotion, Cash Transfers and Early Childhood Development: Experimental Evidence from a Government Program in a Low-income Setting. Policy Research Working Paper 9368. World Bank: Washington, DC.

¹⁹ Premand, P. & Stoeffler, Q. 2020. Do Cash Transfers Foster Resilience? Evidence from Rural Niger. Policy Research Working Paper 9473. World Bank, Washington, DC.

²⁰ Archibald, E., Bossuroy, T. & Premand, P. 2020. *Productive Inclusion Measures and Adaptive Social Protection in the Sahel: A Case Study*. Washington, DC: The World Bank.

²¹ Bossuroy, T., Goldstein, M., Karlan, D., Kazianga, H., Parienté, W., Premand, P., Thomas, C., Udry, C., Vaillant, J. & Wright, K. 2021. Pathways Out of Extreme Poverty: Tackling Psychosocial and Capital Constraints with a Multi-Faceted Social Protection Program in Niger. Policy Research Working Paper 9562. World Bank, Washington, DC.

15. Niger's resilience programme is aligned with the CSP objectives and includes interventions that aim to promote the capacities of households and communities to absorb shocks, adapt to risks, transform livelihoods and – more broadly, in the living environment – to exit poverty. It includes (i) food assistance for assets (FFA), (ii) nutrition and health, (iii) value chain and smallholder agriculture market support (SAMS), (iv) school feeding, and (v) lean season support through seasonal cash transfers. The planning and prioritization of these interventions is supported and guided by the community-based participatory planning (CBPP) process.
16. By introducing a combination of layered and sequenced activities targeting the most vulnerable, the WFP resilience programme in Niger intends to promote the resilience capacities of individuals, households, and communities:²²
 - At the community level, activities intended to promote resilience focus on environmental rehabilitation and food-systems development. These include the FFA and SAMS components.
 - At the individual and household levels, activities promoting resilience include investments in human capital through the provision of an integrated package of school feeding and nutrition services, and lean season support for vulnerable households. These activities complement government efforts to strengthen the resilience of vulnerable communities.
17. More specifically, the WFP resilience programme in Niger includes the following:
 - *The Food Assistance for Assets (FFA) component* is intended to meet the immediate food needs of households while restoring degraded landscapes, improving water harvesting, reducing the risk of environmental disasters, creating productive assets to secure ecosystem services, increasing productivity and yield, supporting economic development, and strengthening social ties between community members and villages. New assets are selected through a community-wide participatory process and are built through asset creation activities tied to cash or in-kind support.²³ Some of the assets are built commonly by a set of villages. Villages that may benefit from the assets and restored landscapes are identified within the sites. The majority of sites in the impact evaluation area include activities related to land rehabilitation for agricultural and/or pastoral purposes, for instance through soil and water conservation activities, protective measures and land restoration (CES/DRS) activities and assisted natural regeneration (RNA) activities. Typical examples of related assets include half-moons, stone bunds, dune fixations, grazing areas, and some activities such as fights against *pergularia/cordifolia*. Approximately half of the sites include activities to increase access to water for domestic, fishing, and agricultural purposes, mainly through pond restoration. Beneficiary households within an asset site are targeted using the Household Economy Approach (HEA). Households are categorized into four socioeconomic groups: Very Poor, Poor, Average, and Well-off. Households in the Very Poor and Poor categories are eligible to participate in FFA activities. They are paid approximately US\$43 per month (25,000 CFA) to work on the assets, typically for two or three months between March and May, with a few exceptions that are context-specific (for example, some FFA measures can start in November/December or continue in July for tree planting and eradication of invasive species).²⁴

²² Throughout the programme, WFP is working with governments and other partners. This includes activities that strengthen the capacity of national and decentralized institutions to design, deliver, and learn from integrated resilience efforts, including through support to Universities (e.g., deployment of second year master students (MSCII) in resilience sites to undertake research and deliver training), and dedicated packages to technical ministries to review national norms and provide technical support.

²³ To select sites, WFP, government technical services, local authorities and the community make the inventory of the potential and constraints for natural resource development based on a geographical approach and community-based participatory planning (CBPP). Once sites have been identified, a group of surrounding villages participates in the community-based participatory planning process which serves to identify which priority interventions, including the livelihood assets, communities will build. This provides a platform for inclusive community engagement, where the most vulnerable, marginalized, and disempowered have a voice in community decisions. Participatory planning facilitates agreements for access to land and water resources for women's groups, youth, refugees, internally displaced persons, returnees, and the very poor.

²⁴ Households received on average approximately 75,000 CFA (approximately US\$130) for FFA participation over three months in 2021. Their average yearly consumption at baseline was 372,000 CFA (approximately US\$640). Therefore,

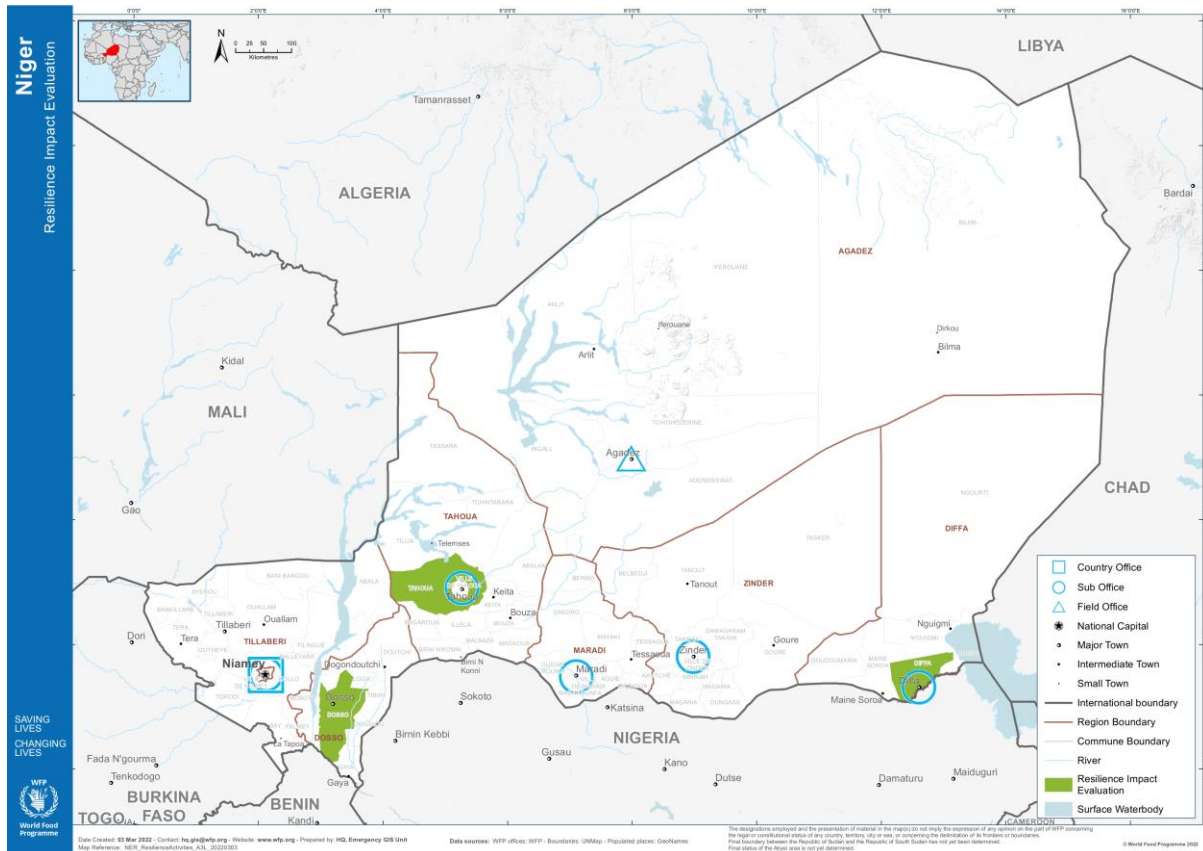
- *The Nutrition/Health component* is intended to address acute malnutrition and prevent all forms of malnutrition through improved nutrient intake and consumption behaviours. WFP implements the following activities within the resilience programme in Niger: i) screening and provision of food supplements to children aged 6–23 months through the seasonal Blanket Supplementary Feeding Programme (BSFP); ii) supporting moderately acutely malnourished children (aged 6–59 months) through the Targeted Supplementary Feeding Programme (TSFP); iii) providing Specialized Nutritious Food (SNF) and/or cash transfers to pregnant or lactating women and girls at each prenatal consultation; iv) sensitization campaigns on improving maternal, infant and young child nutrition and feeding practices for better nutrition outcomes; and v) promoting the use of crops with high nutritional value in restored areas. Specifically, the following support is provided for households with high malnutrition levels participating in FFA activities: i) distribution of Super Cereal and Super Cereal Plus for each child aged 6–23 months in the household; ii) malnutrition diagnosis, treatment, or referral to services for children aged 6–23 months; and iii) and distribution of iron - folic acid for pregnant or lactating women and girls. In addition to the three categories of support above, WFP supports health centres (CSI) to run awareness campaigns and to systematically diagnose, treat, or refer to specialized services children aged 6–23 months.
- *The Smallholder Agriculture Market Support (SAMS) component* is intended to support the development of value chains and improve the sustainable management of FFA assets through a set of complementary activities, such as storage for reduction of post-harvest loss, support for food processing, capacity building, support for the organization of small producers into farmers' organizations, and market access facilitation. The SAMS component targets farmers' organizations and is seen as a key pathway to valorize assets and increase their returns to local communities. The goals are to intensify and improve farmers' smallholder production; contribute to women's empowerment; promote sustainable management of community assets and natural resources; improve the livelihoods of small producers (food production, trade, and so on); support farmers' organizations, especially women's groups to have better access to markets; and increase the income of small producers. In Niger, one of the main activities is the purchase of cereals (millet) and legumes (cowpeas) between November and February from farmer organizations with a production surplus.
- *The School Feeding component* is intended to increase access to education and school retention rates by providing nutritious school meals, take-home rations, and scholarships. Complementary activities include nutrition education, the creation of school gardens, training activities on canteen management, leveraging schools as a platform to contribute to food diversification and deliver messages on hygiene, family practices, and environmental stewardship. The schools in the villages around the sites and their students benefit from the school feeding component. Girls aged 10–19 in FFA households are eligible to benefit from an additional scholarship programme implemented through schools.
- *The Lean Season Support component* comprises unconditional cash/food assistance provided to extreme poor households to offset the peak hunger and malnutrition period. It amounts to an average of US\$52 (30,400 CFA) per month for two or three months between June and August.²⁵ Lean season support helps households to meet their gaps in food needs during the lean season. The support may only be provided in the first year of the resilience programme.

18. In Niger, the integrated resilience programme is currently being implemented in 261 sites (each site is a cluster of multiple villages) across three regions. The impact evaluation will focus on new sites added to the resilience programme in 2021 in the regions of Diffa (communes of Foulatari, Goudoumaria, and N'guelbely), Dosso (communes of Falwel, Loga, and Sokorbe), and Tahoua (communes of Allakaye,

these FFA transfers represent approximately 20 percent of average yearly consumption (75,000/372,000 CFA). The FFA transfer modalities vary by region. In Tahoua, households receive 26,000 CFA per month in cash for two months (April and May). Households receive food transfers for FFA for three months (March–May) in Dosso and Diffa, which are worth 22,600 CFA (approximately US\$39) per month in Dosso and 26,900 CFA (approximately US\$46) per month in Diffa.²⁵ Households in Tahoua receive unconditional lean season transfers in cash for two months (June and July), while households in Dosso and Diffa receive unconditional food transfers for three months (June, July, and August). The transfers are worth, respectively, 19,500 CFA per month (approximately US\$33), 33,300 CFA (approximately US\$57), and 38,400 CFA (approximately US\$66).

Bagaroua, Bambeye, Garhanga, Keita, Tabalak, and Tebaram). Figure 1 shows the communes included in the impact evaluation.

Figure 1: Map of the communes included in the impact evaluation



2.3. THEORY OF CHANGE AND HYPOTHESES

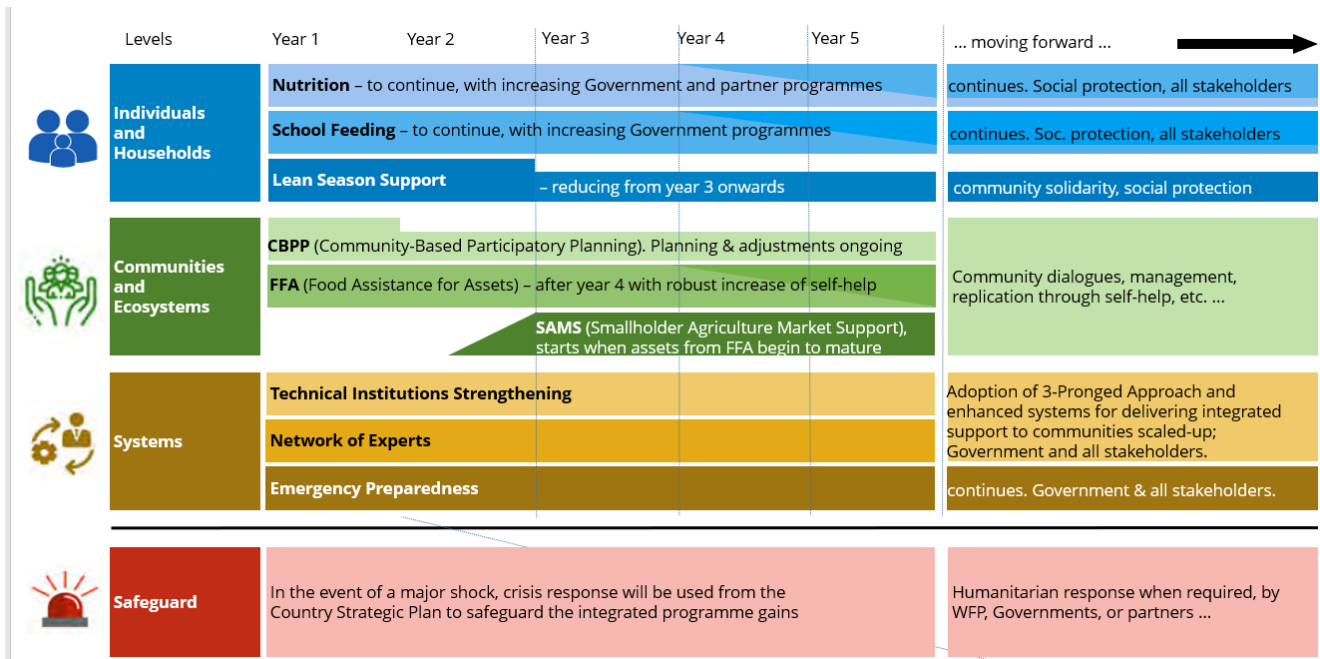
19. The theory of change for the resilience programme assumes that supporting communities through multiple activities focusing on various outcomes will: 1) support people to ensure their short-term well-being; and 2) enhance people’s capacity to maintain and improve well-being while facing shocks and stressors. Figure 2 below summarizes the key intermediate outcomes and the impact the programme focuses on. This is a simplified version of the full programme Theory of Change.

20. The resilience impact evaluation in Niger is intended to test the following hypotheses:

- Hypothesis 1: In the short term, the WFP’s resilience programme will support people to maintain their food security by meeting a household’s immediate food needs that may arise during a shock or stressor. The effect of activities focused on meeting immediate food needs will be reflected mainly in:
 - Household-level food consumption
 - The coping strategies of households.
- Hypothesis 2: In the medium term, the WFP resilience programme will support households by improving capacities associated with maintaining and/or improving food security while experiencing multiple and/or recurring shocks and stressors. These capacities include:
 - Livelihood activities
 - Household assets
 - Financial outcomes (such as income, savings, and expenses)

- Variations in food consumption over time.

Figure 2: Theory of Change of the programme in Niger²⁶



²⁶ 'Scaling up for resilient individuals, communities and systems in the Sahel' – Operational Reference Note, RBD, Dakar - October 2018.

3. Evaluation approach and questions

21. Impact evaluations measure changes in development outcomes of interest for a target population that can be attributed to a specific programme or policy through a credible counterfactual. WFP's ability to establish a credible counterfactual for programme interventions depends on logistical and financial constraints. Impact evaluations are therefore restricted to focusing on a set of questions that can be answered during a programme cycle using credible counterfactuals.
22. Regional discussions, in-country consultations, and subsequent conversations with the programme and monitoring and evaluation teams have led to the adoption of a gradual approach to building an impact evaluation learning agenda that is suitable for Niger's context and WFP's programme implementation plans. Several primary and secondary questions were developed addressing different aspects of the programme. The questions and the agreed methodological approaches are summarized in Table 1 below.
23. The focus will be on documenting impacts on food security and related changes in well-being associated with households' resilience capacities. These indicators include:
 - Consumption and food security
 - Nutritional status
 - Financial outcomes and assets
 - Assets and livelihoods.
24. The evaluation will also directly assess **how** the resilience programme affects households' ability to mitigate the effects of shocks on their food security and welfare. This is mainly achieved by monitoring the following main outcomes using two-monthly surveys:
 - Food consumption
 - Coping strategies
 - Shocks and stressors experienced.
25. The impact evaluation uses a mixed-methods evaluation design, using quantitative and qualitative data. The design includes three complementary components:
26. *Clustered randomized controlled trial design*: The impact evaluation will employ a clustered randomized controlled trial (RCT) design to examine the primary evaluation question outlined in Table 1. Within the design, 91 resilience communities (sites or clusters of villages) in regions will be randomized into treatment and control groups. Participants in the treatment group will be expected to receive the integrated package of interventions, including School Feeding, Food Assistance for Assets, Nutrition/Health, and SAMS activities. The RCT will construct credible counterfactuals to identify the impact of the programme on resilience outcomes.
27. *Heterogeneity analysis*: Not all the resilience programme activities are suitable for randomized assessment. The data collected within the framework of the RCT will also be used to answer two secondary questions to be examined through this impact evaluation using heterogeneity analysis and qualitative analysis. Additional details of the evaluation methodology can be found in Table 1 below, as well as in the following sections.
28. *Qualitative analysis*: Qualitative data will be used to understand how the programme is being implemented, and how the support provided through the programme is being perceived by the beneficiaries. Additionally, the data will be used to generate additional insights about the patterns observed in the quantitative data. The qualitative data will be particularly useful to understand which aspects of the programme are being well implemented, and to identify opportunities for further improvement. Qualitative data will be collected using two main methods: focus groups of willing beneficiaries, and interviews with select village leaders. The topics for the interviews and focus group discussions will be informed by the quantitative survey data and may include: overall awareness of the programme; level of participation in the programme; perceived changes in indicators of key outcomes

of the programme; and feedback on programme implementation. The focus groups and interviews will be planned after sufficient time has passed since the start of implementation (that is, after approximately one year) to enable informed feedback on the programme to be collected.

29. All the quantitative analysis in the impact evaluation will utilize data collected through baseline surveys, endline surveys and high-frequency surveys (conducted every two months) to answer the evaluation questions. The impact evaluation involves baseline and endline data collection, enabling the team to estimate short-term and medium-term impacts (timeline presented in Section 8). Baseline data will be collected before implementation of the integrated resilience package begins. The endline data collection will occur after at least two years of implementation of activities in the treatment groups.
30. The high-frequency surveys will be implemented every two months, starting after the baseline survey, and continuing for a period of at least one year. This high-frequency data collection exercise will focus on collecting data related to food security, shocks experienced, and coping strategies. These surveys, combined with endline data, will enable the evaluation team to observe changes in food security over shorter periods of time more frequently, providing a more nuanced picture of fluctuations in food security in the context of various shocks and agricultural seasons. The high-frequency surveys will enable us to examine the characteristics of households whose food security is less stable and understand in which periods in the year households will require most support to maintain or improve their food security.
31. By virtue of the evaluation design, the data collected will be disaggregated by the gender of the respondent. Importantly, the evaluation does not consider “a ‘household” as one unit, but rather considers individuals within the households separately. As such, some components of the questionnaire are directed at female respondents of reproductive age (such as minimum dietary diversity) or children aged 6 to 23 months (such as vaccination information, and minimum acceptable diet), as well as other age categories. Additionally, the evaluation will also disaggregate households based on their gender composition, to understand how gender influences households’ access to services, income generating opportunities, and their wellbeing.

Table 1: Overview of evaluation questions and methods

Questions answered through a randomized controlled trial (RCT)	Details of evaluation methodology
<p>What is the impact of the integrated WFP resilience package (FFA, SAMS, Nutrition/Health, and school feeding) on the resilience of recipient communities and households?</p>	<p>This question is answered by comparing households in villages that participate in the integrated resilience programme with households in villages that do not participate in the programme.</p>
Questions answered through heterogeneity analysis/qi	Details of evaluation methodology
<p>How does the effectiveness of the WFP's resilience package vary depending on households' initial poverty and food security levels?</p>	<p>This question is answered by disaggregating the data collected based on socioeconomic and demographic characteristics, such as gender. This will enable us to assess social determinants of resilience outcomes, particularly when looking at patterns (such as decision-making of women and shock exposure of different social groups) across the countries under the Climate and Resilience Window.</p>
<p>Does the impact of the programme on key resilience outcomes vary between primary villages (where the FFA sites are located) and secondary villages (located further away from the sites)?</p>	<p>Each impact evaluation site has primary villages where the FFA sites are located, and secondary villages whose households would benefit from the FFA programme. The impact evaluation has identified primary and secondary villages and has included both in the sample. This will enable us to track outcomes for households from both types of villages through heterogeneity analysis.</p>
<p>Is the resilience package having any observable environmental impacts on site-level outcomes, such as vegetation indices, around the sites where FFA activities recuperate land?</p> <p>For example, do impacts differ between sites involving a water project or pond and those that involve land recuperation only?</p>	<p>This question is answered by combining household-level survey data (collected through baseline, endline and high frequency surveys_ with environmental indicators such as Normalized Difference Vegetation Index (NDVI) (that are publicly available). Adding environmental data into the analysis provides a more nuanced picture of the shocks affecting the impact evaluation areas at the time of the evaluation. This may also inform the food security patterns being observed in the survey data.</p>

4. Evaluation methodology

32. The evaluation is being conducted between 2021 and 2023. In 2021, the impact evaluation design was set up in 91 sites in the three regions of Niger.
33. As outlined in Section 3, the impact evaluation design utilizes a cluster RCT design that is complemented by heterogeneity analysis and qualitative analysis.

4.1. CLUSTER RANDOMIZED CONTROLLED TRIAL (RCT) DESIGN:

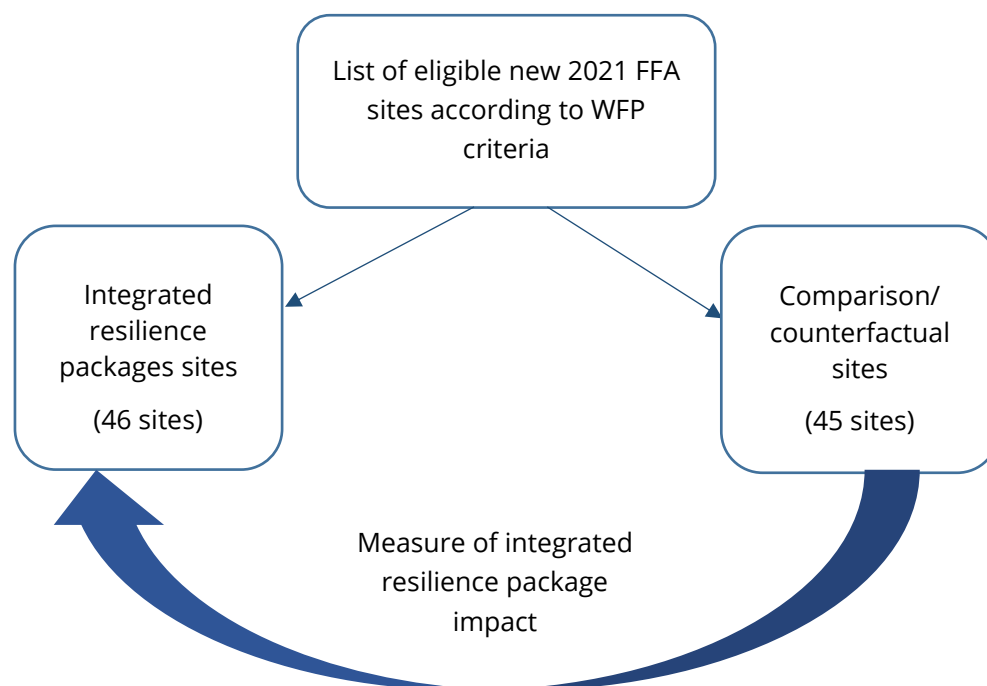
34. In a cluster RCT design, communities are randomly assigned to one of the comparison groups. The approach is depicted in Figure 3 below. The RCT analysis will compare the treatment and control groups to derive an estimate of the credible and unbiased treatment effects of the resilience package.
35. The impact evaluation design follows the implementation modalities and timelines of the WFP resilience programme. In Niger, an important entry point for programme targeting and implementation is sites or clusters of villages. Many activities critical for the programme are being implemented at site level as opposed to household- or individual-level interventions. Therefore, to identify the causal impact of the resilience programme on different comparison groups, the impact evaluation will use a clustered RCT design. In this clustered RCT design, sites will be randomly assigned to one of the comparison groups. This approach is depicted in Figure 3, below.
36. The clustered RCT will make use of a pool of asset creation sites (groups of one or more villages) that meet the criteria for programme participation and are therefore equally eligible to receive support through the programme. This ensures that villages in different comparison groups have similar core characteristics. After WFP's Niger Country Office identifies eligible sites for the resilience programme using the organization's targeting criteria, eligible sites are randomly assigned to comparison groups. Since FFA activities are the entry point for implementing the resilience programme in a village, these activities are used as the reference for randomization in this RCT.

Treatment and control groups

37. WFP pre-selected sites that meet eligibility criteria for the integrated resilience programme (including villages attached to each of these sites). The eligibility of sites was determined by the WFP resilience programme's selection process, which includes a vulnerability assessment and a technical assessment. In the eligible sites, the impact evaluation involved randomly selecting a subset of sites to receive the resilience programme from 2021.
38. As logistical difficulties and financial constraints prevented WFP from providing all eligible sites with support through the programme in 2021, the random selection is an objective and unbiased mechanism to decide which of the sites that meet the same eligibility criteria for support will receive the programme first. With a sufficient sample size, the randomization eliminates any systematic differences between the treatment and control group, and thus creates a valid counterfactual.
39. As a result of randomization, two comparison groups were created:
 - Group A: Sites receiving the integrated resilience programme in 2021 [46 sites]
 - Group B: Sites not receiving the integrated resilience programme in 2021 [45 sites]
40. The Niger WFP Country Office identified 91 sites eligible to receive the resilience programme in 2021. The Country Office had sufficient resources to support 46 of these sites in 2021. As a result, Group A (the treatment group) has 46 sites receiving the programme in 2021. Group B (the control group) has 45 sites that will not receive the programme in 2021, and this group will provide a valid counterfactual for estimating the impacts of the resilience programme.
41. This impact evaluation uses a stratified randomization approach. In this approach, first the sites were assigned into subgroups based on: 1) the communes they belong to, and 2) the average poverty level within a site. Sites are then randomly selected from each of these subgroups. This helps to ensure balance: that is, sites in comparison groups are similar on average. This will also provide some insurance against cases of non-compliance or possibly security issues that may make some sites inaccessible over time.

42. The data collected by the impact evaluation will also enable us to explore the heterogeneous effects of the programme on different groups within the population along various dimensions (such as very poor versus less poor households and male- versus female-headed households), as well as between communities (in particular, communities where the sites are located versus secondary communities located further away). We will be able to address these parts of the evaluation question by repeating the basic analysis, differentiating by category.

Figure 3: Resilience programme experimental design



4.2. HETEROGENEITY ANALYSIS

43. Not all resilience outcomes can be captured by only examining differences between treatment and comparison groups. The impact evaluation uses heterogeneity analysis to understand the secondary evaluation questions identified during the design phase. Heterogeneity analysis harnesses the cluster RCT design as a basis for identifying treatment and comparison communities. The approach allows us to compare wellbeing outcomes (e.g., food security) of household groups with different characteristics or varying levels of vulnerability. The impact evaluation employs heterogeneity analysis to examine the following interactions and outcome areas.

Resilience across varying vulnerability levels

44. Within the impact evaluation sample, the heterogeneity analysis will be used to understand how social or demographic characteristics such as livelihood types, degree of exposure to shocks, gender of the households etc. will affect resilience outcomes. Trends in food security can be observed within different sub-categories of the sample.
45. This type of analysis is made feasible by the extensive baseline survey, which collects a list of outcomes ranging from household's livelihood sources and their frequency of cultivation in a year, to their access to social safety nets and community support mechanisms. The baseline characteristics, combined with two-monthly high-frequency surveys enable us to observe the food security pathways of different sub-categories within the sample.

Impact on village-level environmental outcomes

46. It is possible that larger infrastructure construction projects (such as dam construction, large-scale land restoration and so on) may influence environmental outcomes such as rainfall density and water availability. These environmental outcomes could in turn influence village-level exposure to various

shocks or stressors. The impact evaluation team will examine Water Requirement Satisfaction Index (WRSI) data and Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) Data. The WRSI provides information about crop performance based on the availability of water at different stages of the crop cycle. CHIRPS is a data set includes more than 35 years of rainfall data from select regions across the globe. These data sets will inform the impact evaluation by providing valuable information about crop growth and climatic shocks in the impact evaluation areas. As most geographic areas under the impact evaluation are predominantly agriculture-focused, these datasets will be important for better understanding the shocks experienced by the populations at different periods during the impact evaluation.

4.3. QUALITATIVE ANALYSIS

47. In addition to the quantitative analysis, the impact evaluation will examine important process-related questions such as:
 - 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?
 - How were the intended beneficiaries supported by the programme experience participation in selected interventions?
48. In 2020 and 2021, a barrier to using additional qualitative data collection methods, such as focus group discussions, was the requirement of the Institutional Review Board (Solutions IRB was used for this impact evaluation) to limit “research activities” that increase the chance of group-based spread of COVID-19.
49. If conditions allow, two qualitative data collection activities are planned under the impact evaluation, prior to the endline survey: interviews and focus group discussions. The impact evaluation uses semi-structured interviews with implementing partners to capture information about the process of programme implementation and the experience of programme participants. The structure of the focus group discussions will be informed by the quantitative survey data and may include: overall awareness of the programme; level of participation in the programme; perceived changes in the key outcomes of the programme; and feedback on programme implementation. The focus groups and interviews will be planned after sufficient time has elapsed since the beginning of implementation (that is, approximately one year) to collect informed feedback on the programme.

5. Data collection and measurement

50. This section provides an overview of data collected by the impact evaluation, including the sample sizes and outcomes measured by household surveys.

5.1. DATA COLLECTION FOR THE RCTS

51. This impact evaluation will rely on detailed baseline and endline surveys, as well as shorter two-monthly high-frequency surveys (presented in Section 5). The baseline and endline surveys will enable us to measure outcomes before and after the intervention, and to examine whether the well-being of beneficiaries improved during the programme period. The high-frequency surveys will enable us to capture variations in outcomes, such as food security across seasons, and as households encounter shocks or stressors. The high-frequency surveys will generate additional information about which types of households need assistance and when. We present our sampling strategy for both exercises below.
52. There are two important reasons for having a sufficient sample size in an impact evaluation: 1) ensuring that households in both the treatment and the control groups are, on average, similar across the main characteristics that would influence outcomes; and 2) ensuring that the sample households in the evaluation are, on average, representative of the population they were drawn from.
53. Insufficient sample sizes generate the risk of falsely detecting, or not detecting, programme impacts. As it is logistically and financially impossible to survey the entire population, the power calculations applied for this evaluation follow standard research norms to estimate the minimum sample size needed to minimize the risk of biased estimates.
54. In Niger, each programme site comprises multiple villages at different distances from where the FFA activities occur. Therefore, the sampling for impact evaluation involved an initial listing of villages and households. This was followed by calculating approximate sample sizes required for the baseline, endline, and high-frequency surveys.

Identification and listing process:

55. Listing at village level: In total, 91 FFA work sites are included in the sample. Within these sites, 266 villages are included – 91 primary villages (where the FFA sites are located) and 175 secondary villages (further away from these FFA sites). For the impact evaluation, we sample all primary villages and one secondary village in sites with multiple villages. This gives us 158 villages across 91 sites.
56. Listing at household level: The first step of household-level sampling is identifying households eligible for receiving support through the programme, using the criteria set by WFP. Each village often has households in different poverty categories, and limited resources necessitates WFP to target the most vulnerable. To ensure that the impact evaluation follows WFP's targeting process, we undertook a door-to-door census in each of the 158 villages, leading to the identification of 23,414 households. We then drew up a sample of 10,079 households (up to 80 households per community) to be ranked by community committees in a process that mimics the Household Economic Analysis process used by WFP for targeting.²⁷ This list provided a list of households eligible for programme support. Additional information on the listing process is provided in Tables 7, 8, and 9 in Annex 4.

Baseline and endline surveys

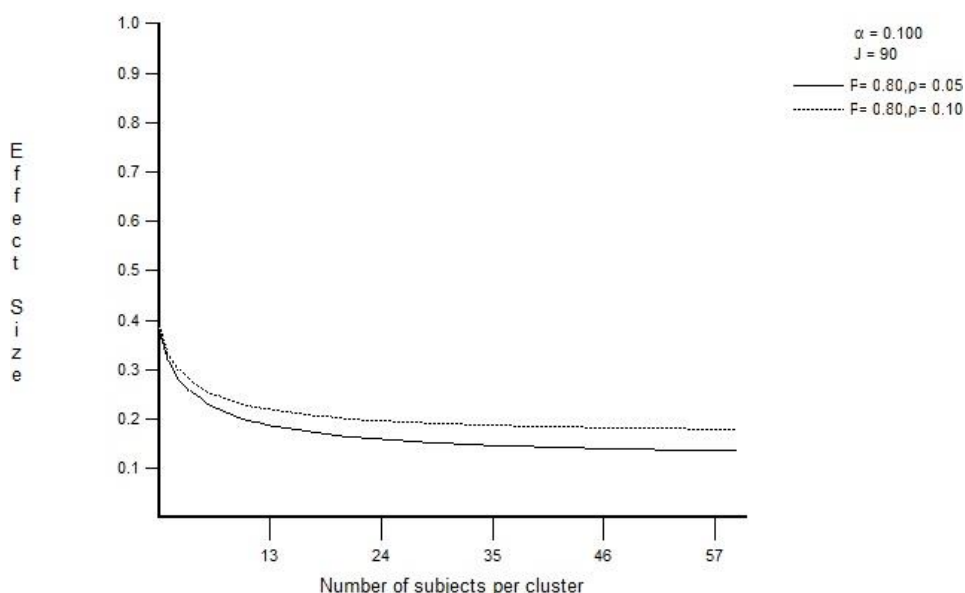
57. The sample size was established after making power calculations using optimal design. Among the households ranked, based on power calculations, we then drew up a subsample of 54 households per

²⁷ As a result of these rankings, 4,755 households (47.2 percent) were considered very poor, 3,841 (38.1 percent) poor, 1,222 (12.1 percent) moderately poor, and 225 (2.3 percent) affluent. Thirty-six households (0.3 percent) could not be ranked due to a lack of consensus in community committees. See Table 8 in Annex 4 for a breakdown.

site for interviews at baseline.²⁸ The 4,892 households selected for the baseline sample included 3,057 households from primary villages (62.5 percent) and 1,835 households from secondary villages (37.5 percent). Of the 4,892 households sampled, 4,714 were found and consented to be interviewed at baseline (a 96 percent response rate).²⁹

58. The Food Consumption Score (FCS) was used as the main outcome for power calculations, as it is a primary outcome for the impact evaluation and will be measured in all surveys. The team used FCS data collected from the Sahel for another study to perform power calculations for the baseline survey. Specifically, we calculated the number of households we need to survey in each village to detect effect sizes of 0.2 standard deviations in Food Consumption Scores (which is a reasonable minimum detectable effect for Niger as determined in the literature), with a power of 0.8. Additional factors considered in the power calculations include intra-cluster correlations (ICC) of 0.05 and 0.1, and a significance level of 0.1.
59. As Figure 4 shows, below, with 90 clusters evenly split between the treatment and control groups, we would need 22 subjects per cluster to capture an effect size of 0.2 with an intra-cluster correlation of 0.1. Eleven subjects are sufficient with an ICC of 0.05. Therefore, the evaluation requires at least 22 subjects per cluster for each subgroup (such as men and women) of interest.

Figure 4: Power calculations using optimal design, power by number of observations per cluster



60. The sample size of 54 subjects per cluster was established to be able to conduct subgroup analysis, that is, to be able to detect impact separately for (i) very poor and (ii) poor households. On average, we have 24 very poor households and 20 poor households per site. With an intra-cluster correlation of 0.1, we can identify a minimum detectable effect of 0.2 standard deviations with a power of 0.79 (among poor households) and 0.81 (among very poor households). With an intra-cluster correlation of 0.05, we

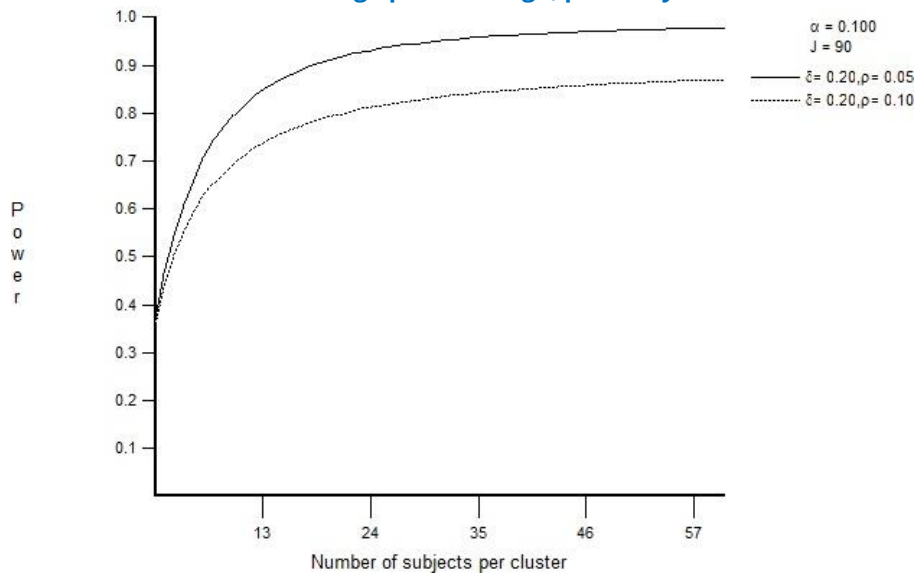
²⁸ Specifically, if there was only 1 village in a site, all 54 households were drawn from the same villages. If there is 1 village with fewer than 54 households in a site, all households in the village were drawn. If there are 2 villages drawn from the site, and both villages have more than 27 households, we draw 27 households per village. If there are 2 villages drawn from the site, but one village has fewer from 27 households, we draw all the households from that village and additional households are added from the second village to reach 54 households in the sample for the site. If there are 2 villages drawn from the site, and both have fewer than 27 households, all households are drawn into the sample.

²⁹ Of the 4,892 households drawn in the baseline sample, 2,249 households (45.9 percent) were considered very poor, 1,890 (38.6 percent) poor, 616 (12.6 percent) moderately poor, and 125 (2.6 percent) affluent. Twelve households (0.3 percent) could not be ranked due to a lack of consensus in community committees. See Table 8 in Annex 4 for a breakdown.

can identify minimum detectable effects of 0.2 standard deviations with a power of 0.91 (among poor households) and 0.93 (among very poor households).

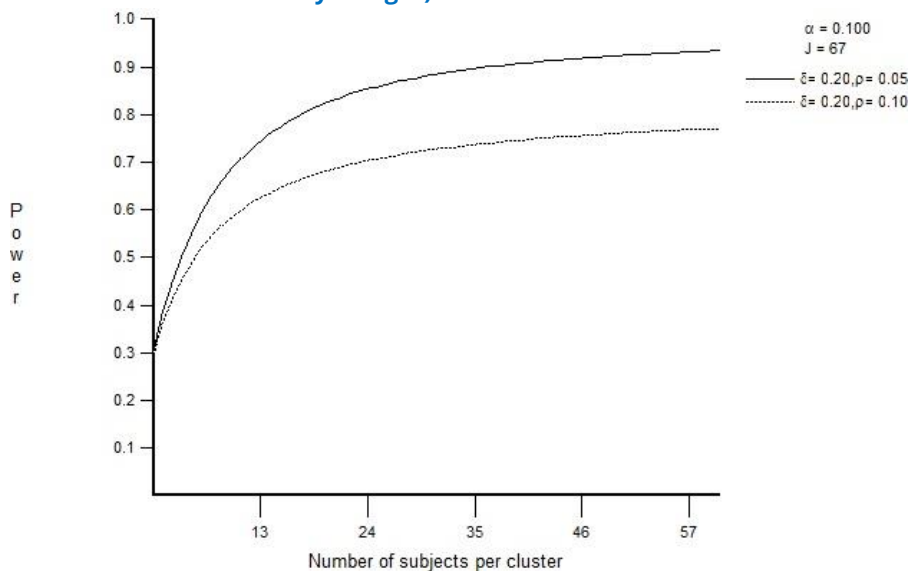
61. Similarly, the sample is powered to detect impacts separately for (i) primary and (ii) secondary villages. For primary villages, we have, on average, 32 households per primary village, which gives a power of 0.84 to detect minimum detectable effects of 0.2 standard deviations for an intra-cluster correlation of 0.1, respectively, and a power of 0.96 for an intra-cluster correlation of 0.05.

Figure 5: Power calculations using optimal design, power by number of observations per cluster



62. However, note that there are not secondary villages for all sites (only 67 sites have secondary villages). We have 20 households per site in secondary villages. We therefore have less power to detect impacts in secondary villages. Specifically, we obtain a power of 0.65 to detect minimum detectable effects of 0.2 standard deviations for an intra-cluster correlation of 0.1, respectively, and a power of 0.81 for an intra-cluster correlation of 0.05.

Figure 6: Power calculations using optimal design, power by number of observations per cluster (for 67 clusters with secondary villages)



63. The baseline report will show the actual minimum detectable effect for key variables based on the observed intra-cluster correlation coefficients at baseline.

High-frequency surveys

64. High-frequency data collection is a relatively new approach to measuring resilience in Niger, and therefore previous datasets that would inform power calculations are not available in the Niger context. Thus, we use a Madagascar dataset to assess the size of the sample required and how frequently data needs to be collected to detect reasonably small changes in outcomes, such as the Food Consumption Score.³⁰ The Madagascar data is unique because it collects three common food security indicators: Household Hunger Scale (HHS), Food Consumption Score (FCS), and Household Dietary Diversity Score (HDDS). We take these data as our starting point and assess the role of survey frequency on power to compute changes in these measures over time.
65. To make recommendations for sample size and power for the high-frequency surveys, we used data collected in Madagascar on 601 households (HHs) from 32 communities surveyed every month for 18 months. We used the first 12 of these 18 months so that we are consistent in using one full year as the relevant period.³¹ Then, using a data expansion approach, we increased the number of clusters to 90 communities and 1,600 households.
66. We model, through simulations, a hypothetical experiment that assigns half of the 90 communities to treatment. All households in treated communities experience one of three treatment effects:
- Increases in the mean of high-frequency measures by X percent of the control mean, keeping other parameters constant
 - Decreases in the standard deviation of food security measures for a household over time by X percent of baseline control SD, keeping other parameters constant
 - Decreases in the share of the year spent in poverty by X percent of the control proportion in poverty (as defined by standard thresholds for each indicator).
67. These simulations enable us to estimate the power needed for detecting the effects of programmes that may make households less food insecure on average but not change variability around that mean (variation in food security) or vice versa. For each of these effects, we replicate the hypothetical experiment with the assigned effect size for a given parameter 1,000 times, regress the measure on treatment, and calculate the proportion of the 1,000 hypothetical experiments in which we can reject the null hypothesis of no impact of treatment at the 10 percent level. This proportion is our estimate of the statistical power of an experiment with this sample size to estimate the effect. The goal of these simulations is to give guidance on how frequently countries need to collect food security data in order to identify impacts on Food Consumption Scores and other measures.
68. Power calculation tables for 90 clusters and 800–1,600 households are provided in Annex 5.
69. Table 10 in Annex 5 presents the results of power calculations needed to detect a 15 percent effect size for each of the three outcome measures with varying frequencies of data. For a 15 percent effect on power, gains in increasing frequency from two-monthly to monthly frequency are small, but the power losses in going from a quarterly to semi-annual schedule are large. We therefore focus on comparisons with the two-monthly and quarterly schedules, and compare effect sizes needed to obtain 80 percent power to guide the decisions on whether to plan for quarterly or two-monthly data collection.
70. Table 11 in Annex 5 repeats the power exercise for different effect sizes for two-monthly and quarterly schedules. We aim for two-monthly data collection for 1,600 households in 90 communities, which is sufficient to detect a 20 percent change in either the mean or standard deviation of food security at 80 percent power for all three measures. Quarterly frequency is sufficient to detect impacts on both

³⁰ Depending on the size of change expected in an outcome during each period, in this case high-frequency rounds, an evaluation may need to survey more or fewer households to detect impact on a specific outcome, such as food security. To estimate expected changes, evaluations try to use previously collected data from the same context when available (such as national surveys). In the absence of available high-frequency data on household food security from Niger, this impact evaluation uses data from Madagascar to support power calculations and estimate the sample sizes needed for high-frequency survey rounds. At the time of power calculations, the high-frequency panel from Madagascar was a rare example of the food security data required.

³¹ Outcomes such as food consumption scores are expected to vary based on seasonal changes and agricultural cycles. Therefore using 12 months of data will allow us to account for the seasonal variations in the outcome.

means and standard deviations of 20 percent of control averages with 80 percent power for only two of these three food security measures.

71. Based on these power calculations, we establish a sample size of 18 households per cluster to detect effect sizes of 0.2 standard deviations with a power of 0.8 surveyed two-monthly. We aim to have two cohorts (A and B) of 819 households (9 in each of the 91 clusters) that would be surveyed every two months using the below schedule.

Table 2: High-frequency data collection schedule

Cohort	Apr 21	May 21	Jun 21	Jul 21	Aug 21	Sep 21	Oct 21	Nov 21
A	819 HHs		819 HHs		819 Hs		819 HHs	
B		819 HHs		819 HHs		819 Hs		819 HHs

5.2. DATA USED FOR HETEROGENEITY ANALYSIS

72. The heterogeneity analysis will utilize the data collected through baseline, endline, and high frequency surveys. Outcomes such as food consumption score, nutrition, health, coping strategies and so on will be used to understand the changes in well-being among households with varying characteristics, and receiving various packages of types of programme support. Additionally, to understand the effect of physical distance from basic service points, behavioural outcomes related to vaccine take-up, school attendance, frequency of seeking medical assistance and so on will be collected in the baseline and endline surveys. Heterogeneity analysis based on access to programme activities will also utilize the data collected from implementing teams and cooperating partners on location of facilities such as schools, health centres, distribution centres, and so on.
73. Additional information required for creating sub-groups of households (such as children of school-going age) will be collected in the baseline and follow-up surveys as relevant. It is important to note that the criteria for creating the sub-groups cannot be predetermined. Therefore, until we collect this information, we will not know how many households fall into each category, or whether we will have enough households in each group to detect any differences.

5.3. OUTCOMES MEASURED

74. The resilience programme will support households' ability to cope when hit by shocks and stressors. Typically, a programme's ability to buffer against shocks is assessed by examining the interaction between the changes in outcome and exposure to shock.³² A growing body of resilience literature has relied on measuring the impacts of resilience programmes at single points in time, and documenting positive gains in well-being.³³ However, households are systematically exposed to seasonal fluctuations and shocks, such as changes in precipitation or agricultural productivity, that affect well-being over time.
75. The impact evaluation considers the fact that people who are poor today may not be the poorest tomorrow. The capacities needed to improve and sustain well-being are also likely to evolve over time, depending on the type and severity of shocks encountered. Evaluating the effect of programmes on resilience requires measuring well-being and absorptive, adaptive, and transformative capacities across seasons, and before and after shocks.

³² Gunnsteinsson et al., 2019. "Protecting Infants from Natural Disasters" *NBER Working Papers*. 35; Macours, Premand, and Vakis. 2020. *Transfers, Diversification and Household Risk Strategies*. Working Paper; Premand and Stoeffler. 2020. Do Cash Transfers Foster Resilience?". Policy Research Working Paper No. 9473. World Bank, Washington, DC

³³ Macours, K., Premand, P., & Vakis, R. 2020.

76. Building on proposals from Barrett and Conostas (2014)³⁴ and Cissé and Barrett (2018)³⁵ to conceptualize resilience as avoidance of poverty in the face of shocks and stressors, each evaluation in the Climate and Resilience Window directly measures welfare dynamics to understand resilience outcomes. These measures are calculated from a minimum set of indicators collected at higher frequencies in each country supported. Annex 2 provides additional details on the resilience measurement approach.
77. A wider range of likely outcomes is considered when answering the main evaluation questions. Annex 3 summarizes and briefly defines the key outcomes of interest for the impact evaluation in Niger.
78. The indicators were selected in collaboration with the WFP Country Office and the following three issues were considered: (i) operational relevance and importance to the programme components; (ii) a review of relevant literature; and (iii) evidence generation across the portfolio of Climate and Resilience Window evaluations.
79. The primary set of outcomes are food security indicators. These include, for example, Food Consumption Score (FCS), Food Insecurity Experience Scale (FIES), and household food consumption expenditures (measured at household and individual levels). A set of secondary outcomes will also be captured to understand the mechanisms of impacts, and other benefits beyond the immediate food security effects of the package of interventions.
80. The outcomes are measured during baseline data collection, high-frequency surveys (two-monthly surveys following the baseline), and at the endline (at least 24 months after the beginning of project implementation). A key feature of the resilience measurement approach adopted for this evaluation is reliance on high-frequency data to explore the dynamics of well-being throughout the evaluation period.

5.4. SURVEY IMPLEMENTATION

81. Data will be collected in baseline, high-frequency and endline surveys. The surveys will be identical in structure for all countries in the Climate and Resilience Window, and will be only adapted to reflect the different contexts.
 - The baseline and endline surveys are estimated to take two hours to administer, on average, for the median household.
 - High-frequency surveys are estimated to last 30 minutes, on average, for the median household. The high-frequency surveys will be implemented every month for a year. The sample will be divided into two cohorts. Each month, one cohort will be surveyed. Therefore, data will be collected for the same cohort every two months.
82. The baseline survey is administered in the first year of the programme (2021), before beneficiaries receive any of the treatments listed above. Baseline data collection activities include creating an exhaustive household listing in all participatory villages and a household classification to identify village distributions of wealth and socioeconomic layers. Baseline surveys took place before beneficiaries were registered with WFP, a few months before the first cash transfer.
83. The data collected at baseline are important to provide information about the pre-programme situation and therefore serves as a point of reference for the impact evaluation. It is also used to verify that indicators that potentially affect the main outcomes of impact (that is, food consumption, and food and nutrition security) are balanced and thus ensures that the randomization process was successful. Furthermore, baseline data provide a last-resort opportunity to assess programme impacts when there is imperfect randomization, in which endline data alone will not be sufficient to assess the programme's impact. Then, baseline data can be used to account for observable differences between treatment and control groups to assess the programme's impact.

³⁴ Barrett, C., & Conostas, M. 2014. Toward a Theory of Resilience for International Development Applications. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (40):14625-14630.

³⁵ Cissé, J., & Barrett, C. 2018. Estimating Development Resilience: A Conditional Moments-Based Approach. *Journal of Development Economics* 135:272-284.

84. We will implement the endline survey after at least two years of programme implementation, to measure changes in the outcomes of interest. We will complement these yearly rounds of data collection with high-frequency surveys that ask a smaller set of questions at more regular intervals (see Table 2).
85. A key feature of high-frequency measurements is capturing intra-annual dynamics of well-being through high-frequency food security data. This strategy enables us to better understand the resilience impacts of these interventions, not only by exploring the difference between beneficiaries and non-beneficiaries, but also by capturing the dynamics of food security throughout the study period. This will help us understand how individuals absorb shocks, adapt to changing situations and improve well-being over time. It is thus an important measurement strategy for understanding how WFP's resilience programme contributes to resilience capacities in Niger.

5.5. MANAGEMENT OF DATA QUALITY

86. Multiple steps are undertaken to ensure the high quality of data collected through the impact evaluation in Niger. Factors to consider and measures taken at each stage of the data collection are summarized below.

Questionnaire development

87. Data will be collected using multi-module household surveys covering a range of outcomes. The planned baseline and endline will be identical in structure and format. The high-frequency data collection will also follow the same format for multiple rounds of data collection throughout the study period. Questions will be repeated across surveys to enable creation of panel outcomes. The data collection instruments will be piloted extensively in each country to ensure context-specific details and option-sets for each question are appropriately identified.

Tracking participants over multiple survey rounds

88. The high-frequency and endline surveys will involve revisiting as many baseline households as possible to create a panel, enabling us to control for differences in initial levels of key outcome indicators. Information required to track households over time will be collected to allow for the possibility of revisiting some or all of the households following the second round of the survey. Collecting identifiable data is necessary to verify the identity of respondents and merge data across survey rounds. Furthermore, respondents will have to be located for subsequent survey rounds. Participants may skip these questions if they are not comfortable answering them. Relevant data protection guidelines will be followed for all data collection exercises (see the section below for additional details). To track respondents over time and construct social networks, the following direct identifiers will be recorded.

High-frequency surveys:

- Names, addresses and phone numbers of study participants.
- Names and phone numbers of alternative contacts to assist in the locating of study participants.
- GPS coordinates of respondents' households.

Baseline and detailed follow-up surveys:

- Names, addresses and phone numbers of study participants.
- Names, approximate addresses, and phone numbers for members of study participants' social networks (who will then be enrolled into the study as part of the first face-to-face follow-up survey).
- Names and phone numbers of alternative contacts to assist in the location of study participants for the next survey round. This will be collected both for the original study participants (treatment and control group) and members of their social network enrolled as part of the first face-to-face follow-up survey.

Enumerator management and training

89. In parallel with the development of the questionnaire, a suitable third-party monitoring agency has been identified to carry out data collection on the ground. The criteria for selecting these agencies in

Niger include: (i) prior experience collecting high-quality survey data; (ii) experience in Niger; and (iii) capacity to deliver multiple household surveys in the country over the study period. While the third party will be responsible for hiring the enumerators and managing them in Niger, all data collection activities will be supported by the DIME and OEV impact evaluation team. Detailed protocols have been developed to guide data collection. These have been developed by the impact evaluation team, which will also lead the enumerator training.

90. Enumerator training will include classroom and field training. Enumerators will be selected based on their performance during the training. The training is divided into three stages and will take approximately one week to complete:
- To 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 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 will discuss best practices for engaging with respondents and recording their answers with 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 a smooth transition in telephone surveys, especially when the survey will be broken up into several telephone calls.

Confidentiality of data

91. As the survey collects information about sensitive topics, strict data confidentiality protocols will be maintained throughout the evaluation. Data will be synced from the field to servers protected by passwords so that individual enumerators do not have access to the data. The data will be de-identified for analysis. Only the Principal Investigators (PIs) will have the key to link anonymized data to individually identifiable information; the PIs will consequently be responsible for ensuring the security of this key. No individual-level results will be reported, and all results will be aggregated to protect the identities of individuals.

Data quality protocols

92. Data will be collected electronically using the computer-assisted personal interviewing (CAPI) platform; CAPI surveys will reduce logical inconsistencies in the questionnaire. It also allows us to programme consistency checks into the survey and perform quality checks daily.
93. High-frequency checks look out for the following situations:
- 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.
94. Any anomalies that we detect through this process will be flagged to the data collection team immediately. In addition, the team will perform a set of back-checks (drawing a random 10–20 percent sample of households and calling them back to validate some of their answers). Cross-checking the data will allow us to provide immediate feedback to the field teams in case of divergences or other problems.

Internal team coordination

95. The evaluation process will be overseen by the project PIs, including coordination with programme counterparts, data collection, and analysis. The analysis will be performed by the project Research Assistant, with close supervision from the PIs, and it will be completed in a replicable and reproducible manner.

5.6. IMPLEMENTATION MONITORING SYSTEM

96. The WFP Niger Country Office, OEV, and DIME are working together to ensure that beneficiaries receive the scheduled WFP programming on time. The Country Office regularly tracks when transfers are made to programme recipients, as well as whether work requirements are being met. OEV and DIME are complementing these efforts by ensuring that the programme variations we introduce are properly followed. More specifically, OEV and DIME monitor treatment compliance in the following ways:
- The evaluation team ensures that the unique identifier used in the survey is aligned with the beneficiary ID used in the programme.
 - The team cross-checks periodically with field teams to ensure that initial randomization plans are adhered to. Any deviation is recorded and documented systematically to be considered during the analysis stage. The high-frequency surveys are used to track self-reported participation in programme activities. This is then cross-checked with administrative data to understand adherence to the randomization and initial programme plan. The team will support the Country Office to establish protocols to obtain relevant administrative data on programme participation.
97. The evaluation team also monitors any new activities that may be introduced into the treatment or control communities and, where possible, captures the impact of these activities through the measurement framework.

6. Data processing and analysis

6.1. DATA CODING, ENTRY, AND EDITING

98. Data will be collected by enumerators recruited by a survey firm contracted by DIME. The criteria used to select the survey firm include experience of collecting surveys in the specific impact evaluation areas, ability to hire enumerators who know the local contexts and languages, and the capacity to securely collect and manage high-quality data. All data will be collected using tablets and will be stored on SurveyCTO servers. As soon as an enumerator marks a filled-out form as "finalized", the form's contents are encrypted. Whenever form data are transmitted via a 3G or other internet network, it is encrypted in transit using SSL as well. Finally, any data downloaded from the server will be encrypted or purged of any personal identifiers before analysis.
99. Daily high-frequency checks on data quality will be implemented by the impact evaluation field coordinator and DIME research assistant, with regular reports to the Impact Evaluation (IE) teams and field teams. The high-frequency surveys will pilot methods for reaching respondents by phone or in-person visits to determine which method is more cost-effective for minimizing non-response.

6.2. PROGRAMME-SPECIFIC QUANTITATIVE DATA ANALYSIS

Resilience package

100. To measure the impact of the FFA and resilience package against the control group, our primary means of analysis is a simple regression of resilience outcomes on treatment status. A dummy variable (1/0) will be used for randomized treatment at the community level (community receives FFA/resilience or is assigned to the control group).

Descriptive targeting analysis

101. The study will document the profile of select beneficiaries along with a wide range of indicators collected at baseline. This will provide descriptive information to WFP about the efficiency of its targeting protocols, and its ability to identify the poorest households. The first rounds of high-frequency data collection may include questions about household satisfaction with targeting to shed further light on the legitimacy of targeting among local populations.

Identifying the profile of the households who benefit most from the programme

102. Finally, after follow-up data is collected, the study will help to assess how programme targeting could be improved to select the households that benefit the most from the programme. This will be based on recent statistical methods that can identify and compare the households that benefit the most from WFP assistance, before analysing whether their profiles correspond to the profile of households selected by their respective communities.

Heterogeneity to shocks

103. A feature underlying household resilience is ability to anticipate, absorb, adapt to (for example, avoid), cope with, and recover from shocks, while improving the trajectory for well-being (for example, transform). Many programmes are designed to help households mitigate the impacts of shocks, but evaluating this ability can be difficult. Typically, the ability of a programme to buffer against shocks is assessed by interacting a treatment effect with a variable measuring exposure to a shock (Gunnsteinsson et al., 2019; Macours et al., forthcoming; Premand and Stoeffler, 2020). However, evaluations that measure impact through only a baseline and endline only capture a single period of the recovery trajectory, meaning that most evaluations fail to measure the full depth of welfare costs associated with shocks, full recovery, or both. Moreover, the shocks are rarely pre-specified in experiments, meaning that the literature on shock mitigation may be vulnerable to publication bias.
104. To determine the differential impact of the programme, based on whether a household was exposed to a shock, we will run a regression interacting programme participation with a list of pre-specified context-specific shocks that will include both natural events (for example, droughts as defined by rainfall during main cultivation months falling below a defined threshold); conflict (as defined by a recorded conflict in standardized data, such as the Armed Conflict Location & Event Data Project (ACLED)); and economic shocks. In Niger, based on previous work in the region, information is collected

on shocks related to natural disasters, increases in price, family, conflict, and other miscellaneous events. The high-frequency data will enable us to run this regression for multiple points in the year, while accounting for different types and severities of shocks.

Sampling and specification

105. The sampling frame will be the lists of project sites and households as provided by the Niger WFP team. The sample will be households identified to receive benefits. Identification of recipients before implementation in all treatment arms will ensure that we can estimate intent-to-treat effects on recipient households, or likely recipient households, in pure control groups even in the event of endogenous take-up. Across all specifications, we use double-selection Lasso to select controls for precision, and we control for baseline measures of our outcomes when they are available through an analysis of covariance (ANCOVA) specification. We cluster standard errors at the community level whenever the treatment of interest is assigned at the community level. In the event of non-random attrition, we will report Lee bounds on primary impacts.

6.3. PROGRAMME-SPECIFIC QUALITATIVE DATA ANALYSIS

106. The impact evaluation will use semi-structured interviews to capture information about the programme implementation process and the experience of programme participants. The questions will be developed based on the information gaps identified through the high-frequency surveys and discussions with the Country Office. We will be collecting qualitative information relating to the implementation process as described in Section 5. We will be asking the beneficiaries if, in their view, the programme has had a positive or negative impact on outcomes. The qualitative data collected, specifically on programme implementation, will be used to support and strengthen the analysis of high-frequency survey data, and to better understand the changes in outcomes observed in the impact evaluation.

7. Ethical considerations

107. A key goal of WFP's Impact Evaluation Strategy is to increase the use of rigorous evidence in delivering interventions both in the countries directly involved in this evaluation and in other parts of the world where humanitarian and development programmes are delivered. Guided by this overarching principle, the evaluation will take into account several ethical considerations and put in place several relevant practices.

IRB Approval

108. The evaluation team obtained international approval (on 12 November 2020) from an Institutional Review Board for the Climate and Resilience Impact Evaluation Window design as well as the specific design and measurement elements in Niger. In addition, the evaluation team will also obtain approvals from local institutions in Niger (based at the Health Ministry of Niger). This will ensure that the evaluation complies with local regulations and does not violate any local laws.

Communication with participants

109. Given that the evaluation is taking place in a context of heightened inter-communal tensions and extreme vulnerability, an evaluation risk is the perception that some groups receive benefits at the expense of others solely for research. To mitigate this risk, DIME and WFP are working together to ensure transparent and clear communication to communities.

Informed consent

110. The evaluation and survey teams will ensure that enumerators are fully trained to obtain informed oral consent from all evaluation participants. Every participant must consent to take part in our surveys. We are very explicit that refusal to respond to our survey does not come with any consequences for their participation in WFP's resilience programming. The head of the household is the primary respondent for the survey. While most survey questions are addressed to the head of the household, a few questions may be directed to other household members, including women (such as questions on women's empowerment and food consumption for children aged from 6–23 months). To avoid respondent discomfort during surveys, we will take precautions to ensure that questions are being asked bearing in mind the privacy and comfort of respondents:

- Participants may skip any questions they do not wish to answer, or withdraw from the survey at any time.
- Interviews will be conducted at the participants' homes to help them to be comfortable answering questions.

111. Finally, all enumerators will complete training lasting 1–2 weeks. Following the training, the surveys will be piloted in the impact evaluation areas. The goals of the training are to ensure that enumerators follow survey best practices in terms of protocols and ethics, and that questions are asked in a uniform and contextually appropriate manner.

Confidentiality

112. The evaluation team will ensure complete anonymity and confidentiality of all data collected from study participants. This means that the identity of study participants will remain hidden in all forms of data construction and analysis, and sensitive information will not be shared with anyone outside the evaluation team.

Transparency in evaluation designs

113. To increase the transparency of the work, the evaluation is registered through the American Economics Association's (AEA's) trial registry.

Considerations for rewarding participation

114. The evaluation team considered providing small cash transfers to participants in the high-frequency survey. However, following discussion with the Country Office, it was decided it would be preferable to provide small in-kind payments (such as soap), which will be provided for each round of high-frequency data collection.

8. Risks and limitations

Risks to internal validity

115. The primary risks to an internally valid, causal estimate of programme impact, in combination and in isolation, are statistical power for estimating multiple treatment arms with limited scope for ex ante identification of programme sites/participants combined with rigorous methods to create a counterfactual, such as phase-in randomization. When sites and participants have already been selected, options to create counterfactuals are limited, and with a small number of planned expansion sites, not all possible treatment arms can be implemented simultaneously. To address this problem, we use the most rigorous impact evaluation method available – a randomized controlled trial. In addition, as with any in-field RCT, spillover across communities and differential attrition are potential risks for the evaluation. However, the team will work closely with the implementing partners on the ground to monitor potential spillover risks and design clear and direct implementation protocols. We expect differential attrition to be less common than in other contexts, as the control group has been made aware that they will receive the FFA intervention in the programme's second year.
116. Another related risk is low response rates due to respondent fatigue. The provision of small in-kind payments (such as soap) will mitigate this risk, along with close monitoring of response rates and data quality. Previous work in Niger has shown that differential attrition is not a major concern.

Risks to external validity

117. One of the evaluation's limitations may be that the results of a single study might not be externally valid. However, we can indicate the robustness of our findings in external contexts through the formal synthesis of findings from all the countries that participate in the Climate and Resilience Window. Furthermore, comparing the findings to a similar ongoing study in Mali will be particularly interesting as there may be more similarities in programme mechanisms and impacts within the same region.
118. The use of coordinated survey instruments and data collection protocols will help to ensure that the data collected from Niger will be comparable to other countries in the window and in other WFP-supported evaluation windows, to maximize the potential for externally valid inferences.

Risks due to instability

119. A further risk is that a crisis (such as conflict, political instability or a 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 by supporting the planning and implementation of operational activities and the timely launch of procurement processes. Furthermore, field coordinators will work closely with DIME, WFP, and the implementing partners to ensure that programme activities are being carried out according to the planned standards and protocols, and to alert the evaluation team in a timely fashion about deviations and other implementation challenges.

Risks due to COVID

120. As a result of COVID-19, in the early phase of the impact evaluation, the Country Office implemented all its programmes with third-party NGOs that are now responsible for all field-related activities. This creates additional monitoring challenges as the evaluation team must ensure that the NGOs comply with the original design (registering dual-headed households, respecting the randomization of communities to treatment arms, and delivering cash and assets on time). The evaluation team has developed a strong working relationship with the Country Office and frequently communicates with the Country Office and the NGOs to monitor these dynamics.
121. In addition, traditional in-person surveys may become difficult to implement if national authorities require social distancing. Survey activities will comply with national policies and, if social distancing is required, alternative means of data collection, such as remote surveys by telephone or similar, will be used. These forms of data collection are currently being explored by WFP in multiple countries.

Risks due to low programme participation

122. In contexts of insecurity, participation in programme activities may be reduced due to difficulties accessing activity areas (such as, asset creation sites and nutrition centres). It is also possible that beneficiaries find alternative income sources that are more suitable for their needs and choose not to participate in the resilience programme. If programme participation is low, it is difficult to detect statistically significant effects of the programme based on the original survey sample. It is also crucial to detect any reduction in participation early in the evaluation, and try to identify possible reasons for this.

123. The evaluation team is working closely with the Country Office to establish programme monitoring systems to closely track participation in various programme activities. Data obtained through administrative sources will be cross-checked with self-reported programme participation through high-frequency surveys to better understand programme implementation progress.

9. Quality assurance and peer review

124. WFP's Impact Evaluation Quality Assurance System (IEQAS) provides 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 materials to guide evaluation teams and partners throughout the evaluation process. Quality assurance measures will be systematically applied throughout the evaluation phases. These include preparation and selection, design, data collection,³⁶ consistency of programme implementation with the evaluation design, analysis, and reporting.
125. Climate and Resilience Window pre-analysis plans, which include each country using a similar impact evaluation design, are reviewed by the Steering Committee and Technical Advisory Group, and by external quality support peer-reviewers before registration. Following registration, country-specific evaluation reports published by WFP – including inception, baseline, and final reports – are prepared by the evaluation team. All country-specific evaluation reports are reviewed by the Evaluation Committee (see Table 5) and shared with the window's Steering Committee for comments. Final evaluation reports are also reviewed by external peer reviewers. In addition to WFP-published reports, the impact evaluation team produces a window-level meta-analysis and peer reviewed journal articles. All reports and articles are reviewed by the Head of Impact Evaluation.
126. The WFP Director of Evaluation approves all the reports before they are 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 managed by the Office of Evaluation. The overall rating category of the reports will be made public alongside the evaluation reports.

³⁶ This includes using high-frequency data quality checks routinely throughout the data collection phases, and ensuring that the baseline and endline reports adhere to predesignated standards set by the Office of Evaluation.

10. Communication Plan

127. In Niger, the WFP integrated resilience programme is intended to address chronic food insecurity and to support communities to respond and adapt to climatic shocks and conflicts. The evidence generated from the impact evaluation will inform the programme's future scale-up or expansion. The impact evaluation will also provide insights on the most effective approaches for targeting the most vulnerable and providing support at the most effective times.

128. More broadly, the impact evaluation evidence will also contribute to the planning of the next Country Strategic Plan for WFP in Niger by supporting the Country Office to identify what activities or *combination of activities* have the greatest impact on resilience, and how this varies across regions. Finally, WFP's resilience programme in Niger is part of the regional WFP resilience initiative, and the impact evaluation will contribute to resilience learning in the region, as well as to the development of a regional data ecosystem.

129. Considering these objectives, the impact evaluation team developed a communication plan to ensure timely dissemination of the evidence, and to allow for its use in programme design and delivery. DIME and WFP will ensure that the WFP Regional Bureau and country offices are full partners in discussing and using the evidence created by the impact evaluation. More specifically, there are three complementary avenues envisioned for dissemination:

- **Active engagement with programme teams:** This includes evaluation workshops and country-level engagements. On completion of the evaluation phases, we will work closely with all stakeholders to elaborate relevant and visually appealing policy briefs, social media communications, and dissemination events. The evaluation team will produce a report to be shared with operational teams and policymakers in each country to summarize learning, solicit suggestions and improvements, and generate new uses for the resulting data.
- **Harnessing WFP's and DIME's global networks:** The impact evaluation team is working closely with many different stakeholders in the development arena. The network brings together governments, donors, and academics. Evaluation results will be disseminated widely across the community of practice through workshops. In addition, we plan to make our findings broadly available to other World Bank teams and cash transfer/social protection-related projects to emphasize the role of community targeting of public goods. DIME hosts or participates in multiple workshops each year on using impact evaluations to improve project learning. Each of these workshops will be an opportunity to share evaluation findings and lessons with other agriculture and rural development projects from both within and outside the World Bank and WFP.
- **Academic publications:** Papers are planned based on the RCT in Niger. The impact evaluation team will engage broader academic communities to both contribute to and shape the knowledge generated by the impact evaluation. All data collected as part of the set of evaluations will be made available online through the IE database, following the World Bank's open data policy.

11. Organization of the evaluation

130. The impact evaluation will be delivered through a partnership between WFP and the World Bank's Development Impact Evaluation Department (DIME). DIME and WFP will deliver the impact evaluation through the existing Memorandum of Understanding between the Office of Evaluation of WFP and the World Bank. Key governing and management structures within the partnership are outlined below.

EVALUATION TEAM

131. The evaluation team will consist of Principal Investigators, and focal points from the World Bank DIME and WFP. The composition of the team is summarized in table 3 below.

132. The responsibilities of the evaluation team include:

- Preparation of the impact evaluation concept note and work plan
- Delivery of all activities set out in the impact evaluation workplan
- Monitoring and reporting on the progress made in delivering the workplan to the evaluation Steering Committee.
- Preparing Annual Progress Reports

Table 3: Evaluation team and main counterparts

NAME	ROLE	ORGANIZATION/UNIT
Patrick Premand	Principal Investigator (Lead Researcher)	World Bank/DIME
Guigonan Serge Adjognon	Co-Principal Investigator	World Bank/DIME
Marcus Holmlund	Resilience Window Manager	World Bank/DIME
Chloë Fernandez	Research Officer	World Bank/DIME
Jonas Heirman	Evaluation Officer	WFP/OEV
Hanna Paulose	Evaluation Officer	WFP/OEV
Ola Eltoukhi	Evaluation Analyst	WFP/OEV
Pulkit Bajpai	Research Assistant	World Bank/DIME
Abdoul-Aziz Adama	Field Coordinator	World Bank/DIME

EVALUATION COMMITTEE

133. The evaluation committee will include the evaluation field coordinator, and representatives WFP Niger country office Sudan and the Regional Bureau in Dakar. The committee will be responsible for monitoring the progress and advising on broad strategic issues at each stage of the impact evaluation (e.g., concept notes, data collection, reviewing reports etc.) The committee will meet annually or on the side-lines of the evaluation learning workshops. Key members of the evaluation committee are listed in Table 4 below.

Table 4: Evaluation committee

NAME	ROLE	ORGANIZATION/UNIT
Koffi Akakpo	Head of Programme	WFP Niger
Anna Law	Head of Research, Assessments and Monitoring (RAM)	WFP Niger
Claudia Schwarze	Regional Evaluation Officer	WFP Regional Office in Dakar
Edoxi Kindane	Evaluation Officer	WFP Regional Office in Dakar
Abdoul Aziz Adama	Field Coordinator	DIME

WORK PLAN AND DELIVERABLES

Table 5: Milestones, deliverables, and estimated timeline

MILESTONES	DELIVERABLES	TIMELINE
Agreement on the IE design	Concept Note	August 2020
Data collection plan and pilot	TORs Questionnaires	November 2020–February 2021
Data collection (baseline)	Cleaned data Dictionaries	February–September 2021
First data analysis	Presentation Data file Do files Baseline report	November 2021–January 2022
Implementation of intervention aligned to evaluation	Rollout plan Monitoring reports verifying treatment and control status	Starting February 2021
High-frequency survey data collection plan	TORs Questionnaire Sampling plan	March 2021
Data collection (high-frequency surveys)	Cleaned data dictionaries Datasets	Two-monthly April 2021–August 2022
Draft Inception Report	Peer Reviewed Methodology Note	October 2020
Final Inception Report	Published on WFP.org	September 2022
Follow-up data collection plan	TORs Questionnaire	September 2022
Baseline Report	Report published on WFP.org	October 2022
Data collection (follow-up)	Cleaned data Dictionaries	July 2023/aligned with the programme timeline
Final report and policy notes	Technical note Policy note Data file Do files	August 2023/aligned with programme timeline
Dissemination of findings	Presentations	December 2023/aligned with programme timeline

Annexes

ANNEX 1: WINDOW SUMMARY

134. The Climate and Resilience Impact Evaluation Window was developed by WFP's Office of Evaluation (OEV) in partnership with WFP's Asset Creation and Livelihoods Unit, and Climate and Disaster Risk Reduction Unit, as well as the World Bank's Development Impact Evaluation (DIME) department. The Window will coordinate a portfolio of impact evaluations to measure the impact of WFP's resilience programmes on household resilience across a set of countries.
135. Windows are coordinated portfolios of impact evaluations on a specific evidence area – in this case, climate and resilience. Windows are part of WFP's impact evaluation strategy and are coordinated by WFP's Office of Evaluation and the Development Impact Evaluation (DIME) department of the World Bank. They inform country offices about the effectiveness of their programming and contribute to a global evidence base by examining similar questions about the same concept in multiple programming contexts. Through a coordinated portfolio of impact evaluations, windows aim to increase the power of evidence generated and expand its ability to be generalized across contexts, thereby improving external validity.
136. The concept of resilience has gained attention in the development and humanitarian sectors because it recognizes that a household's well-being depends on social, economic, human, and environmental capital, as well as exposure to – and ability to cope with – shocks and stressors. Therefore, it is centred around addressing shorter-term humanitarian needs while simultaneously supporting communities to face future crises induced by climate change, conflict, and other factors. Many institutions, including the United Nations Children's Fund (UNICEF), the World Bank, and the World Food Programme (WFP), have increasingly used the concept as a basis for their programming.
137. WFP's resilience policy uses the definition of the Technical Working Group of the Food Security Information Network (FSIN) for its resilience programming, which defines resilience as “the capacity to ensure that shocks and stressors do not have long-lasting adverse development consequences”.³⁷ The set of capacities required before, during, and after the onset of shocks and stressors supports the ability to: (i) absorb shocks and stressors; (ii) adapt to change through making proactive choices; and (iii) transform, thus changing the available choices. The capacities contribute to maintaining development gains during shocks and stressors.
138. To strengthen resilience, WFP employs an integrated approach to programming, in which multiple forms of support are provided to the same community. These integrated packages of interventions aim to improve food security and nutrition by smoothing and improving food consumption in the short term, while supporting livelihoods and addressing barriers to development (e.g., better climate information, access to markets, education, WASH, and community ownership and leadership) in the long term. Rigorous evidence is needed on how these interventions contribute to resilience in order to design programmes that simultaneously address the root causes of food insecurity and malnutrition while meeting immediate food needs.
139. The Climate and Resilience Impact Evaluation Window will support resilience programme teams in designing impact evaluations to understand how the integrated packages of interventions, and activities within the package, contribute to resilience. Currently, resilience impact evaluations in four countries (Mali, Niger, Rwanda and South Sudan) are part of the window, with the expectation that at least one more impact evaluation will be added.
140. Each window is guided by one or more pre-analysis plan(s) (PAP). The first Climate and Resilience Window PAP describes how evaluations will estimate the impacts on resilience of experimentally varying livelihoods, education, health, and complementary activities. Resilience is measured in the window through baseline, endline, and high-frequency surveys that capture changes in household well-being, defined in terms of food consumption and food and nutrition security. Climate and Resilience

³⁷ FSIN Resilience Measurement Technical Working Group. 2014. Resilience Measurement Principles: Toward an Agenda for Measurement Design. Rome, FAO & WFP.

Window impact evaluations also examine the timing and sequencing of activities, as well as their targeting modalities, to understand if, and how, programme designs can be most effective.

ANNEX 2: DEFINING RESILIENCE

141. This annex describes how we plan to conceptualize resilience by measuring dynamic outcomes, such as food security, school attendance, and labour outcomes. The material in this annex is closely adapted from the WFP climate and resilience pre-analysis plan.

Defining resilience through high-frequency measurement

142. Measurement of resilience has mostly taken one of three approaches in the literature. The first is to define ex ante characteristics of households that are expected to be associated with lower resilience, and construct a “resilience index”. This is the approach of the Food and Agriculture Organization’s (FAO’s) Resilience Index Measurement and Analysis (RIMA) tool or the Technical Assistance to NGOs (TANGO) resilience index, as well as examples of resilience evaluations that use characteristics like diversification of livelihood strategies as a proxy for resilience. The second is to regress outcomes on measures of shocks to isolate the contribution of shocks to food security. The third is to use the measurement of different households’ food security at different times to impute a given household’s food security path and then measure the parameters of the imputed distribution.

143. Our measurement framework extends these existing imputation-based measures of food security dynamics by allowing idiosyncratic shocks that are not shared across households. The measures of interest are closely related to proposed measures of vulnerability, but we aim to measure underlying consumption smoothing behaviour rather than the welfare consequences of such behaviour. Resilience is best described not by a single index, but by the following simple structural equation for household welfare:

$$y_{it} = \alpha_i + f_i(d) + \delta it + \epsilon_{it} \quad y_{it} = \alpha_i + f_i(d) + \delta it + \epsilon_{it}$$

where y_{it} is a measure of well-being, such as aggregate consumer expenditure, food security, or poverty status, for an observation unit i at time t . Since the programmes included in the study primarily focus on improving food security and nutrition outcomes, selected food security indicators will be used as measures of well-being. The four components of this equation determine a household’s ability to avoid food insecurity over time and can be estimated as a regression of household food security on time and survey dates. To understand this equation, imagine using this framework to estimate a household’s level of resilience. Specifically, α_i , the household-specific fixed effect, measures a household’s reference level of food security. The second term is a function of the calendar date on which food security is measured, and measures seasonality. The third term is a trend measuring how quickly a household is improving food security over time t . Finally, ϵ_{it} measures exposure to shocks not systematically correlated with survey dates. Figure 9 shows how this looks in a plot, where we measure a household’s consumption or food security status in every period from $t = 0$ to some period $t = T$.

144. Impact evaluations typically focus on measuring a household’s consumption at one point in time, with the view that a single observation is a sufficient statistic for that household’s reference level of well-being for a given year. In panel A, the red and blue households differ only in their value of α . The household whose consumption is depicted by the red line is always “more food insecure” than the household whose consumption trajectory is shown by the blue line, meaning that for any given food security threshold, the blue household will be food insecure if, and only if, the red household is also food insecure.

145. However, the average food security of the household over the period (α_i) only captures one feature of the consumption function that is important for welfare analysis. The blue household in panel B has a steeper δ , indicating a steeper trend in food security, meaning that this household will move above the poverty line and/or farther away from it. The blue household in panel C has a seasonal pattern with greater variability than the household with a red line. Seasonality could lead to households falling below a food security threshold in the lean season. In panel D, both the red and blue household experience a shock at the same point.

146. Given the structure of the equation of motion for consumption above, each component could be estimated if data were collected every day from $t=0$ to T . However, such data are virtually impossible to

collect and may not be necessary to distinguish impacts arising from influencing different components of the well-being equation. We propose to operationalize resilience measurement by repeated sampling of the same household on different dates within a pre-defined period and estimating key household-specific parameters of the structural consumption equation from this sample of consumption at different dates.

Operationalizing feasible measures of resilience

147. These impact evaluations will estimate welfare trajectories within a one-year period following the start of a programme. Figure 9 shows a hypothetical consumption path for a household in a period $t = 0 \dots T$. The dynamics shown could represent either a seasonal consumption path with one lean season and one peak season, or a household that experiences one positive and one negative shock.

148. The first measure of the consumption equation we are concerned with is the household's intra-annual reference level of consumption – this is α_i in the structural equation. If we observed a household's value of consumption every day, this would be measured as a household's average food security status over the period – as shown by m in Figure 10 (Panel A). Next, we consider the household's intra-annual standard deviation, the average of the household's deviations from the reference mean (Figure 10 – Panel B). The standard deviation captures the combined influence of both $f(d)$ and (ϵ) on household welfare trajectories. This single indicator summarizes the variability associated with both seasonality and shocks within the period. The third measure is the time trend. However, by limiting the comparison within a year, we do not consider a year-on-year trend in welfare. The final measure we consider is the share of the period the household spends below a poverty line or food security range. This is the number of days covered below the poverty line divided by the total number of days in the period of interest (Figure 10 – Panel C). Resilience is then defined as the ability of a household to avoid poverty over time, which we operationalize in the following ways:

- A household with a higher m is on average higher above or less below the food security threshold. So, households with higher m are more resilient than households with lower m . The intra-annual reference mean of food security is measured by $m_i^{\wedge} = \frac{1}{n_i} \sum_{t=0}^T y_{it}$ $m_i^{\wedge} = \frac{1}{n_i} \sum_{t=0}^T y_{it}$
- Conditional on m , having a higher standard deviation will increase the likelihood of falling below a food security threshold, the share of time spent below the poverty threshold, and/or the number of days that are relatively far below the food security threshold. Conditional on m , households with a higher standard deviation are less resilient. The intra-annual reference standard deviation of food security is measured by $s_i^{\wedge} = \frac{1}{n_i} \sqrt{\sum_{t=0}^T (y_{it} - m_i)^2}$ $s_i^{\wedge} = \frac{1}{n_i} \sqrt{\sum_{t=0}^T (y_{it} - m_i)^2}$
- Households that spend more time below the threshold are less resilient than households that spend less time below the line. The share of observations below a poverty line is measured by

$$share_i^{\wedge} = \frac{1}{n_i} \sum_{t=1}^T \mathbb{1}(y_{it} < y^-)$$

where n_i is the number of times a community, household, or individual i is surveyed; T is the length of the period over which resilience is measured, y_{it} is a measure of household food security status, and y^- is a threshold below which a unit is considered poor or food insecure. These three measures, defined for a selected set of food security indicators, will be our main welfare outcomes. Below, we consider power and describe how frequently we need to measure outcomes to detect changes in these outcomes associated with interventions.

149. Figure 11 shows what the measures look like for the household with the hypothetical sinusoid function shown so far, assuming a quarterly data collection schedule in which food security status is observed at three-month intervals. For this household, the reference level of consumption m (shown by the red dashed line) is simply the average of the four points. The intra-annual standard deviation is estimated by calculating the standard deviation of the four points, the average of the solid red lines. The range is the difference between the highest of the four values and the lowest, the difference between the dashed black lines. And the share of the period spent below the poverty line is the number of observations that fall below the poverty line (the grey dashed line) divided by the total number of observations (i.e., the number of grey dots divided by the number of blue dots).

Figure 7: Examples of capacities over time

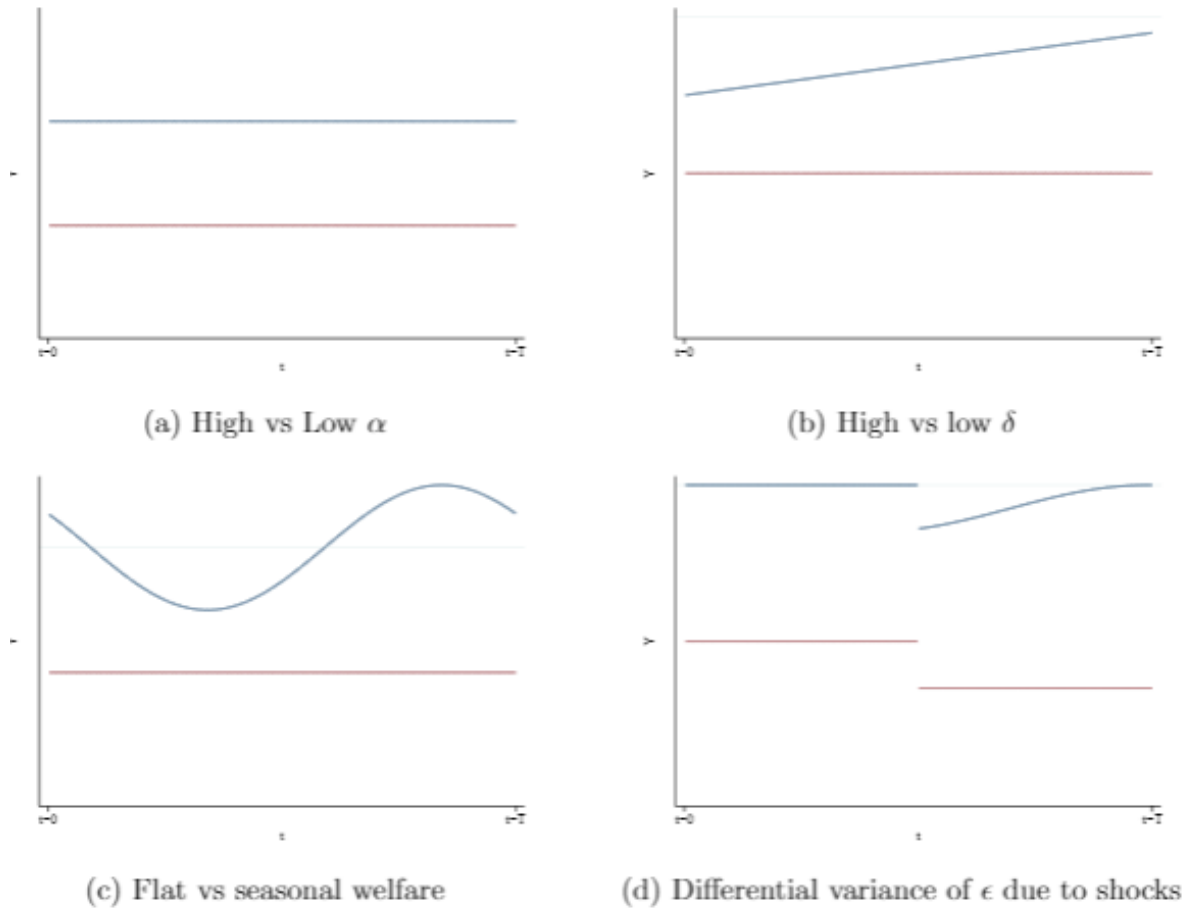


Figure 8: Measures of capacities

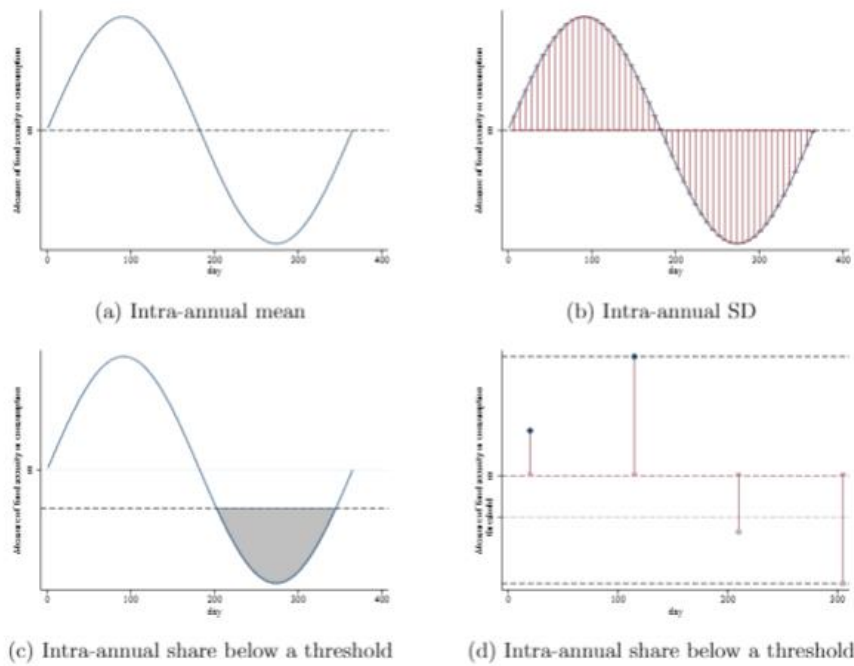
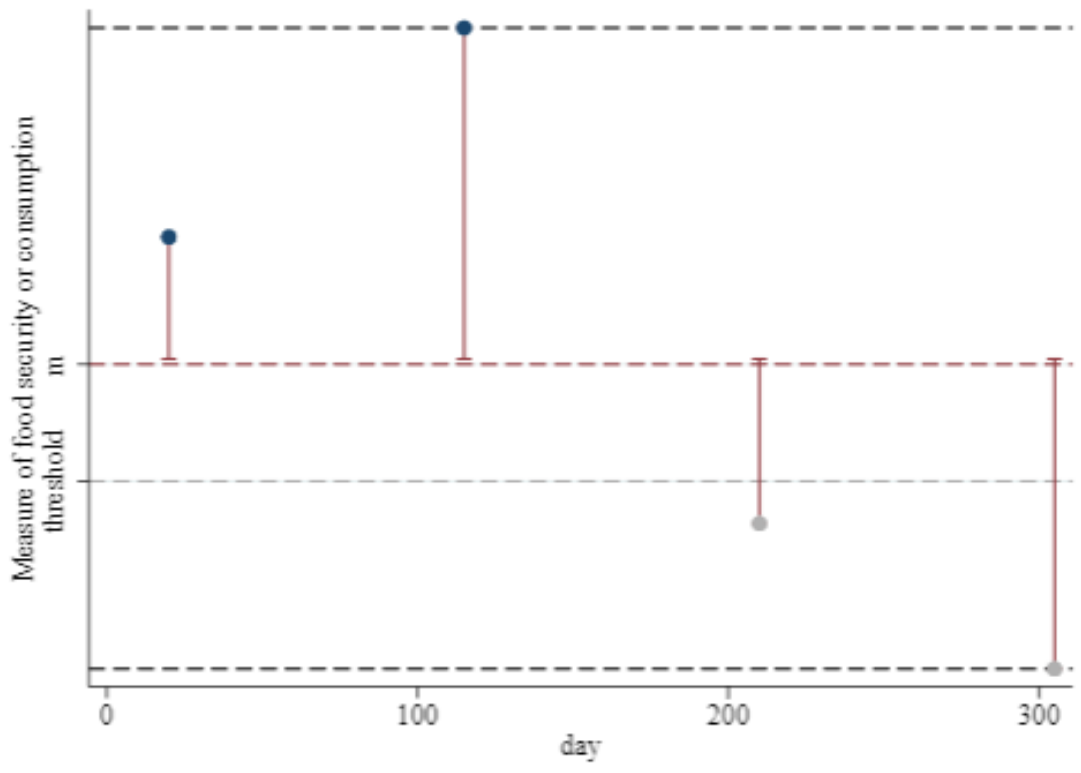


Figure 9: Feasible measurement of capacities



ANNEX 3: MAIN OUTCOMES OF INTEREST

Table 6. Main outcomes of interest

Outcome type	Outcome name	Definition	Measurement level	Source
Primary	Consumption and food security	FCS/FIES/ expenditure.	Household/ individual	Baseline, endline, and high-frequency surveys
Secondary	Time use	Activities and time spent at points of the day for selected household members	Household/ individual	Baseline, endline, and high-frequency surveys
Secondary	Assets	Number and value of assets owned by the household from a contextually pre-defined list.	Household	Baseline and endline surveys
Secondary	Income-generating activities	Participation in non-farm business, agriculture and livestock, or wage employment and revenue from these activities.	Household/ individual	Baseline, endline, and high-frequency surveys
Secondary	Shocks and coping mechanisms	Shocks encountered by the household, including the severity of shocks, and coping strategies used. Selection of shocks from a pre-defined list.	Household/ individual	Baseline, endline, and high-frequency surveys
Secondary	Financial outcomes	Current savings levels, the number of loans they have taken and their current outstanding debt, insurance products currently owned, and cash transfers undertaken, including remittances.	Household/ individual	Baseline and endline surveys
Secondary	Migration	Migration of household members	Household/ individual	Baseline and endline surveys
Secondary	Psychosocial well-being	Stress, life satisfaction, self-efficacy, aspirations, Center for Epidemiological Studies Depression scale.	Household/ individual	Baseline and endline surveys
Secondary	Women's empowerment	As defined by CBT/Gender window (normative and positive time use and agency).	Household/ individual	Baseline and endline surveys

Secondary	Social capital	Social cohesion, closeness of community index, financial support index, collective action index.	Household/ individual	Baseline and endline surveys
Secondary	Safety nets	Amount and source of transfers from other NGO and gov. sources	Household/ individual	Baseline and endline surveys
Secondary	Reservation wages	The minimum hourly wage selected household members would accept to engage in short-term labour and the duration they would be willing to work.	Household/ individual	Baseline, endline, and high-frequency surveys

ANNEX 4: VILLAGE AND HOUSEHOLD LISTING

Table 7: Number of households listed, ranked, and interviewed at baseline – across all 91 *treatment and control* sites (and the 158 sampled villages associated with these sites)

Regions	# of HH listed	# HH ranked	#HH interviewed at baseline
Diffa	1,453	1,315	936
Dosso	3,139	1,551	595
Tahoua	18,822	7,213	3,183
Total	23,414	10,079	4,714

Table 8: Number of households listed, ranked, and interviewed at baseline – across the 46 *treatment* sites (and the 80 sampled villages associated with these sites)

Regions	# of HH listed	# HH ranked	# HH interviewed at baseline
Diffa	643	566	436
Dosso	1,641	948	315
Tahoua	9,629	3,741	1,612
Total	11,913	5,255	2,363

Table 9: Number of households by poverty status, for all households ranked and for households in the baseline sample (treatment and control sites).

Regions	# of HH ranked	Ranking distribution	# HH interviewed at baseline	Baseline distribution
Very poor	4,755	47.2%	2,249	45.9%
Poor	3,841	38.1%	1,890	38.6%
Moderate	1,222	12.1%	616	12.6%
Affluent	225	2.3%	125	2.6%

No consensus	36	0.3%	12	0.3%
Total	10,079	100%	4,892	100%

ANNEX 5: POWER CALCULATIONS

Table 10: Power calculations to detect a 15 percent effect size

FCS			
Frequency	Mean	SD	Share of obs < threshold
Monthly	0.997	0.932	0.526
Bi-monthly	0.996	0.884	0.420
Quarterly	0.958	0.712	0.365
Semi-annually	0.962	0.387	0.302

HDDS			
Frequency	Mean	SD	Share of obs < threshold
Monthly	0.991	0.925	0.861
Bi-monthly	0.973	0.847	0.713
Quarterly	0.972	0.700	0.802
Semi-annually	0.948	0.455	0.531

HHS			
Frequency	Mean	SD	Share of obs < threshold
Monthly	0.421	0.832	0.330
Bi-monthly	0.394	0.697	0.307
Quarterly	0.372	0.537	0.285
Semi-annually	0.262	0.268	0.245

Table 11: Power calculations to detect different effect sizes for two-monthly and quarterly schedules

FCS							
Bimonthly				Quarterly			
Effect size	Mean	SD	Share of obs < threshold	Effect size	Mean	SD	Share of obs < threshold
15%	0.996	0.884	0.420	15%	0.958	0.712	0.365
20%	1.000	0.990	0.600	20%	1.000	0.922	0.546
25%	1.000	1.000	0.777	25%	1.000	0.984	0.683
30%	1.000	1.000	0.881	30%	1.000	1.000	0.851

HDDS							
Bimonthly				Quarterly			
Effect size	Mean	SD	Share of obs < threshold	Effect size	Mean	SD	Share of obs < threshold
15%	0.973	0.847	0.713	15%	0.972	0.700	0.802
20%	0.999	0.973	0.915	20%	1.000	0.892	0.968
25%	1.000	0.998	0.992	25%	1.000	0.971	0.997
30%	1.000	1.000	1.000	30%	1.000	0.996	1.000

HHS							
Bimonthly				Quarterly			
Effect size	Mean	SD	Share of obs < threshold	Effect size	Mean	SD	Share of obs < threshold
15%	0.394	0.697	0.307	15%	0.372	0.537	0.285
20%	0.523	0.892	0.423	20%	0.480	0.752	0.362
25%	0.603	0.982	0.527	25%	0.548	0.859	0.436
30%	0.725	0.992	0.654	30%	0.653	0.940	0.563

ANNEX 6: QUESTIONNAIRES

A link to the baseline questionnaire is available [here](#). The modules included in the survey are summarized in the table below.

Table 12: List of modules included in the baseline questionnaire

Module	Description
A	Introduction
B	Consent
C	HH roster
D	Education & employment
E	Income Generating Activities - non-ag business
F	Income Generating Activities - ag & livestock
G	Food Consumption Score (FCS)
H	Food Insecurity Experience Scale (FIES)
I	Consumption expenditure (food & non-food)
J	Asset Index + Access to Basic Services
K	Psychosocial & Mental Health
L	Shocks
M	Coping Strategies
N	Migration
O	Financial outcomes (savings, loans, insurance, cash transfers)
P	Time-use
Q	Safety Nets
R	Social Capital
S	Women's Empowerment
T	Women's Dietary Diversity
U	Child Health
	End of Survey

ANNEX 7: DETAILED STAKEHOLDER ANALYSIS

150. Stakeholders and users of this evaluation are defined as those actors who may influence the evaluation, and those who may be influenced by it. This includes internal, external, and national actors and programme beneficiaries. The WFP country office in Niger is intended to be the primary user of this evaluation. In addition, the WFP Niger office provides technical guidance at the national level to inform national policy and dialogue on social protection, and the country office has expressed interest in using the results of this evaluation to support this technical advisory capacity.

151. The various categories of stakeholders include:

- Internal Niger-based stakeholders: the Country Director and Deputy Director, the Head of Programme, and all technical and management personnel.
- Internal stakeholders outside Niger: the Office of Evaluation, the Regional Bureau for Central and Western Africa in Dakar, and the Resilience Divisions and WFP.
- Population groups in need (affected populations): resident communities and migrants of different sexes and age groups.
- External stakeholders, including international non-governmental organizations, donors, United Nations agencies, and forums in Niger.
- National stakeholders, including national and subnational government actors, and non-governmental organizations.

152. The main users of the evaluation (country office management and 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.

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

- Reviewing and commenting on the 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.
- Reading the final evaluation report and other evaluation communication products.

154. More detailed information about evaluation users is provided in Table 11, below. This table introduces all the 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 13: 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 a 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 it to involve them in the evaluation? (High, medium, low)
WFP internal stakeholders					
WFP Country Office	The main implementer of the programme under evaluation	To inform upcoming country strategic plan and relevant programming	The Country Office is responsible for implementing the programme according to the evaluation design. It actively provides 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 country offices 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, with regular meetings to provide feedback on tools and outputs	High
Office of Evaluation	Coordination of impact evaluation and liaisons with the country office	As coordinators of the impact evaluation and for WFP, in alignment with 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 country office and impact evaluation team	From the scoping stage	High
External stakeholders					
Affected communities	Affected communities, including men, women, boys, and girls, will be the primary participants of the intervention	Beneficiaries will likely have a strong interest in any changes in targeting, reach, or effectiveness of future programming because of the evaluation and recommendations. Women and girls have a particular stake in the results meant to shed light on recommendations for improving gender equality	Beneficiaries and non-beneficiaries alike will provide the primary sources of data on effectiveness	From the targeting and selection stage	High
Government at the central level:	National and local government structures	WFP has an established relationship with the national Government to provide	The evaluation receives national-level clearance before inception	At the initial scoping for the	Medium

Regional governments in Diffa, Dosso, Maradi Cellule Filets Sociaux	provide ethical and administrative clearance for programming and evaluation efforts and oversee local development initiatives and national social protection programmes.	technical support on food security and nutrition; evaluation results will support these efforts		intervention, targeting for the intervention & dissemination of findings	
Local non-governmental organizations: Volontaires Nigériens pour le Développement (VND NUR) Agence pour le bien être (APBE) Masnat Garkua	As implementing partners for the programme under evaluation	Evaluation results can inform their own livelihood and gender transformation programming	As cooperating partners, the NGOs listed are responsible for ensuring the programme is implemented in line with the evaluation design	At the initial scoping for the intervention & dissemination of findings	High
International non-governmental organizations: CARE	CARE is an implementing partner for the programme under evaluation	Evaluation results can inform their own livelihood and gender transformation programming	As a cooperating partner, CARE is responsible for ensuring the programme is implemented in line with the evaluation design	At the initial scoping for the intervention & dissemination of findings	High
World Bank	Development Impact Evaluation (DIME) department	In line with the Office of Evaluation-DIME partnership, DIME is interested in producing and disseminating the evaluation results as part of a broader research portfolio	As the primary investigators and research analysts	At the initial conceptualization of the window	High
Donor: Federal Ministry for Economic Cooperation and Development (BMZ)	Primary funder of the intervention	As a user of the evaluation	BMZ is informed at key milestones in the evaluation. They have an interest in using the results as evidence for other funded projects	At the proposal stage of the intervention	Medium

ANNEX 8: DETAILED EVALUATION PROCESS

Table 14: Detailed evaluation timeline

Phase 1 – Preparation	Involved	Estimated Date
Regional discussion on Impact Evaluation	CO/OEV	March 2019
Discussions with BMZ	BMZ/CO/OEV/DIME	May 2019
Impact Evaluation approved by BMZ	BMZ	August 2019
Impact Evaluation Workshop	CO/DIME/OEV	September 2019
Impact Evaluation Inception mission to Niger	CO/OEV/DIME	February 2020
Phase 2 - Inception report		
Regional discussion on Impact Evaluation	DIME/OEV	March 2020
Partnership Steering Committee	OEV	May 2020
Impact evaluation design discussions & adapting to COVID-19	CO/OEV/DIME	April–September 2020
Inception Report drafting, quality assurance and peer review	OEV/DIME	January–December 2021
Dissemination of the inception report with country office, regional bureau, evaluation committee, window's reference group, steering committee, online/social media as appropriate	DIME/OEV	March 2022
Phase 3 – Baseline and high-frequency data collection		
Pre-baseline wealth ranking and baseline data collection preparations	CO/OEV/DIME	October–December 2020
Baseline data collection	DIME	December–January 2021
Data collection through high-frequency surveys	DIME/OEV/CO	April 2021–August 2022
Phase 4 – Baseline report		
Data analysis and baseline report drafted by the impact evaluation team, and submitted for quality assurance and revisions	DIME/OEV	March 2022
Publication of the baseline report	OEV	March 2022
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 appropriate	DIME/OEV	March 2022
Phase 5 – Programme implementation		
Randomization	DIME	March 2021
Assignment intervention and comparison sites	DIME/CO	March 2021
Rolling out programme activities as per randomization	CO	April 2021

Monitoring programme activities verifying treatment and control status	CO/DIME	April 2021–Endline
Phase 6 – Endline data collection		
Preparation of data collection tools, including survey questionnaire, digital devices, sampling strategy, and training material.	DIME/OEV/CO	December 2022
Pilot and finalization of data collection tools	DIME	March 2023
Recruitment of enumerators/data collection firm	CO	April 2023
Enumerators training	CO	April 2023
Data collection process and live monitoring data quality checks	DIME	May 2023
Feedback/data sharing mechanisms, as appropriate/possible		July 2023
Phase 7 – Final evaluation reports		
Data analysis and final evaluation report drafted by the impact evaluation team, and submitted for quality assurance and revisions	DIME/OEV	April 2022
Publication of the final evaluation report	OEV	June 2023
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 appropriate	OEV/DIME/CO	June 2023
Final evaluation report reviewed by post hoc quality assessment	OEV	June 2023
Phase 8 – Management response		
Based on the findings, country office to develop a management response	CO	August 2023
The Office of Evaluation to review and, if needed, respond to the management response	OEV	Sept 2023
Publication of the management response	OEV	October 2023
Phase 9 – Dissemination and learning		
Webinar presenting the findings	OEV/DIME	August 2023
Blogs, summary briefs, other relevant communication products	OEV/DIME	September 2023
Considerations for academic publications	DIME/OEV	November 2023

ANNEX 9: REFERENCES

- Aker, J. C., Boumnijel, R., McClelland, A. & Tierney, N.** 2016. Payment Mechanisms and Antipoverty Programs: Evidence from a Mobile Money Cash Transfer Experiment in Niger. *Economic Development and Cultural Change*, 65(1): 1–37.
- Annan, F. & Sanoh, A.** 2018. Social Protection in Niger: What Have Shocks and Time Got to Say? Policy Research Working Paper 8455. World Bank, Washington, DC.
- Archibald, E., Bossuroy, T. & Premand, P.** 2020. Productive Inclusion Measures and Adaptive Social Protection in the Sahel: A Case Study. Washington, DC: The World Bank.
- Asfaw, S., Pallante, G. & Palma, A.** 2018. Diversification Strategies and Adaptation Deficit: Evidence from Rural Communities in Niger. *World Development*, 101: 219–234.
- Barrett, C. & Constan, M.** 2014. Toward a Theory of Resilience for International Development Applications. *Proceedings of the National Academy of Sciences of the United States of America*, 111(40): 14625–14630.
- Bossuroy, T., Goldstein, M., Karlan, D., Kazianga, H., Parienté, W., Premand, P., Thomas, C., Udry, C., Vaillant, J. & Wright, K. 2021. Pathways Out of Extreme Poverty: Tackling Psychosocial and Capital Constraints with a Multi-Faceted Social Protection Program in Niger. Policy Research Working Paper 9562. World Bank, Washington, DC.
- Cissé, J. & Barrett, C.** 2018. Estimating Development Resilience: A Conditional Moments-Based Approach. *Journal of Development Economics*, 135: 272–284.
- FSIN. 2021.** Global Report on Food Crises: Joint Analysis for Better Decisions. [\[Link\]](#)
- FSIN Resilience Measurement Technical Working Group.** 2014. Resilience Measurement Principles: Toward an Agenda for Measurement Design. Rome, FAO & WFP.
- Gunnsteinsson, S., Adhvaryu, A., Christian, P., Labrique, A., Sugimoto, J., Shamim, A. & West, K.** 2019. Protecting Infants from Natural Disasters: The Case of Vitamin A Supplementation and a Tornado in Bangladesh. NBER Working Paper 25969.
- Hoddinott, J., Sandström, S. & Upton, J.** 2018. The Impact of Cash and Food Transfers: Evidence from a Randomized Intervention in Niger. *American Journal of Agricultural Economics*, 100(4): 1032–1049.
- Macours, K., Premand, P. & Vakis, R.** Forthcoming. Transfers, Diversification and Household Risk Strategies: Experimental Evidence with Lessons for Climate Change Adaptation. *The Economic Journal*.
- Phadera, L., Michelson, H., Winter-Nelson, A. & Goldsmith, P.** 2019. Do Asset Transfers Build Household Resilience? *Journal of Development Economics*, 138(C): 205–227.
- Premand, P. & Barry, O.** 2020. Behavioural Change Promotion, Cash Transfers and Early Childhood Development: Experimental Evidence from a Government Program in a Low-income Setting. Policy Research Working Paper 9368. World Bank: Washington, DC.
- Premand, P. & Stoeffler, Q.** 2020. Do Cash Transfers Foster Resilience? Evidence from Rural Niger. Policy Research Working Paper 9473. World Bank, Washington, DC.
- UNHCR Operational Portal.** Figures as of 31 January 2022. [\[Link\]](#)
- World Bank.** 2009. World Development Report 2010: Development and Climate Change. Washington, DC.
- World Bank.** 2016. Republic of Niger: Priorities for Ending Poverty and Boosting Shared Prosperity Systematic Country Diagnostic (SCD) Concept Note. World Bank: Washington, DC.
- World Food Programme.** N.d. Niger. [\[Link\]](#)
- World Food Programme.** 2021. Niger Country Brief. January 2021. [\[Link\]](#)
- World Food Programme.** 2022. Niger Country Brief. January 2022. [\[Link\]](#)
- World Food Programme Niger.** 2021. Annual Country Report 2020. [\[Link\]](#)

Office of Evaluation

World Food Programme

Via Cesare Giulio Viola 68/70
00148 Rome, Italy
T +39 06 65131 wfp.org