Food and Agriculture Organization of the United Nations

Impact evaluation of FAO's cash-for-work programme in Somalia

Findings from the baseline survey

Impact evaluation of FAO's cash-for-work programme in Somalia

Findings from the baseline survey

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS
Rome, 2024

Required citation:
FAO. 2024. *Impact evaluation of FAO's cash-for-work programme in Somalia – Findings from the baseline survey*. Rome.
https://doi.org/10.4060/cd2760en

Abstract

This report presents the results from the analysis of baseline data collected for the evaluation of the Food and Agriculture Organization of the United Nations (FAO)'s cash-for-work programme in Somalia. This programme provides temporary employment opportunities in the maintenance or rehabilitation of public infrastructure. Its primary objective is to offer income support to vulnerable individuals while facilitating the development of essential community assets. To evaluate the impact of the programme, we use a longitudinal, quasi-experimental design study, with one baseline and three follow-up surveys at 12, 24 and 36 months. The evaluation focuses on 14 districts in Somalia, Somaliland and Puntland. Within each district, we selected four treatment villages and four comparison villages. The baseline data collection took place in January/February 2023. We calculate five main programme outcomes as summary indexes: resilience, food security, income and livelihoods, agency, social cohesion. For all of them, we do not detect any statistically significant differences between the two evaluation arms. There are small imbalances between the treatment and comparison groups on three characteristics. The treatment arm reports a larger number of income sources, a larger number of months with members participating in wage labour and greater participation in social networks than the comparison group. These results were expected, since the baseline survey took place around three weeks after the public works activities started, though payments were not made. The analysis of sociodemographic characteristics points to a relatively successful targeting approach. We observe a relatively large share of people of working age and a minor share of elderly people. Households lacking adult members are only 1 percent of the sample. Further, one out of ten households report having a member with a moderate or severe disability.

Contents

Abstract ... iii
Acknowledgements ... vii
Abbreviations ... viii
Executive summary ... ix
1. Introduction ... 1
2. Theory of change and hypotheses tested ... 3
3. Study design ... 7
3.1 Comparison group selection ... 7
3.2 Data collection ... 8
3.3 Survey instruments ... 9
4. Descriptive analysis ... 11
4.1 Methodology ... 11
4.2 Demographics ... 12
4.3 Goal and outcome indicators ... 14
4.4 Resilience ... 18
4.5 Food security ... 19
4.6 Income and livelihoods ... 20
4.7 Psychological well-being ... 22
4.8 Social cohesion ... 23
5. Baseline data econometric analysis ... 27
5.1 Relationship between resilience and outcomes ... 27
5.2 Relationship between resilience and climate ... 29
6. Estimation strategy ... 33
6.1 Power and attrition ... 33
7. Conclusions ... 37
References ... 41
Appendix 1. Additional tables and figures ... 44

Figures and tables

Figure 1. Theory of change for the effect of the cash-for-work programme on households' resilience 5
Figure 2. Household size distribution 12
Figure 3. Sample distribution by age and gender 13
Figure 4. Kernel density of resilience scores 19
Figure 5. Food consumption score categories by treatment group 20
Figure 6. Livelihood source categories by treatment group 22
Figure 7. Relationship between resilience and programme outcomes 28
Figure 8. Relationship between minimum detectable effect and intracluster correlation coefficient 34
Figure 9. Relationship between minimum detectable effect and autocorrelation parameter 35

Table 1. Sample size of households, individuals and communities, by treatment group and district 8
Table 2. Average household characteristics by treatment group 14
Table 3. Adult members (15+) disability status by treatment group 14
Table 4. Main programme outcomes by treatment group 18
Table 5. Resilience measurements by treatment group 18
Table 6. Food security status by treatment group 19
Table 7. Income variables by treatment status 20
Table 8. Livestock ownership by treatment status 22
Table 9. Agency variables by treatment status 23
Table 10. Agency variables by gender of the main respondent 23
Table 11. Social cohesion by treatment status 25
Table 12. Social cohesion by gender of the main respondent 25
Table 13. Correlation between subjective resilience and programme outcomes 28
Table 14. 3-months Standardized Precipitation Index by year, season and region 31
Table 15. Correlation between subjective resilience, climate and weather shocks 32

Appendix table 1. Topics in the survey questionnaires 44
Appendix table 2. Nine-capacities model questions 44
Appendix table 3. Agency questions 45
Appendix table 4. Attitudes statements to construct social cohesion index 45
Appendix table 5. Unadjusted and Holm- Sidak adjusted p-values for baseline comparisons 45

Acknowledgements

Silvio Daidone, Fatimetou El Bah, Navya Sureka and Clemencia Cosentino carried out this study. We gratefully acknowledge the useful support and comments received from colleagues from the FAO Office of Evaluation, the FAO Somalia Country Office and the FAO Rural Transformation and Gender Equality Division.

Abbreviations

DiD	difference-in-difference
FAO	Food and Agriculture Organization of the United Nations
ICC	intracluster correlation coefficient
IPC	Integrated Food Security Phase Classification
LASSO	least absolute shrinkage and selection operator
MDE	minimum detectable effect
RCS	resilience capacity score
SPI	standardized precipitation index
TLU	Tropical livestock units

Executive summary

1. This report presents the results from the analysis of baseline data collected for the evaluation of the Food and Agriculture Organization of the United Nations (FAO)'s cash-for-work programme in Somalia.

Background

2. In 2023, FAO Somalia contributed to the Humanitarian Response Plan and the Drought Response and Famine Prevention Plan through a programme titled "Providing Emergency Life-saving Food and Livelihood Support to Drought-affected Communities in Somalia." Aimed at improving food security for the most drought-affected populations in rural Somalia, this programme is organized around the following five interventions:

   i. Cash +
   ii. Unconditional cash transfers (UCT) and cash-for-work
   iii. Transitional Cash and Livelihood Programme
   iv. Somalia Water and Land Information Management
   v. Desert Locust Control and Surveillance

3. At the geographical level, the cash-for-work intervention targets rural communities and villages located more than 20 km away from the main district town in IPC3+ status, without cash or in-kind support from other humanitarian organizations and where there is a communal infrastructure that needs to be rehabilitated. Households with able-bodied members are eligible for the intervention, while community-based targeting aims to prioritize the poorest and most vulnerable households (with ill family members, elderly, disabled, female-headed).

4. Cash-for-work provides temporary employment opportunities in public projects, such as road construction and maintenance or rehabilitation of public infrastructure. Its primary objective is to offer income support to vulnerable individuals while facilitating the development of essential community assets. Local communities identify assets in need of maintenance or rehabilitation, including small-scale water and soil conservation structures like water catchments, irrigation canals, contour bunds, and feeder roads. In 2023, FAO engaged 11 300 households across 21 districts in cash-for-work activities to provide them with a direct source of cash while also contributing to the rehabilitation of critical productive, rural, communally-owned infrastructure.

Impact evaluation design

5. To evaluate the impact of FAO's cash-for-work programme, we use a longitudinal, quasi-experimental design study, with one baseline and three follow-up surveys at 12, 24 and 36 months. The evaluation focuses on 14 districts in Somalia, Somaliland and Puntland. Within each district, we selected four treatment villages and four comparison villages.

6. The comparison group was selected using propensity score matching, which was based on the following characteristics: distance from permanent and functioning strategic water source, distance from a major town, distance from a police station, distance from the primary road and distance from health centre/post. Villages included in the matching exercise are located in rural areas of targeted districts and where infrastructure (such as feeder roads, contour bunds, or water catchments) existed within a 6 km radius of the village, and that were not among those already rehabilitated by the cash-for-work programme in the 2023 programme round.

7. At follow-up, we plan to estimate the cash-for-work impacts using a difference-in-difference (DiD) statistical model that compares change in outcomes between baseline and follow-up and between cash-for-work treatment and comparison groups. The combination of propensity score matching with DiD helps us correct for possible imbalances across the two groups. Specifically, we will use inverse probability weighting with regression adjustment, that has the convenient property of being doubly robust.

Baseline survey findings

8. The baseline data collection took place in January/February 2023. A total of 1 697 households were interviewed in 113 clusters, representing 9 158 individuals.

9. We calculate five main programme outcomes as summary indexes: resilience, food security, income and livelihoods, agency, social cohesion. For all of them, we do not detect any statistically significant differences between the two evaluation arms.

10. There are small imbalances between the treatment and comparison groups on three characteristics. The cash-for-work treatment arm reports a larger number of income sources, a larger number of months with members participating in wage labour and greater participation in social networks than the comparison group. These results were expected, since the baseline survey took place around three weeks after the public works activities started, though payments were not made. We can control these small differences in modelling.

11. The analysis of sociodemographic characteristics points to a relatively successful targeting approach. We observe a relatively large share of people of working age and a minor share of elderly people. Households lacking adult members are only 1 percent of the sample. Further, one out of ten households report having a member with a moderate or severe disability.

12. Results indicate that the evaluation will be able to rely on these two groups – and the baseline data collected – to assess impacts over time.

1. Introduction

1. Somalia currently faces severe drought conditions caused by five consecutive poor rainy seasons. Most of the country experienced a hotter and drier than normal dry (Jilaal)[1] season, prolonging the severe effects of the historic drought on pastoral communities. The below-average rainfall has had detrimental effects on harvests, livestock and income, leading to an unprecedented level of food insecurity (FEWSNET and FSNAU, 2022). With a consistent decline in food production, Somalia heavily depends on food imports to mitigate the food deficit. However, the conflict in Ukraine has significantly worsened the situation, as Russia and Ukraine previously supplied over 90 percent of the country's wheat (FAO, 2022a). High global food and transportation prices compounded the effect of five consecutive seasons marked by below-average harvests and excessive livestock losses, aggravating the food security situation for the most vulnerable households in Somalia. The Famine Early Warning Systems Network and the Food Security and Nutrition Analysis Unit estimated that 6.5 million Somalis would face crisis (Integrated Food Security Phase Classification [IPC] Phase 3) or worse acute food insecurity outcomes in 2023.

2. Somalia received large-scale humanitarian assistance to mitigate the worsening food security situation. The Food and Agriculture Organization of the United Nations (FAO) in Somalia contributed to the Humanitarian Response Plan 2023 and the Drought Response and Famine Prevention Plan through the "Providing Emergency Life-saving Food and Livelihood Support to Drought-affected Communities in Somalia" programme, aimed at improving food security for the most drought-affected populations in rural Somalia (FAO, 2022b). The programme is organized around five interventions:

    i. Cash+
    ii. Unconditional cash transfers and cash-for-work[2]
    iii. Transitional Cash and Livelihood Programme
    iv. Somalia Water and Land Information Management
    v. Desert Locust Control and Surveillance

3. Several donors support this work. While the United States Agency for International Development's Bureau of Humanitarian Assistance is FAO Somalia's primary Cash+ donor, other resource partners such as the Swedish International Development Cooperation Agency and Germany's Federal Foreign Office provide resources for agricultural complementary livelihood inputs. This report seeks to establish guidelines to conduct an impact evaluation of a key element of the second component of the short-term response initiatives of the Somalia Humanitarian Response Plan, namely the cash-for-work programme.

4. The cash-for-work intervention provides temporary employment opportunities in public projects, such as road construction, maintenance and rehabilitation of public infrastructure. The primary

[1] Somalia has two rainy and two dry seasons. The *gu* rains (April–June) are followed by the *hagaa* dry season (July–September), and *deyr* rains (October–November) are followed by the *jiilaal* dry season (December–March).

[2] Within the same cash-for-work locations, FAO also targets the most food-insecure households. These households are unable to actively participate in the public works activities but receive unconditional cash transfers for the same duration as the cash-for-work households, albeit for a smaller amount. This includes, for instance, households headed by ill family members, pregnant women, the elderly and people living with disabilities.

objective of cash-for-work programmes is to offer income support to vulnerable individuals while facilitating the development of essential community assets. In 2023, FAO engaged 11 300 households across 21 districts in cash-for-work activities to provide them with a direct source of cash, while also contributing to the rehabilitation of critical productive, rural, communally-owned infrastructure. Guided by the criteria provided by FAO, local communities identify assets in need of rehabilitation. These assets include small-scale water and soil conservation structures like water catchments, irrigation canals, contour bunds and feeder roads. In addition, public works may also involve small-scale bush clearing and afforestation activities along rehabilitated water catchments.

2. Theory of change and hypotheses tested

5. The cash-for-work programme revolves around its dual objectives: providing cash income to poor and vulnerable rural households affected by climate shocks, while simultaneously restoring community assets that enhance local productivity and foster long-term development.

6. Figure 1 illustrates the theory of change for this two-pronged approach in addressing poverty and food insecurity. The left side of the figure represents the safety net component, which entails providing cash assistance to beneficiary households as remuneration for their participation in public works. The main objective of this component is to ensure that eligible households can meet their immediate needs and have access to essential resources, particularly food. However, we expect that the impact of the cash component will go beyond mere income generation. Drawing from existing literature, several hypotheses can be supported. First, the cash-for-work programme has the potential to enhance household's capacity to manage risks, leading to greater diversification of livelihood strategies. This assertion aligns with findings from Pavanello *et al.* (2016) who examined the effects of cash transfers on community interactions. Their findings indicate that a predictable cash flow improves strategic livelihood choices, stimulates productive investments, and encourages beneficiaries' entry into risk-sharing agreements and economic collaboration networks. Similarly, Pace *et al.* (2022) conducted a study on the impact of a social protection programme in Zimbabwe, focusing on the effect of cash transfers on food security and livelihood. Their research revealed a significant shift from survival-led diversification to opportunity-led diversification among beneficiary households. This shift indicates that the programme not only provides immediate support but also leads to a transformation in livelihood strategies, enabling households to pursue more sustainable and growth-oriented economic activities.

7. Furthermore, beyond improving material well-being, the cash component can alleviate the psychological constraints that impede decision-making and perpetuate household poverty and vulnerability. Haushofer and Shapiro (2016) present compelling evidence of this causal link by investigating the impact of cash transfers on psychosocial well-being and economic outcomes of poor and vulnerable households in rural Kenya. Beneficiary households reported improvement for several indicators, including better mental health and reduced levels of stress. Similar results are found in other countries like South Africa (Ohrnberger *et al.,* 2020) or Malawi (Angeles *et al.*, 2019), where national social cash transfer programmes significantly improved mental health outcomes. Drawing from a mixed-method evaluation of a cash transfer programme in Kenya and qualitative research conducted in Ghana, Zimbabwe, and Lesotho, Attah *et al.* (2016) found that cash transfers can lead to improvements in psychosocial well-being, subsequently contributing to improved educational performance, increased participation in social life and empowerment in decision-making processes.

8. Lastly, the cash component of the cash-for-work programme also has the potential to strengthen social cohesion and positive community dynamics through participation in social networks of reciprocity and formal group membership. Fisher *et al.* (2017) conducted a study exploring the livelihood impacts of cash transfers in sub-Saharan Africa. They highlight the positive effects of social transfers on strengthening bonding social capital and breaking patterns of exclusion. The research by Valli *et al.* (2019) further supports these findings. In their study, the authors conducted a cluster randomized control trial to assess the impact of a short-term transfer programme on social cohesion among Colombian refugees and poor Ecuadorians in urban areas of northern Ecuador. The study demonstrated that the programme had a positive influence on social cohesion among Colombian refugees within the hosting community. Specifically, recipients reported an

increase in personal agency, more accepting attitudes towards diversity, higher confidence in institutions and greater levels of social participation.

9. The right side of Figure 1 illustrates how the productive component of the cash-for-work programme strengthens the resilience of local economies, potentially benefiting both recipients and non-recipients. When public works coincide with the slack agricultural season, the cash-for-work programme creates off-farm labour opportunities, diversifying beneficiaries' livelihoods in the short term. This diversification can potentially contribute to enhanced social cohesion and trust. Andrews and Kryeziu (2013) highlight three potential pathways that contribute to this outcome.

10. First, through the incorporation of participatory elements in programme design, excluded groups can have a platform to voice their concerns and interact with local government officials, thereby fostering social inclusion and trust. For instance, communities in villages where the public works programme is implemented participate in various aspects of programme implementation, including household identification, project selection and project monitoring (Conning and Kevane, 2002). This active engagement empowers community residents by granting them improved access to information and decision-making processes, further enhancing their sense of ownership and participation. A mixed methods evaluation of cash-for-work programmes for Syrian refugees and their local neighbours in Jordan found that these interventions strengthened the sense of belonging and trust of participants and non-participants, refugees and locals, and in particular women (Zintl and Loewe, 2022). A similar story emerges from the analysis of the relationship between Malawi's largest and oldest public works programme and social cohesion, specifically within community cooperation for the common good (Beierl and Dodlova, 2022). Second, the empowerment of marginalized groups – such as women, ethnic minorities and disadvantaged youth – through employment opportunities, can result in long-term effects on equity and inclusion. The provision of jobs in times of crisis and recovery can restore a sense of dignity and social identity, particularly in low-income and post-conflict settings. Last, in the medium term, the rehabilitation of community assets is expected to have a positive impact on the economic environment, leading to improved land quality, higher agricultural yields, and enhanced access to markets.

11. In studying the effect of the FAO Somalia cash-for-work on households' resilience, and following the theoretical framework just described, we formulate the following hypotheses (H):

i. H1: Cash-for-work transfers will increase household resilience and reduce the adoption of negative coping strategies.

- Cash-for-work transfers can mitigate the impact of shocks on household consumption and income. They reduce households' vulnerability to adverse weather events and decrease the likelihood of resorting to negative coping mechanisms, such as reducing portion size or frequency. In addition, transfers will directly alleviate the credit and liquidity constraints faced by poor rural households.

ii. H2: Cash-for-work transfers will increase food security of beneficiary households.

- The most immediate impact of the cash-for-work payments is an increase in household income, which plays a significant role in addressing food security. In the short term, beneficiaries have greater access to an adequate supply of safe and nutritious food that fulfils their dietary needs. This, in turn, improves physical well-being, reduces days of work lost, and ultimately increases labour productivity.

iii. H3: Cash-for-work transfers will increase income diversification.

- Cash-for-work transfers facilitate diversification by allowing temporary labour market opportunities and enabling households to invest in income-generating activities that offer a higher return than subsistence agriculture.

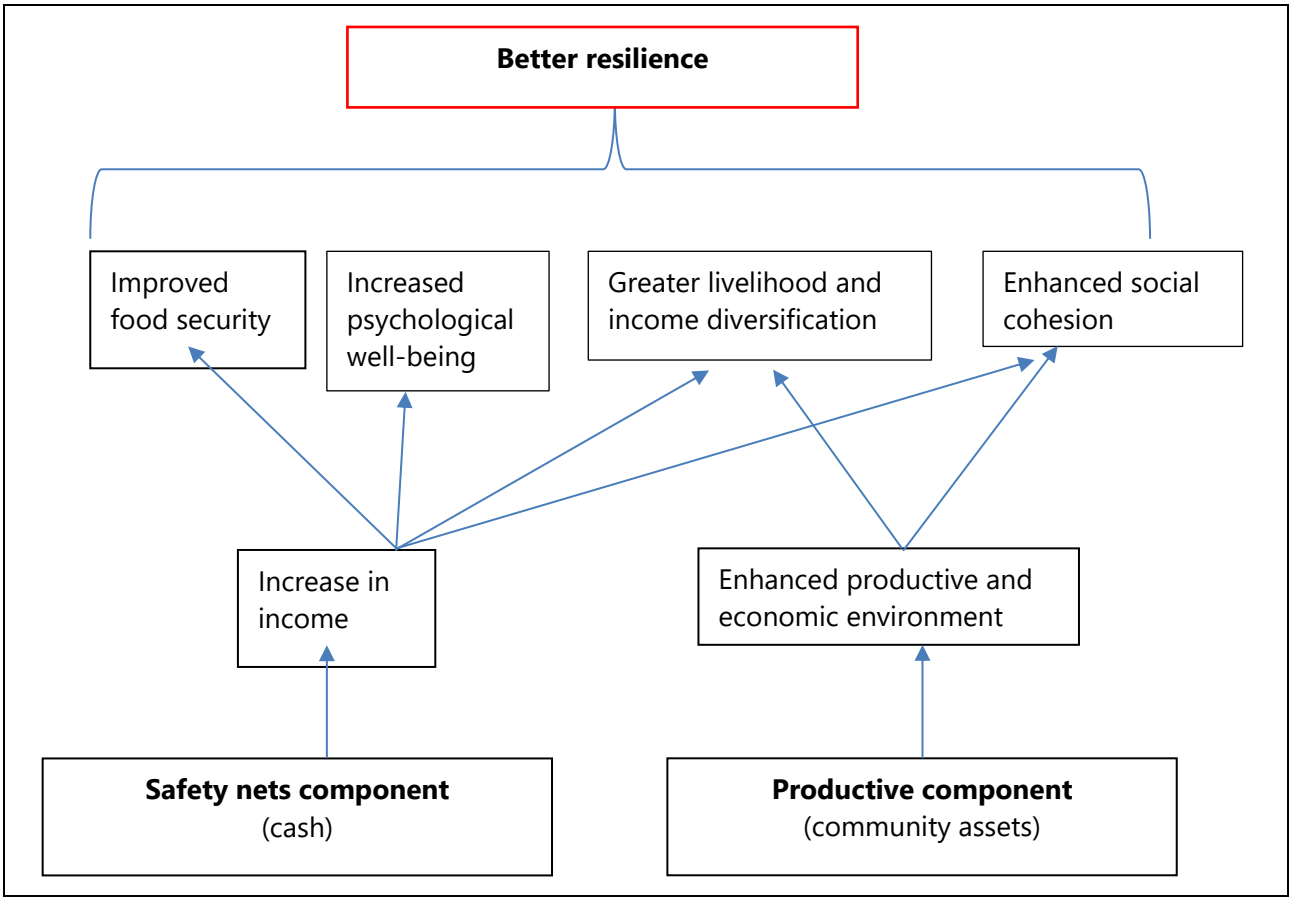
iv. H4: Cash-for-work will improve psychological well-being.

- Cash-for-work transfers can have psychological benefits, by allowing recipients' flexibility concerning spending modalities. In the short-term we expect beneficiaries to increase their agency, perceiving greater ability and a sense of control over the outcomes of events in their life.

v. H5: Cash-for-work will improve social cohesion.

- Cash-for-work increases the capacity of individual and household beneficiaries to participate in cultural, social and familial activities, fostering "bonding" social capital. This programme strengthens or reintegrates existing social networks by enabling recipients to join community events, share food and borrow when in need thanks to their enhanced capacity to repay.

Figure 1. Theory of change for the effect of the cash-for-work programme on households' resilience



Source: Authors' own elaboration.

3. Study design

12. To evaluate the impact of the cash-for-work programme we use a longitudinal, quasi-experimental design focusing on 14 districts in the regions of Somalia, Somaliland and Puntland. Within each district, we selected four treatment villages and four comparison villages. The selection of villages for participation in the public works was initially done based on specific eligibility criteria, with a key factor being the availability of able-bodied members within the household. FAO Somalia provided a list of villages involved in the public work along with household beneficiaries from these villages. From this list, the Evaluation Team selected 15 households per village, ensuring a balanced representation of beneficiary characteristics within the sample.

3.1 Comparison group selection

13. For its humanitarian response plan, FAO primarily targets rural communities and villages located more than 20 km away from the main district town and not receiving support from other humanitarian organizations. Since randomization was not possible, the Evaluation Team used propensity score matching to construct a comparison group of villages from the same districts where the cash-for-work programme was implemented.

14. The selection of villages for the comparison group was carried out by the Evaluation Team based on the following characteristics: distance from permanent and functioning strategic water source, distance from a major town, distance from a police station, distance from the primary road and distance from health centre/post. The identification of these villages was carried out as follows:

    i. The Evaluation Team calculated a propensity score for all villages in the sample, both treated and not treated, using the aforementioned criteria. This score helped identify comparison villages that shared similar characteristics with the treated villages in our sample.

    ii. The final list of comparison villages included multiple replacements to account for cases where it might not be feasible to reach a village or where the necessary conditions for inclusion in the sample were lacking (as established in the verification step described below).

    iii. The list of comparison villages, along with their potential replacements, was then shared with the data collection service provider who engaged with the chiefs of these villages to verify the presence of specific infrastructure, namely feeder roads, contour bunds or water catchments.

    iv. If any of the mentioned infrastructure existed within a 6 km radius of the village and were not among those already rehabilitated by the cash-for-work programme in the current project round, the village was confirmed to be eligible for the comparison group. In case the required infrastructure was unavailable, the service provider proceeded to the next village on the list.

15. Once a village was confirmed, enumerators collected information from the village authorities or chiefs regarding the approximate number of households in the village and a list of households with labour capacity. From this list, 15 households were selected at random by the service provider for the interview.

3.2 Data collection

16. The cash-for-work impact evaluation involves several rounds of data collection. At present, we foresee a baseline (the focus of this report) and three follow-up surveys at 12, 24 and 36 months. The baseline data collection took place from 26 January to 11 February 2023. The fieldwork was conducted by a team consisting of one project coordinator and 25 field enumerators.

17. The collection of baseline data was a necessary step in the evaluation process for several reasons. In addition to enabling us to measure the impacts of the programme, the baseline data allowed us to i) document the initial living conditions and circumstances of both cash-for-work participants and the comparison group prior to any cash transfers being disbursed; and ii) provide a comprehensive and detailed sociodemographic description of the beneficiaries, enabling the Evaluation Team to identify and account for any differences that might exist between the treatment and comparison groups at the start of the evaluation. By checking for these differences during the analysis, the evaluation results become more reliable and credible.

18. The Evaluation Team originally aimed to gather a total of 1 680 observations, which were supposed to be distributed as follows:

    i. 840 observations in 56 villages (with 15 observations per village), covering four villages per district in 14 districts;

    ii. 840 observations in 56 neighbouring villages (also with 15 observations per village), again covering four villages per district in 14 districts.

19. We surpassed this target by conducting a total of 1697 interviews in 113 villages, representing 9 158 individuals. This was achieved by including an additional comparison village in the district of Xudun.

Table 1. Sample size of households, individuals and communities, by treatment group and district

District	Comparison	Treatment	Total
Borama	60	60	120
	423	*389*	*812*
	(4)	(4)	(8)
Burco	60	60	120
	366	*425*	*791*
	(4)	(4)	(8)
Laas Caanood	60	60	120
	245	*283*	*528*
	(4)	(4)	(8)
Xudun	75	59	134
	322	*255*	*577*
	(5)	(4)	(9)
Ceerigaabo	60	61	121
	259	*295*	*554*
	(4)	(4)	(8)
Qardho	60	60	120
	347	*398*	*745*
	(4)	(4)	(8)
Burtinle	60	60	120
	330	*363*	*693*
	(4)	(4)	(8)
Eyl	60	60	120
	377	*361*	*738*

District	Comparison	Treatment	Total
	(4)	(4)	(8)
Hobyo	60	59	119
	276	*354*	*630*
	(4)	(4)	(8)
Jariiban	60	60	120
	367	*340*	*707*
	(4)	(4)	(8)
Dhuusamarreeb	60	60	120
	287	*313*	*600*
	(4)	(4)	(8)
Cabudwaaq	60	60	120
	265	*245*	*510*
	(4)	(4)	(8)
Cadaado	62	61	123
	280	*300*	*580*
	(4)	(4)	(8)
Belet Weyne	60	60	120
	380	*313*	*693*
	(4)	(4)	(8)
Total	857	840	1 697
	4 524	*4 634*	*9 158*
	(57)	(56)	(113)

Note: Figures in italic (grey shaded rows) represent individuals. Number of clusters in parentheses.
Source: Authors' own elaboration.

3.3 Survey instruments

20. The household questionnaire covered questions across various topics to gather a wide range of relevant information. Topics covered in the questionnaire include:

i. Household composition: collecting data on household members, such as age, gender, and their relationship to the main respondent.

ii. Land ownership and crop production: gathering information about land ownership, tenure arrangements, and the household's engagement in crop cultivation.

iii. Livestock possession: obtaining information about household's ownership and management of livestock.

iv. Consumption: understanding the household's consumption patterns and expenditure on various goods and services.

v. Coping strategies: learning about the strategies employed by the household to cope with various challenges, including economic shocks or food insecurity.

vi. Access to basic services: assessing the household's access to essential services such as water supply and sanitation facilities.

vii. Social cohesion: obtaining information about the dynamics of social relationships and support networks within the household and the community.

viii. Resilience to shocks: assessing the household's resilience to weather-related shocks, such as droughts or floods.

21. The design of the household survey was guided by three core principles:

i. Include key indicators to enable an assessment of the programme against its stated objectives. These core indicators include food security, transfers, access to markets, and savings among other important factors.

ii. Whenever possible, rely on existing survey questions used in similar surveys conducted in the country. This ensured that the questions were appropriate and relevant for local conditions, and facilitated the comparison of data with national-level information.

iii. Limit the length of the survey to prevent interviewer and respondent fatigue and promote high data quality by increasing the willingness of respondents to provide accurate and detailed information.

4. Descriptive analysis

4.1 Methodology

22. This baseline report provides a snapshot of the livelihoods and family characteristics of future recipients before they started receiving the payments for the work carried out to rehabilitate existing community infrastructure. We describe demographic features of the sample and several characteristics that relate to either the goals of the programme or potential mediators for expected outcomes.

23. To test for statistical differences, we carry out simple t-tests of the mean difference between the comparison and cash-for-work treatment groups in the main outcome variables, outputs and other sociodemographic characteristics. Specifically, we test the null hypothesis:

$$H_0: \mu_C = \mu_T \qquad (1)$$

where μ_C and μ_T are the group means for the comparison and the cash-for-work beneficiaries groups respectively. Because of their non-random allocation, we expect to find significant differences between the two groups.

24. The tables with the baseline comparisons across the treatment groups are organized as follows, from left to right. Under the column *All* we provide the overall sample average. Under the columns *cash-for-work* and *comparison* we report the group means for the cash-for-work and the comparison groups respectively. In the fourth and fifth columns we provide the mean difference and the p-value of the related t-test for the null hypothesis that the sample mean of a given variable is the same in the two treatment arms. Finally, some tables include an additional column reporting the total number of observations used to calculate the mean difference and carry out the test.[3]

25. The analyses conducted include multiple comparisons, as the same null hypothesis is tested for multiple outcomes. As a result, p-values obtained to assess whether differences are statistically significant are likely incorrect. Therefore, the typical 5 percent threshold for a Type I error (that is, the probability of falsely rejecting the null hypothesis of no average difference) used in single hypothesis testing must be replaced by the family-wise error rate (FWER) to account for the probability of incorrectly rejecting even one null hypothesis in a sequence of hypotheses. To control for the FWER, we employ the Holm-Sidak correction (McDonald, 2008).[4] The results reported in Appendix table 4 show small but significant differences in only two characteristics (main respondent's age and number of income sources). These differences will be controlled for in future impact analyses.

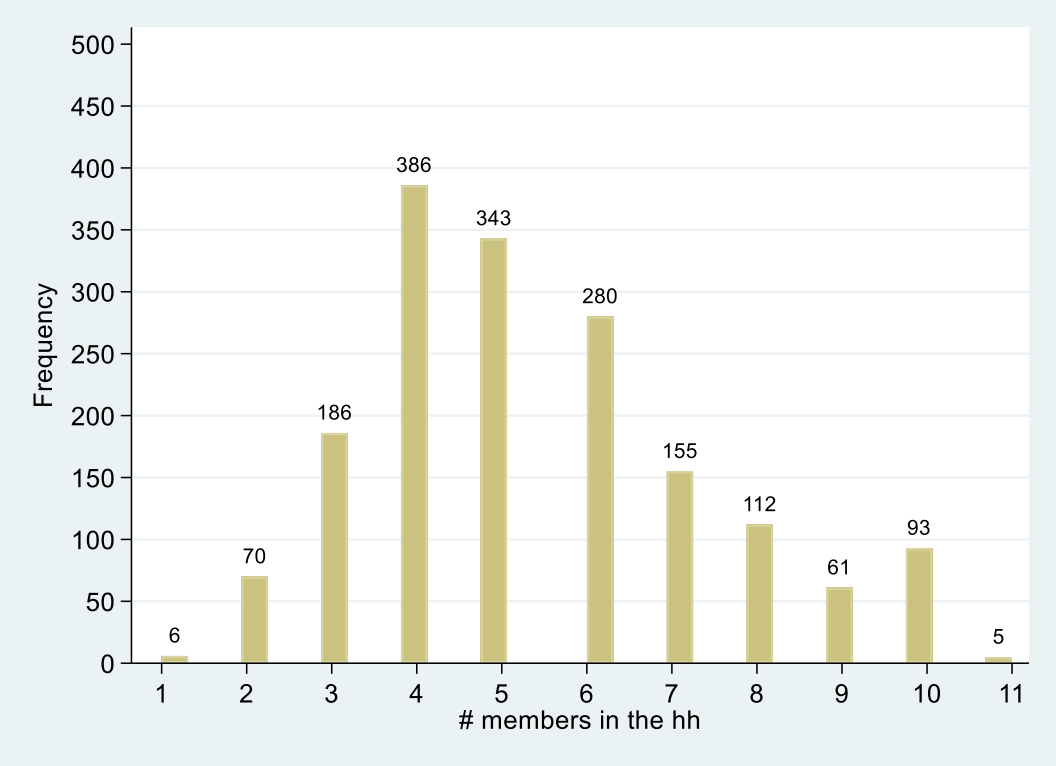
[3] When the last column with the total number of observations is not included, we provide the sample size for all the listed variables in the table notes. For the outcomes, we omit reporting the overall mean, as the variables represent standard deviations from the comparison group, and they do not have a practical economic meaning.

[4] The procedure behind the Holm-Sidak test is to first find all the p-values for all of the individual tests we were performing and then rank them from smallest to largest. We compare the smallest to $\alpha=\alpha_T/k$ where α_T is the Type I error rate. If we fail to reject the null hypothesis for the first step, then we stop here. If we reject it, then we compare the next smallest to $\alpha=\alpha_T/(k-1)$. Again, we stop here if we fail to reject the null hypothesis; if we do reject it, we continue and use $\alpha=\alpha_T/(k-2)$.

4.2 Demographics

26. The sample contains 1 697 households, with 840 in the cash-for-work treatment group and 857 in the comparison group. The median household size is five people, with a standard deviation of 2.03 for the average size of 5.4 persons per household. The distribution of households by size is presented in Figure 2, where we can observe that most households are comprised of three to seven members.

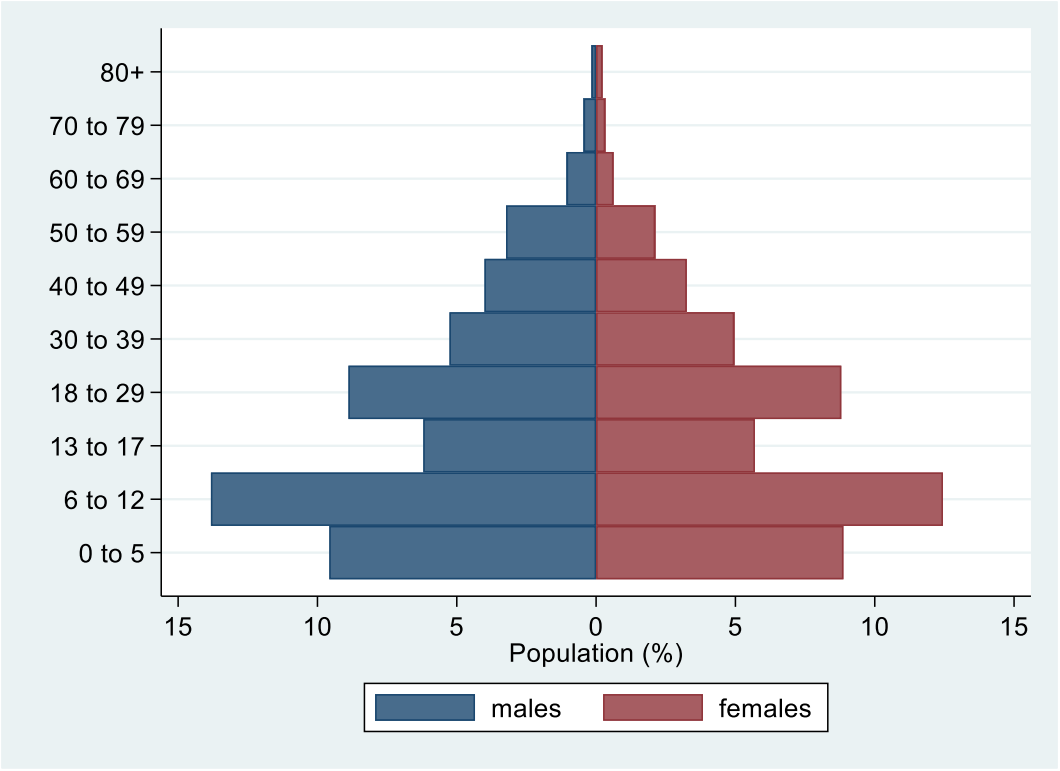
Figure 2. Household size distribution



Note: HH=household
Source: Authors' own elaboration.

27. Figure 3 shows the distribution by age and gender of the evaluation sample. There are a couple of features worth mentioning concerning the demographic distribution of these households. First, both men and women are well represented in the sample, although more men were sampled for the study in virtually all age brackets. Second, the presence of a relatively large share of people in working age and a minor share of elderly people. This suggests that targeting was successful, as the cash-for-work programme focused on both men and women with labour capacity.

Figure 3. Sample distribution by age and gender



Source: Authors' own elaboration.

28. Table 2 shows the mean sociodemographic characteristics of the sample. Despite the lack of randomization, we observe a remarkable similarity between the two groups. In fact, at the 5 percent significance level, only the age of the main respondent and the number of household members aged 0-17 are statistically different. Overall, households report approximately 2.1 adult members (1.1 males and 1 females). Households lacking adult members are only 1 percent of the sample. The share of dependents in the households is about 56 percent, whereas the cash-for-work group report a slightly higher share than the comparison group (57.6 percent *vis-à-vis* 54.5). The majority of the main respondents are male (63 percent) and they are married (89 percent), with low levels of education (only one completed year of education irrespectively of treatment status).

29. The disability indicator provides another indication of targeting attainment. We measured disability for household members aged 15 and above, following the Washington Group Disability Questions. The indicator included in Table 2 corresponds to the number of adult household members reporting moderate or severe disability.[5] Besides the substantial equality across treatment arms, we notice that this number is extremely low. The average 0.173 of at least moderately disabled members are approximately equivalent to one out of ten households reporting to have people with significant health issues. We analysed disability also by looking at the individual-level information, which we included in Table 3. Following the taxonomy created by the Washington Group, 89 percent of the members aged 15 years and above were labelled

[5] The six questions enquire about difficulties members may experience when performing certain activities due to a health problem. Specifically, they refer to difficulty: i) seeing; ii) hearing; iii) walking/climbing steps; iv) remembering or concentrating; v) with self-care, such as washing all over or dressing; vi) communicating. Each question has four response categories, which are read after each question: i) no difficulty; ii) some difficulty; iii) a lot of difficulty; iv) cannot do it all.

without disability, since they responded "*no difficulty*" to all six functioning domains. Those who responded "*cannot do at all*" to any functioning domain were labelled with severe disability and they represent 6 percent of the adult member's sample. We also created a severe disability score, assigning a relatively greater weight to those who responded either "*a lot of difficulty*" or "*cannot do it all*" to any functioning domain. However, because of the polarization of the answers, especially towards the "*no difficulty*" category, this score showed very little variability. Finally, and irrespectively of the indicator chosen, we notice no significant different between the cash-for-work and comparison groups.

Table 2. Average household characteristics by treatment group

Variables	All	Cash for work	Comparison	Difference	P-value
# members in the household	5.397	5.517	5.279	0.238	0.211
# household members 0-17 yrs	3.056	3.198	2.917	0.280	0.048
# male household members 18-59 years	1.154	1.121	1.187	-0.065	0.253
# female household members 18-59 years	1.032	1.046	1.019	0.028	0.684
# household members 60+ years	0.154	0.151	0.156	-0.005	0.842
Equivalence scale	3.361	3.399	3.323	0.076	0.487
# disabled household members	0.173	0.169	0.176	-0.007	0.888
Households with only elderly (>59) and children (<18)	0.011	0.010	0.012	-0.002	0.693
Share of dependents in household	0.560	0.576	0.545	0.031	0.087
Female main respondent	0.373	0.387	0.359	0.028	0.436
Main respondent age	38.379	36.381	40.338	-3.957	0.000
Main respondent is married	0.893	0.890	0.896	-0.006	0.766
Main respondent yrs of education	1.151	1.113	1.189	-0.076	0.752

Note: The total sample for all variables included is 1 697 observations.
Source: Authors' own elaboration.

Table 3. Adult members (15+) disability status by treatment group

Variables	All	Cash for work	Comparison	Difference	P-value
No disability	0.893	0.895	0.891	0.004	0.863
Mild disability	0.033	0.032	0.034	-0.002	0.757
Moderate disability	0.013	0.010	0.015	-0.006	0.223
Severe disability	0.061	0.063	0.059	0.004	0.864
Disability score	12.372	12.474	12.274	0.200	0.967

Note: The total sample for all variables included is 3 973 observations (adult members aged 15 and above, responding to the six Washington Group Disability Questions).
Source: Authors' own elaboration.

4.3 Goal and outcome indicators

30. In this section we describe indicators that correspond to the goal and outcomes included in the project's logframe. This will allow to assess the programme against its main stated objectives, namely: resilience, food security, income diversification, agency and social cohesion.

31. Many of these outcomes are multidimensional concepts and many indicators can be used to measure them. Testing the relationship between the programme and each of these indicators can be problematic for two reasons: i) testing many relationships at the same time may suffer from the multiple comparison problem described earlier; and ii) it prevents generalizing about the

effects of the programme on a specific outcome area. Both potential limitations are addressed by using a summary index approach – that is, by using summary indices for the main outcome variables, which facilitates generalizations about programme effectiveness.

32. The main criticism related to the summary index approach is that these indexes are a "black box," failing to unpack the programme's mechanisms (Schwab *et al.*, 2020). However, using this approach does not prevent researchers from getting deeper into the details, reporting and discussing the individual indicators. This is reflected in the most recent development literature, which disentangles summary index findings (Haushofer and Shapiro, 2016; Janzen *et al.*, 2021). In this section of the baseline study, we will show the balance for the summary indexes, while their single components will be discussed in separate sections.

33. We calculate summary indexes by adopting the standardized weighted mean approach of Anderson (2008), where we use the comparison group as the default reference group for the standardization.[6] These standardized summary indexes following Anderson do not have a specific meaning as they merely reflect deviations from the comparison group and can be thus interpreted as effect sizes.

4.3.1 Resilience index

34. The resilience index is a comprehensive measure that evaluates a household's ability to withstand and recover from shocks caused by various factors. In this study we consider three specific shocks: droughts, floods and pests/diseases. Therefore, the summary index is a standardized weighted average of three resilience capacity scores (RCS), which are derived from the nine capacities model developed by Jones and d'Errico (2019). The nine resilience-related capacities are: i) absorptive capacity; ii) transformative capacity; iii) adaptive capacity; iv) financial capital; v) social capital; vi) political capital; vii) learning; viii) anticipatory capacity; ix) early warning. Each of the three RCS are calculated from nine substatements using a five-point Likert scale (ranging from "strongly disagree" to "strongly agree") to capture the household perception of existing resilience capacities (Appendix table 1). The statements are numerically converted (Strongly disagree=1; Disagree=2; Neutral=3; Agree=4; Strongly agree=5) and then used to compute an overall resilience score for each household as an equally weighted average of the nine answers. The RCS is standardized by min/max normalization, transforming the results in a score that ranges from 0 (not at all resilient) to 100 (fully resilient):

$$RCS_x = \frac{\frac{\sum_{i=1}^{9} Q_{xi}}{9} - \min(Q)}{\max(Q) - \min(Q)} \times 100 = \frac{\frac{\sum_{i=1}^{9} Q_{xi}}{9} - 1}{5 - 1} \times 100$$

35. An alternate resilience score was also computed using the resilience index measurement and analysis (RIMA) approach (FAO, 2016). RIMA is context-and-shocks-specific and estimates household resilience to food insecurity with a quantitative approach to establish a cause-effect relationship between resilience and its critical determinants.

4.3.2 Food security index

36. The food security index is a standardized weighted average of the (positively coded) food consumption score and household dietary diversity score; and the (negatively coded) coping strategies index.

[6] See Schwab *et al.* (2020) for a detailed step-by-step guide to construct such summary indexes following Anderson.

37. The food consumption score is one of the most commonly used food security indicators in protracted crises (WFP, 2008). It aggregates household-level data on the diversity and frequency of food groups consumed over the previous seven days, which is then weighted according to the relative nutritional value of the consumed food groups. For instance, food groups containing nutritionally dense foods, such as animal products, are given greater weight than those containing less nutritionally dense foods, such as tubers. Based on this score, a household's food consumption can be further classified into one of three categories: poor, borderline, or acceptable. The food consumption score is a proxy indicator of household caloric availability.

38. The household dietary diversity score can be described as the number of food groups consumed by a household over a given reference period and is an important indicator of food security for many reasons. A more diversified household diet is correlated with caloric and protein adequacy, percentage of protein from animal sources, and household income (Swindale and Bilinsky, 2006). The household dietary diversity score provides a glimpse of a household's ability to access food as well as its socioeconomic status based on consumption during the previous 24 hours (Kennedy *et al.*, 2010).

39. The coping strategies index is a tool that measures what people do when they cannot access enough food. It consists of a series of questions about how households manage to cope with a shortfall in food for consumption, and results in a simple numeric score. It can be used for a variety of purposes. These include providing a quick, current status indicator of the extent of food insecurity and measuring the impact of food assistance programmes (Maxwell and Caldwell, 2008).

4.3.3 Income diversification index

40. The income diversification index is a standardized weighted average of the (positively coded) number of:

i. income sources;

ii. months in i) a food retail enterprise; ii) a non-food retail business; iii) a processing enterprise (such as a restaurant or carpentry shop); and iv) a business in the services sector;

iii. crops harvested;

iv. crops sold;

v. livestock income sources (livestock heads sold and livestock by-products sold);

vi. different public transfer sources received;

vii. private transfer forms;

viii. months of adult wage labour.

4.3.4 Agency index

41. As a proxy for psychological well-being we construct a summary index of agency, using a reduced number of questions adopted for the internality, powerful others and chance scale proposed by Levenson (1981). It consists of three subscales:

i. "Internality" measures the extent to which a person believes their own actions influence their life;

ii. "Powerful Others" measures the extent to which an individual feels others influence their life; and

iii. "Chance" measures how much the individual feels luck or chance is driving their life (Appendix table 2).

42. For each of these subscales we generate a score, summing up the answers to nine statements (three for each subscale) using a four-point Likert scale (Strongly disagree=1; Disagree=2; Agree=3; Strongly agree=4). The final summary index of agency is a standardized weighted average of the (positively coded) internal locus of control score; and the (negatively coded) chance score and powerful other score.

4.3.5 Social cohesion index

43. The social cohesion index is calculated as a standardized weighted average of the following (positively coded) indicators:

i. Trust score: It refers to the level of trust the main respondent has in their friends, coworkers, neighbours, people in the village, community leaders, police, political parties and humanitarian organizations (1=Little / no trust; 2=Some trust; 3=A lot of trust).

ii. Attitudes score: It relates to the attitudes of the main respondent *vis-à-vis* people in their village. These attitudes are expressed in terms of agreement/disagreement about a series of statements (Strongly disagree=1; Disagree=2; Neutral=3; Agree=4; Strongly agree=5), which refers to people behaviour (Appendix table 3).

iii. Help score: It indicates the help provided by the neighbours or the people living in the village on a range of domestic chores. Specifically, it refers to the number of days spent in the week prior to the survey helping the household in the following tasks: i) cooking meals; ii) cleaning the dwelling/clothes; iii) fetching water and/or firewood; iv) buying food.

iv. Group score: It encompasses the main respondent's and their family member's participation in the following groups/associations: i) farmers; ii) women support; iii) youth; iv) business association; v) religious; vi) savings.

v. A binary variable that takes a value of 1 if the household receives food or cash transfers from other people within the community, and 0 otherwise.

vi. A binary variable that takes a value of 1 if the household makes food or cash transfers to other people within the community, and 0 otherwise.

44. The four scores (trust, attitudes, help, group) are constructed following the min/max approach already discussed.

45. Table 4 describes the main outcome variables by treatment group. Since the summary indexes are standardized relatively to the comparison group, the comparison group mean equals to zero (and the unreported standard deviation is equal to one). While the reported mean value for the cash-for-work does not have an immediate interpretation, a positive (negative) sign is indicative of larger (smaller) values of the index relative to the comparison group. So, for instance, the summary index of food security for the cash-for-work group equals 0.084, meaning that this index is 0.084 standard deviations larger than in the comparison group. Despite the lack of randomization, the two treatment groups are perfectly balanced with respect to all the summary indexes, with the cash-for-work group showing statistically insignificant larger values for all outcomes, with the exception of the agency index.

Table 4. Main programme outcomes by treatment group

Summary index	Cash for work	Comparison	Difference	P-value	Observations
Resilience	0.0113	0.0000	0.0113	0.9407	1696
Food security	0.0843	0.0000	0.0843	0.4271	1697
Income diversification	0.0027	0.0000	0.0027	0.9793	1697
Agency	-0.0183	0.0000	-0.0183	0.8444	1691
Social cohesion	0.0863	0.0000	0.0863	0.4690	1697

Note: Difference=mean difference between cash for work & comparison group.
Source: Authors' own elaboration.

4.4 Resilience

46. Table 5 provides descriptive evidence concerning the three RCS measures, which compose the summary index, while we also look also at FAO's RIMA resilience capacity index. The cash-for-work group reports higher values for all these indicators, though these differences are statistically insignificant. These results are reinforced by the non-parametric Wilcoxon rank-sum tests of the equality of mean ranks (unreported).

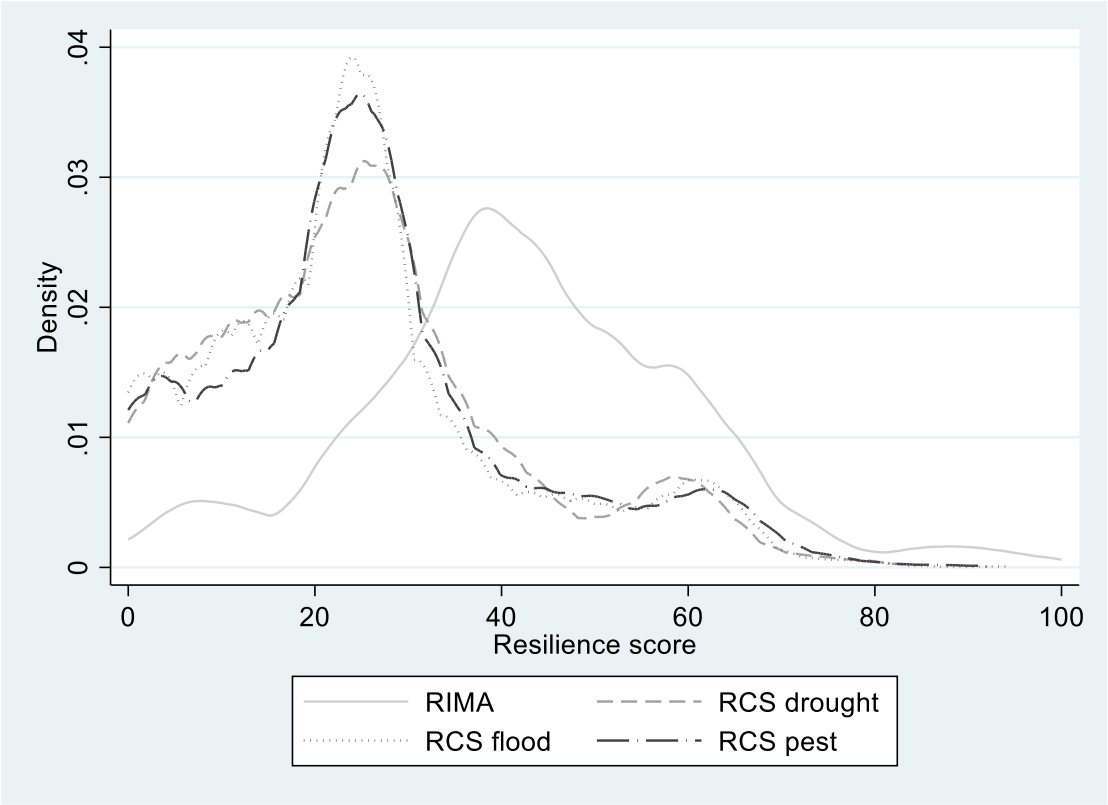
Table 5. Resilience measurements by treatment group

Variables	All	Cash for work	Comparison	Difference	P-value	N
Summary index of resilience	0.006	0.011	0.000	0.011	0.941	1696
Resilience capacity score for drought	25.202	25.331	25.076	0.256	0.919	1697
Resilience capacity score for flood	24.805	24.934	24.679	0.255	0.923	1696
Resilience capacity score for pests	25.841	25.901	25.783	0.118	0.964	1697
RIMA resilience capacity index	42.519	43.022	42.027	0.995	0.691	1697

Note: Difference=mean difference between cash for work & comparison group; N=observations; RIMA=resilience index measurement and analysis.
Source: Authors' own elaboration.

47. The three RCS measures are highly correlated among each other, with a correlation coefficient (ρ) ranging between 0.88 and 0.91. The correlation between RIMA and the RCS measures is instead lower, with ρ≈0.25. Further, the RIMA index has much greater variability, as shown by the kernel densities in Figure 4, with a mode at a value of 25, and are skewed to the right.

Figure 4. Kernel density of resilience scores



Note: RIMA= resilience index measurement and analysis; RCS=resilience capacity score
Source: Authors' own elaboration.

4.5 Food security

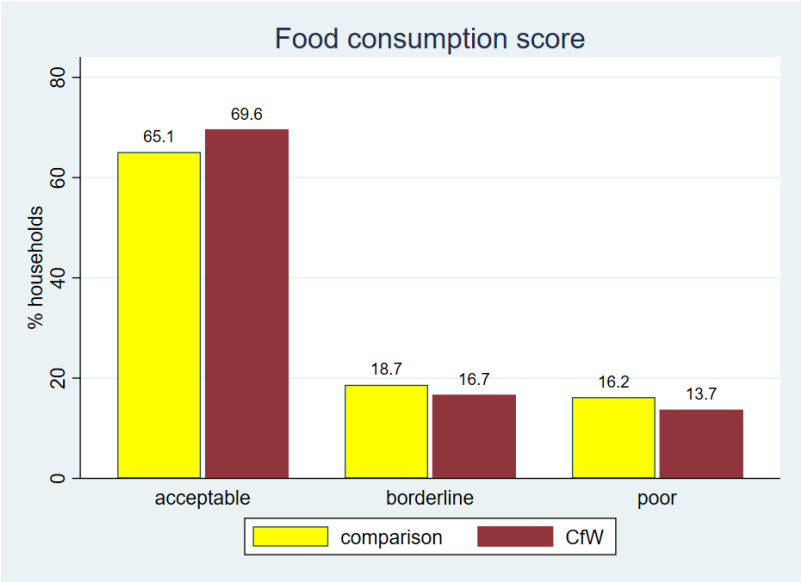
48. Somali population is currently facing extreme levels of food insecurity. As mentioned in section 1, 6.5 million people are facing crisis (IPC Phase 3) or worse acute food insecurity outcomes in 2023. Table 6 shows the mean values for the three food security variables that compose the summary index used as outcome indicator in this study by treatment group. Surprisingly, given the nature of the evaluation design, we do not find any statically significant difference between the comparison and the cash-for-work group means. If any, the cash-for-work group reports a larger share of households with an acceptable food consumption score (i.e. with a value >35) than the comparison group (see Figure 5). As opposed to this, we observe a slightly larger share of comparison households in a borderline or poor food consumption score status, though these two last differences are not statistically significant.

Table 6. Food security status by treatment group

Variables	All	Cash for work	Comparison	Difference	P-value	N
Summary index of food security	0.042	0.084	0.000	0.084	0.427	1697
Food consumption score	47.717	48.162	47.280	0.882	0.797	1697
Household dietary diversity score	5.296	5.344	5.249	0.096	0.744	1697
Coping strategies index	21.845	21.549	22.135	-0.587	0.720	1697

Note: Difference=mean difference between cash for work & comparison group; N=observations.
Source: Authors' own elaboration.

Figure 5. Food consumption score categories by treatment group



Note: CfW=cash-for-work
Source: Authors' own elaboration.

4.6 Income and livelihoods

49. Table 7 shows the baseline evidence concerning the labour and non-labour income variables collected in the survey. Most of these indicators are substantially equal across treatment arms. The only important difference, which is statistically significant at the 5 percent level, refers to the number of months worked by the adult household members in wage labour. This result is not surprising, because the cash-for-work activities started three weeks prior to the beginning of the survey fieldwork. This result also drives the statistically significant difference in the number of income sources across the two treatment arms: households in the cash-for-work group report on average 2.3 income sources, whereas those in the comparison group have 1.8 only.

Table 7. Income variables by treatment status

Variables	All	Cash for work	Comparison	Difference	P-value
summary index of income diversification	0.001	0.003	0.000	0.003	0.979
# income sources	2.083	2.312	1.859	0.453	0.001
# months food-retail business in operation	0.207	0.170	0.244	-0.074	0.516
# months non-food retail business in operation	0.074	0.094	0.054	0.040	0.283
# months processing/manufacturing business in operation	0.019	0.033	0.005	0.029	0.159
# months services business has been in operation	0.127	0.127	0.126	0.001	0.986
# crops harvested after 2022 *gu* season	0.314	0.314	0.314	0.000	0.997
# crops sold after 2022 *gu* season	0.122	0.129	0.116	0.013	0.833
# of forms of transfers from public institutions received	0.942	1.051	0.834	0.217	0.126
# livestock income sources	0.523	0.602	0.446	0.157	0.118
# of forms of private transfers received	0.587	0.624	0.551	0.073	0.518
# months last year adult HH members worked in wage labour	1.631	2.335	0.942	1.393	0.012

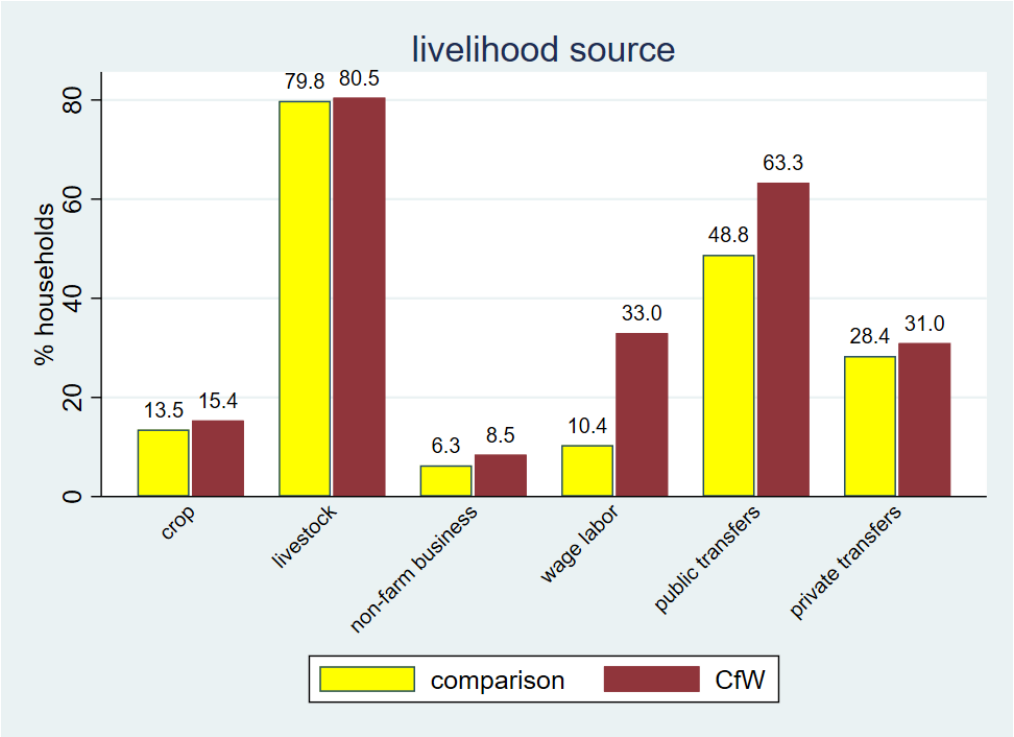
Note: Difference=mean difference between cash for work & comparison group. HH=household. The total sample for all variables included is 1 697 observations.
Source: Authors' own elaboration.

50. The same livelihood pattern is observed in Figure 6, where we show the share of households for each income source. While for crop and livestock production, non-farm businesses and private transfers there is no discernible difference between the two treatment arms, households in the cash-for-work group are more likely to be engaged in wage labour and to receive some forms of public transfer than those in the comparison group. One striking statistics in this graph concerns the proportion of households raising livestock, which is around 80 percent in both groups. This finding seems to contradict the low average number of livestock income sources reported in the previous table. However, this is probably indicative of a high share of households owning/herding livestock as a form of precautionary saving or for their own consumption of milk and by-products, with very little engagement with markets.

51. Livestock ownership is explored in greater detail in Table 8, where we report several intensive margins indicators. Tropical Livestock Units (TLU) indicate the number of animals owned by the household, which are weighted to aggregate the different species of animals using a common unit (typically a cattle with a body weight of 250 kg).[7] As shown in the table, we notice that on average the TLUs in this sample are 5.8. This is equivalent to a small-size herd for cattle and camel owners, and a medium-size herd for goats and sheep owners. These results are confirmed by looking at the number of animals by livestock species. On average, sample households own 23 goats and ten sheep, and approximately two camels. Raising poultry is also quite uncommon in the project area, with only 0.5 chicken owned by the households. For all the livestock indicators, we also notice lack of statistically significant differences between the cash-for-work and the comparison group.

[7] In this study we use the following conversion factors: Camels=1; horses=1; cattle=0.7; donkeys=0.6; sheep=0.1; goats=0.1; poultry=0.01. The results do not change if we adopt different weighting schemes normally adopted in the literature, giving more weight to the cattle.

Figure 6. Livelihood source categories by treatment group



Note: CfW=cash-for-work
Source: Authors' elaboration from survey data.

Table 8. Livestock ownership by treatment status

Variables	All	Cash for work	Comparison	Difference	P-value
Tropical Livestock Units	5.8	5.5	6.1	-0.6	0.504
Households owning livestock	0.801	0.805	0.798	0.007	0.865
# cattle	0.5	0.6	0.4	0.2	0.500
# sheep	10.2	8.9	11.4	-2.5	0.215
# goats	23.7	22.0	25.2	-3.2	0.529
# camels	1.9	1.8	2.0	-0.2	0.629
# donkeys	0.3	0.3	0.3	0.0	0.717
# chicken	0.5	0.6	0.5	0.1	0.513

Note: Difference=mean difference between cash for work & comparison group. The total sample for all variables included is 1 697 observations. Households not owning livestock are included in the average statistics reported for the Tropical Livestock Units and the number of animals.
Source: Authors' own elaboration.

4.7 Psychological well-being

52. Locus of control is the degree to which people believe that they have control over the outcome of events in their lives, as opposed to external forces beyond their control. The original scale proposed by Levenson consisted of 24 questions, with eight questions in each of the three subscales (chance, powerful others and internal locus of control). In this paper, we use a pared down version of the scale, with three questions in each subscale, and therefore a total of nine. Some of the questions were in fact irrelevant for a poor rural setting in Somalia, while at the same time it allowed to keep the survey instrument at a manageable length. Table 9 reports the summary statistics of these constructs by treatment arm. As for the summary index, the three

subscales of chance, powerful others and internal locus of control do not show any significant differences among the two intervention groups. We find the same result also if we look at the agency score, which elaborates the same questions used for the subscales by adopting the min/max approach used for the RCS.

Table 9. Agency variables by treatment status

Variables	All	Cash for work	Comparison	Difference	P-value	N
Summary index of agency	-0.009	-0.018	0.000	-0.018	0.844	1691
Agency: chance	53.922	54.848	53.014	1.833	0.366	1697
Agency: powerful others	61.495	61.654	61.339	0.315	0.873	1692
Agency: internal locus of control	59.408	58.403	60.392	-1.989	0.352	1696
Agency score	58.311	58.352	58.272	0.080	0.936	1697

Note: Difference=mean difference between cash for work & comparison group.
Source: Authors' own elaboration.

53. Interestingly, if we look at the same variables by the gender of the main respondents, we notice that there are some statistically significant differences between males and females at the subscale level, though these are difficult to interpret. Male respondents have higher internal locus of control, which aligns with conventional wisdom that men can exert greater control over their lives through their own individual actions. However, they also report a lower value for both the chance and the powerful other subscale.[8] Therefore, the overall agency score is higher for the female respondents, even though only at 10 percent statical significance level.

Table 10. Agency variables by gender of the main respondent

Variables	All	Female	Male	Difference	P-value	N
Summary index of agency	-0.009	-0.021	-0.002	-0.019	0.788	1691
Agency: chance	53.922	55.722	52.851	2.871	0.022	1697
Agency: powerful others	61.495	64.161	59.906	4.256	0.002	1692
Agency: internal locus of control	59.408	56.688	61.027	-4.340	0.002	1696
Agency score	58.311	58.929	57.944	0.985	0.181	1697

Note: diff=mean difference between males & females' group.
Source: Authors' elaboration from survey data.

4.8 Social cohesion

54. Social cohesion is a multidimensional concept and its measurement at both the micro and macro level is therefore a result of an aggregate of several indicators. As mentioned by Valli *et al.* (2019), the framework provided by Babajanian (2012) is probably the most comprehensive in terms of its relationship with social protection programming, with a focus on domains such as empowerment, community operation, solidarity and social participation. In this study the operational definition of social cohesion follows to a broad extent that approach. We exclude some components unlikely to be affected by social protection, such as conflict and stability or shared values and goals, and complement it with a set of questions dedicated to interpersonal and institutional trust. As shown in Table 11, the treatment and the comparison groups look substantially identical in most outcomes. The only variable where we observe a statistically significant difference is the groups' participation score, which summarizes the participation of the main respondents and/or family

[8] The responses for the chance and powerful others questions were rescaled, such that a higher value for the subscale means greater agency.

members in any of six categories of groups/associations (see subsection 4.3). For this variable, the main respondents included in the cash-for-work group suggest their households have a higher chance of participation in any of these networks. This result is not completely surprising, probably households that participate in these groups have better connections with the village leaders and therefore capitalize on their acquaintances and knowledge to enrol in the cash-for-work programme. One additional interesting result from this table concerns two variables reflecting the occurrence of private transfers within the community/village. Irrespective of the (lack of) statistical difference between the two treatment arms, the share of households that reported receiving or making a transfer in cash or in-kind is non-trivial (25 vs 34 percent respectively). This is indicative that even in a context of high poverty, which is subject to severe weather shocks, cash and food sharing represent not only a way to distribute resources within a community, which strengthens the nature and quality of interpersonal and social relations, but also a significant form of coping strategy for many families.

Table 11. Social cohesion by treatment status

Variables	All	Cash for work	Comparison	Difference	P-value
Summary index of social cohesion	0.043	0.086	0.000	0.086	0.469
People's attitudes score	67.524	67.180	67.860	-0.680	0.589
People's help score	1.624	1.727	1.523	0.205	0.428
Trust score	56.320	55.420	57.202	-1.782	0.469
Group's participation score	0.570	0.685	0.459	0.226	0.029
Private transfers received	0.256	0.261	0.252	0.009	0.866
Private transfers made	0.341	0.343	0.340	0.003	0.947

Note: Difference=mean difference between cash for work & comparison group. The total sample for all variables included is 1 697 observations.
Source: Authors' own elaboration.

55. Similarly to the agency domain, we look at the same indicators by gender of the main respondent (Table 12). Most of the components do not show any discernible difference between male and female respondents. Interestingly, the only indicators where we observe some statistically significant differences are the two variables indicating whether the household receives or makes a cash or in-kind transfer within the community. Considering that the main respondent is probably the main decision-maker within the household, this is indicative of a lower potential for women to contribute to these networks of reciprocity.

Table 12. Social cohesion by gender of the main respondent

Variables	All	Female	Male	Difference	P-value
Summary index of social cohesion	0.043	0.019	0.057	-0.038	0.557
People's attitudes score	67.524	68.197	67.123	1.074	0.213
People's help score	1.624	1.670	1.597	0.073	0.688
Trust score	56.320	56.615	56.145	0.470	0.730
Group's participation score	0.570	0.518	0.602	-0.083	0.126
Private transfers received	0.256	0.227	0.273	-0.046	0.068
Private transfers made	0.341	0.297	0.367	-0.070	0.017

Note: diff=mean difference between males & females' group. The total sample for all variables included is 1 697 observations.
Source: Authors' elaboration from survey data.

5. Baseline data econometric analysis

56. In this section of the report, the focus is to understand the relationship between: i) resilience and the four socioeconomic outcomes that in our theoretical framework represent its immediate determinants; ii) resilience and climate shocks, which influence the humanitarian response in the country.

5.1 Relationship between resilience and outcomes

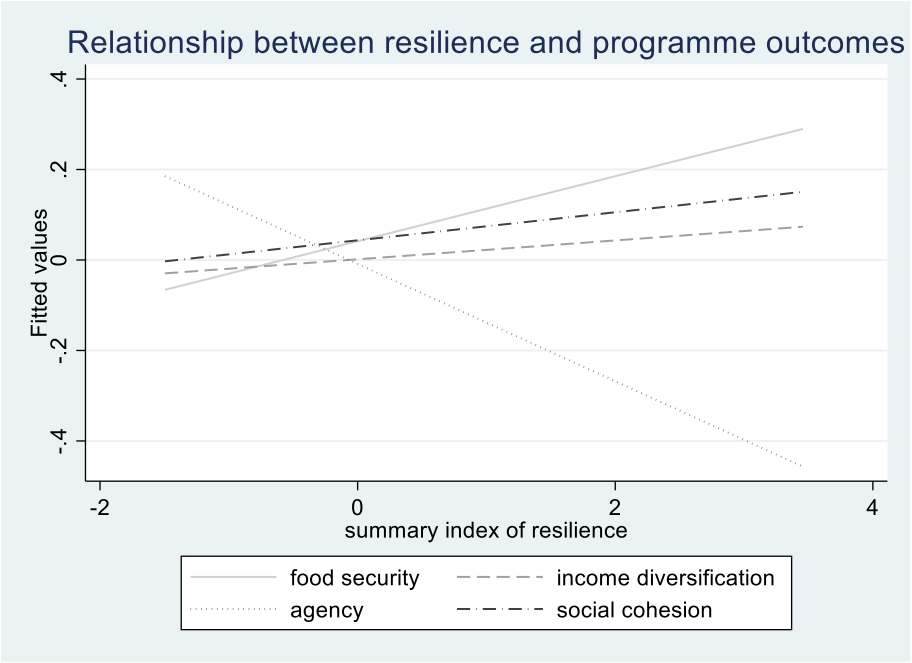
57. We use a simple ordinary least squares (OLS) regression framework to estimate the associations between resilience, measured as an index of three RCS as described in section 4.3, and its most immediate socioeconomic determinants identified in the theoretical framework: food security, income diversification, agency and social cohesion. The OLS regression can be expressed as follows:

$$Res_{iv} = \alpha + \boldsymbol{\beta Y}_{iv} + \boldsymbol{\gamma X}_{iv} + \mu_r + \varepsilon_{iv} \qquad (1)$$

where Res_{iv} is the resilience index for household i living in village/cluster v. **Y** is the vector of socioeconomic determinants, and **β** is the vector of its associated coefficients. **X** is a vector of control variables containing: i) the number of household members in different age brackets (0-5, 6-12, 13-17, 18-59 women, 18-59 men, 60+); ii) sociodemographic characteristics of the main respondent (age, gender, marital status and whether he/she has completed primary education); and iii) indicators of wealth (TLUs, summary index of agricultural assets, summary index of access to basic services). Finally, μ_r represents regional fixed effects.

58. Before analysing the findings of the econometric analysis, we briefly have a look at the relationships between resilience and each outcome one-by-one. As shown in Figure 7, three outcome indicators have the expected positive, albeit weak, correlation with the summary index of resilience. The agency index instead has a strong and negative correlation with resilience. This is a counterintuitive finding, since one would expect that the ability of an individual to make a positive adjustment to adverse events is directly connected to their control over the outcomes of events in their life. We analyse these relationships in a formal econometric framework, estimating equation (1). The weak association between three outcomes (food security income diversification and social cohesion) and resilience is substantially confirmed, as we find mostly positive and statistically not significant coefficients, irrespective of the model chosen (Table 13). However, we confirm the strong negative correlation between agency and resilience. For this reason, we unpack the agency index into the three main subcomponents: chance, powerful others and locus of control. We notice that the first two components are negatively related to resilience, while the locus of control subindex is positively associated. Therefore, survey respondents show greater resilience, not only if they believe that outcomes in their lives are determined by their own actions, but also by chance and other powerful people.

Figure 7. Relationship between resilience and programme outcomes



Source: Authors' elaboration from survey data.

Table 13. Correlation between subjective resilience and programme outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome variables						
Food security	0.0665	0.0637	0.0708	0.0837*	0.00567	0.00182
	(1.75)	(1.59)	(1.74)	(2.17)	(0.16)	(0.06)
Income diversification	0.0165	0.0211	0.0263	-0.00193	-0.0376	-0.0397
	(0.41)	(0.53)	(0.65)	(-0.05)	(-1.22)	(-1.63)
Agency	-0.133***	-0.130***	-0.131***	-0.115***	-0.106**	
	(-4.11)	(-4.03)	(-4.17)	(-3.82)	(-2.94)	
Social cohesion	0.0442	0.0307	0.0241	0.0113	0.138**	0.0544
	(0.82)	(0.59)	(0.47)	(0.24)	(3.34)	(1.57)
Agency: chance						-0.00822***
						(-3.84)
Agency: powerful others						-0.00753***
						(-4.26)
Agency: locus of control						0.00681***
						(4.36)
Control variables						
Demographic	No	Yes	Yes	Yes	Yes	Yes
Main respondent	No	No	Yes	Yes	Yes	Yes
Wealth	No	No	No	Yes	Yes	Yes
Regions	No	No	No	No	Yes	Yes
N	1690	1690	1690	1690	1690	1690

Note: statistical significance: *** p<0.01; **p<0.05, * p<0.1. The dependent variable is the resilience capacity score.
Source: Authors' elaboration from survey data.

5.2 Relationship between resilience and climate

59. Climate shocks like droughts and floods are the hallmark of Somalia's prolonged humanitarian crisis. These extreme weather events disrupt livelihoods, especially agricultural activities, affecting household's food security and resilience. In this section, we seek to understand sensitivity, a specific component of climate vulnerability, which indicates the degree to which a system is either positively or negatively affected by climatic stresses (IPCC, 2012). More specifically, to determine the effects of climate on household's resilience, we would like to assess the following relationship:

$$\text{Res} = f(\mathbf{C}, \mathbf{X}) \tag{2}$$

linking a vector of climatic variables (**C**) and other variables (**X**) to resilience, which is measured at the household level. The vector **C** may include rainfall precipitations, and other weather events. The vector **X** includes any characteristics that are correlated with **C** and affect resilience, possibly by conditioning the climate response. In the absence of panel data, the simplest approach to estimate equation (2) exploits cross-sectional variation in climate (occurring at one point in time) to estimate the marginal effect of long-run changes in the distribution of rainfall (Mendelsohn *et al.*, 1994):

$$y_{il} = \alpha + \beta \boldsymbol{C_{lc}} + \gamma \mathbf{X}_{il} + \varepsilon_{il} \tag{3}$$

where i indexes households, l geographic locations, e.g. regions/districts where the households live.

60. An important empirical question involves the choice of weather variables and their functional form. Since we are interested in both the long-run effects of changes in climate and short-run effects of weather shocks, we follow what Kolstad and Moore (2020) defined the "partitioning variation" approach, which decomposes the variation in the outcome variable into the part associated with climate and the part associated with inter-annual fluctuations around the climatological average. We use the word climate to describe the distribution of outcomes, while weather refers to a particular realization from that distribution (Dell *et al.,* 2014). Therefore, we assume that outcomes are a function of both long-run moving averages of and short-run standardized z-scores of rainfall precipitations. Formally, we rewrite equation (3) as:

$$y_{il} = \alpha + \beta_1 \overline{RG}_{ll^{40}} + \beta_2 \overline{RD}_{il^{40}} + \beta_3 \widetilde{RG}_{il^{40}} + \beta_4 \widetilde{RD}_{il^{40}} + \gamma \mathbf{X}_{il} + \varepsilon_{il}, \tag{4}$$

where $\overline{RG}_{il^{40}}$ and $\overline{RD}_{il^{40}}$ are the average rainfall precipitations in the gu and deyr seasons respectively, computed over the 40 years prior to the survey. $\widetilde{RG}_{il^{40}}$ and $\widetilde{RD}_{il^{40}}$ are instead the standardized weather realizations:

$$\widetilde{RG}_{il^{40}} = \frac{RG_{il} - \overline{RG}_{il^{40}}}{\gamma_{il^{40}}} \qquad \widetilde{RD}_{il^{40}} = \frac{RD_{il} - \overline{RD}_{il^{40}}}{\delta_{il^{40}}} \tag{5}$$

where RG_{il} and RD_{il} are the realized weather outcomes (rainfall precipitations during the last gu and deyr season) in location l, while $\gamma_{il^{40}}$ and $\delta_{il^{40}}$ are the long-run standard deviations of rainfall in the gu and deyr seasons, respectively. In equation (4) the dependent variable is a linear function of both climate and the deviation of weather from its long-run mean. Extensions of this model may entail: i) the formulation of a second-order Taylor approximation, by including quadratic and interaction terms to reflect nonlinearities (Kelly *et al.,* 2005; Moore and Lobell, 2014); ii) including weather realizations from previous years; iii) including interaction terms between weather realizations and indicator functions expressing weather shocks.

61. Equations (5) identify the standardized precipitation index (SPI) in the gu and deyr seasons, transforming the amount of precipitations to a standard Gaussian variate with zero mean and standard deviation of one. An important advantage of the SPI is that it can be computed at different time scales. This way it is possible to incorporate the influence of the past values of the variable in the computation enabling the index to adapt to the memory of the system under study. The magnitude of this memory is controlled by parameter scale. For example, a value of six would imply that data from the current month and of the past five months will be used for computing the SPI value for a given month. By default, all past data will have the same weight in computing the index. A 1-month SPI typically compares well to the percent of normal precipitation for the month, it is a short-term value and during the growing season can be important for correlation of soil moisture and crop stress. In the analysis we adopt the 3-month SPI, which provides a comparison of the precipitation over a specific 3-month period with the precipitation totals from the same 3-month period for all the years included in the historical record.[9] A 3-months SPI (SPI3) reflects short- and medium-term moisture conditions and provides a seasonal estimation of precipitation. A SPI3>1.5 typically reflects severely wet conditions and can be used to identify weather shocks such as floods, while a SPI3<-1.5 reflects severely dry conditions, identifying weather shocks such as droughts.[10]

62. Table 14 reports the mean and standard deviation for the SPI3 data two years prior to the data collection for both the deyr and the gu seasons.[11] We notice that the 2022 gu season has been particularly harsh for households relying on farming, since the average SPI3 was approximately -1.5. Further, there is a lot of heterogeneity across the regions. For instance, while on average the 2022 deyr season did not differentiate much from normal long-run conditions, households in Laas Caanood suffered heavy rainfall precipitations (SPI3=1.84). Similarly, the deyr 2021 had below-average rainfalls, with SPI3=-0.57, but in some regions like Borama precipitations were quite abundant (SPI3=1.02).

63. Finally, we look at the relationship between resilience, climate and weather shocks in a formal regression framework (Table 15). The dependent variable across the five columns is the subjective resilience score. In column 1 we followed equation (4) and included the gu and deyr long-run precipitations mean and the 2022 SPI3 in the regression, without controlling for any other observables. Results are mainly non-significant, though resilience is positively and significantly associated with rains in the deyr season. In column 2 we included the 2021 SPI3 too. In this model, both long-run means are positive and statistically significant, as expected. The 2021 SPI3 in the gu season is the other significant result.

[9] The 3-months SPI refers to the months of April, May and June for the gu season and to the months of October, November and December for the deyr season.

[10] A SPI3 between 1 and 1.5 is indicative of moderately wet conditions, while a SPI3 between -1 and -1.5 is also used to highlight moderately dry conditions.

[11] We extract rainfall precipitations from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), (USC, 2024).

Table 14. Three-months standardized precipitation index by year, season and region

Region	2021		2022	
	gu	deyr	gu	deyr
Borama	0.418	1.024	-1.087	-0.447
	(0.357)	(0.139)	(0.381)	(0.063)
Burco	0.505	0.192	-1.286	0.063
	(0.315)	(0.219)	(0.321)	(0.225)
Laas Caanood	0.309	-0.229	-0.45	1.842
	(0.402)	(1.073)	(0.276)	(0.605)
Xudun	0.009	-0.289	-0.888	0.797
	(0.283)	(0.671)	(0.543)	(0.721)
Ceerigaabo	0.137	0.321	-1.27	0.245
	(0.374)	(1.461)	(0.366)	(0.436)
Qardho	-0.518	-1.12	-2.269	-0.141
	(0.276)	(0.25)	(0.258)	(0.278)
Burtinle	0.48	-1.431	-1.033	0.567
	(0.158)	(0.091)	(0.449)	(0.599)
Eyl	-0.346	-0.993	-2.562	0.801
	(0.265)	(0.31)	(0.317)	(0.35)
Hobyo	-0.222	-1.293	-1.907	0.016
	(0.49)	(0.078)	(0.223)	(0.211)
Jariiban	0.501	-0.888	-1.543	0.266
	(0.363)	(1.06)	(0.349)	(0.358)
Dhuusamarreeb	0.012	-1.239	-1.917	0.019
	(0.108)	(0.204)	(0.119)	(0.086)
Cabudwaaq	0.475	-0.656	-1.261	0.065
	(0.255)	(0.165)	(0.258)	(0.093)
Cadaado	0.304	-0.046	-1.659	0.093
	(0.207)	(1.466)	(0.159)	(0.175)
Belet Weyne	-0.64	-1.36	-1.807	-0.013
	(0.225)	(0.014)	(0.092)	(0.092)
Total	0.102	-0.572	-1.496	0.298
	(0.291)	(0.514)	(0.294)	(0.307)

Note: Standard deviations in parentheses.
Source: Authors' elaboration from survey data and CHIRPS data.

64. However, the SPI3 is a difficult indicator to interpret because the extreme events occur on both sides of the distribution. Therefore, in columns from 3 to 5 we have included interaction terms between the SPI3s and the occurrence of severely wet or dry conditions, which are identified with an absolute value of the SPI3 above 1.5. We notice that once we include regional dummies (column 5) the gu long-run rainfall mean becomes non-significant. Further, the only result always significant across these specifications is the interaction term between the 2022 SPI3 in the deyr season and the occurrence of severely wet conditions. By looking at the summary statistics provided in Table 14, we notice that even if rains in 2022 were very close to the normal long-run mean, in the region of Laas Caanood the value of SPI3 was quite high, and above 1.8. This result entails that the subjective resilience score was negatively affected by a major flooding event that had occurred in the very recent past. Other important weather shocks did not have any major association with resilience if they occurred more than three months prior to the survey.

Table 15. Correlation between subjective resilience, climate and weather shocks

	(1)	(2)	(3)	(4)	(5)
gu long-run rainfall mean	-0.000183	0.0164*	0.0219**	0.0213**	0.0158
	(0.00595)	(0.00746)	(0.00755)	(0.00739)	(0.0134)
deyr long-run rainfall mean	0.0347***	0.0225**	0.0146	0.0128	0.00486
	(0.00714)	(0.00845)	(0.00920)	(0.00894)	(0.0136)
3-months SPI deyr 2022	-0.121	-0.0505	0.245	0.262	0.129
	(0.124)	(0.122)	(0.148)	(0.151)	(0.141)
3-months SPI gu 2022	-0.156	0.00929	0.461	0.473	0.200
	(0.110)	(0.116)	(0.328)	(0.330)	(0.290)
3-months SPI deyr 2021		-0.0655	-0.155	-0.147	-0.224
		(0.0812)	(0.118)	(0.117)	(0.184)
3-months SPI gu 2021		-0.645***	-0.732***	-0.695***	-0.163
		(0.165)	(0.160)	(0.154)	(0.160)
SPI3 gu 2022 * gu 2022 severely dry			-0.236	-0.241	-0.182
			(0.186)	(0.184)	(0.152)
SPI3 deyr 2022 * deyr 2022 severely wet			-0.467*	-0.468*	-0.693**
			(0.214)	(0.224)	(0.236)
SPI3 deyr 2021 * deyr 2021 severely dry			-0.219	-0.249	-0.0933
			(0.193)	(0.189)	(0.220)
SPI3 deyr 2021 * deyr 2021 severely wet			0.185	0.166	0.0680
			(0.124)	(0.122)	(0.302)
socio-demographic	No	No	No	Yes	Yes
regions	No	No	No	No	Yes
Observation	1696	1696	1696	1696	1696

Note: statistical significance: *** p<0.01; **p<0.05, * p<0.1. Robust standard errors in parentheses. SPI3: 3-months standardized precipitation index. The dependent variable is the resilience capacity score.
Source: Authors' elaboration from survey data.

6. Estimation strategy

65. The cash-for-work impact evaluation consists of longitudinal quasi-experimental design, with one treatment and one comparison group. The treatment group includes the population of households eligible for public works, which include able-bodied household members. Given the impossibility of village randomization, for the purpose of the evaluation, FAO has targeted extra villages that serve as comparison group, extracting them from the same districts where the cash-for-work takes place and where FAO could rehabilitate at least one infrastructure within a 6-km radius, including water catchments, feeder roads, irrigation canals and contour bunds.

66. We will estimate the cash-for-work impacts using the following difference-in-difference (DiD) statistical model that compares change in outcomes between baseline and follow-up and between treatment and comparison groups:

$$\Delta Y_{ij} = \alpha + \beta CfW_j + \gamma \boldsymbol{X}_{ij0} + \boldsymbol{\delta} + \boldsymbol{\varepsilon}_{ij}$$

where ΔY_{ij} is the difference in the outcome between baseline and follow-up for household i in village j, CfW_j is a binary variable taking the value 1 for treatment village clusters (0 otherwise). X_{ij0} is a vector of household characteristics to control for observable differences across households at the baseline, which could have an effect on Y. These factors are not only those for which some differences may be observed across treatment and control at the baseline, but also "good controls" which could have some explanatory role in the estimation of Y (Angrist and Pischke, 2009). δ is a vector of district dummies, while β represents the intention-to-treat estimator of the programme impact. The outcomes of interest are resilience, food security, income diversification, agency, and social cohesion indexes.

67. For estimation, we will cluster standard errors at the village level, which is the unit of programme implementation, to account for the intracluster correlation and provide consistent estimates of standard errors. To account for possible imbalances across treatment arms, we will combine the DiD model with propensity score matching (PSM). This approach of inverse probability weighting with regression adjustment has the convenient property of being doubly robust (Wooldridge, 2010): if either the propensity/treatment model or the outcome equation is correctly specified, the estimator will be consistent.

6.1 Power and attrition

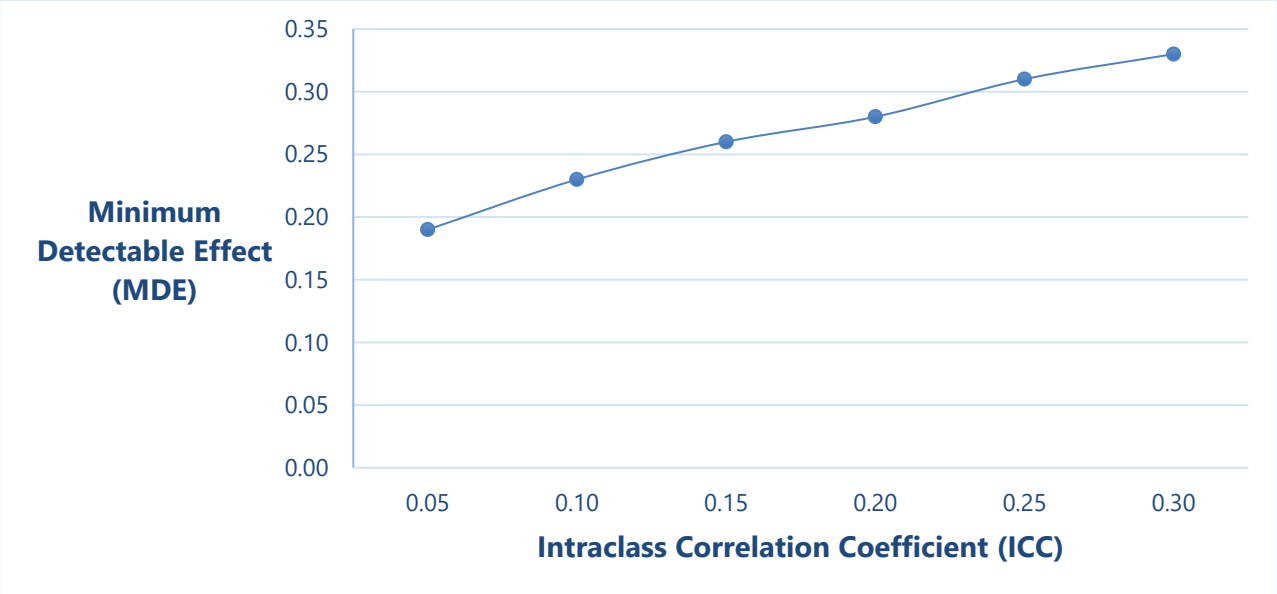
68. While carrying-out power calculations for randomized control trials is relatively straightforward, doing power calculations for quasi-experimental impact evaluations requires more judgement. Recently, guidance became available for regression discontinuity designs (Cattaneo *et al.,* 2019). For PSM and DiD designs, the literature is instead scarce. For PSM designs, McKenzie (2011) suggests to proceed as in the case of a balanced randomized control trial and take a control group sample that is 20–200 percent larger than in the pure experiment case. Intuitively, the problem with matching is that many units in the comparison group may lay outside the common support and cannot be used as "good" matches for the treated units.

69. As far as the DiD model is concerned, recently Hu and Hoover (2018) studied power / sample size estimation methods for non-randomized DiD designs. Further, Schochet (2022) developed new closed-form variance expressions for power analyses for commonly used DiD panel data estimators, with formulas also accounting for other key design features that arise in practice, such as autocorrelated errors, unequal measurement intervals, and clustering due to the unit of treatment assignment. Besides the cluster size, the number of clusters and the intracluster

correlation coefficient (ICC), other parameters must be considered in a DiD design. For the cash-for-work impact evaluation, we rule out staggered timing (so we have one treatment group occurring after the baseline), while we assume autocorrelation follows an AR(1) process with autoregressive parameter equal to 0.4. While the power literature in short panels assumes constant autocorrelations for pooled estimators (Frison and Pocock 1992; McKenzie 2012), here correlations are larger for cluster observations closer in time than further apart. We do not have a precise idea concerning this parameter for the main outcomes of this study. However, McKenzie (2012) provides a useful indication for a number of similar economic outcomes, though in different countries and contexts. Education outcomes such as math and language test scores tend to have autocorrelation coefficients above 0.5 and 0.6 even when the time interval of the measurement is one or two years. Instead, income and expenditure measured at six-months intervals have lower coefficients and vary between 0.1 and 0.4. Food security is a multidimensional concept entailing not only food access, but also availability, utilization and stability. Probably the food insecurity indicators in the Somali context are likely to be closer to the lower bound of these estimates, while the diversification indicators to the upper bound.

70. Figure 8 displays the outcomes of the minimum detectable effect (MDE) analysis, where the calculated values are based on a range of ICC coefficients spanning from 0.05 to 0.30. These ICC values were selected after calculating them on the outcome indexes constructed with the baseline survey data.[12] With the exception of the resilience index, for which we found an extreme ICC equal to 0.67, the other indicators were all comprised between 0.17 and 0.29. The figure visually represents how the MDE, which is expressed in effect size, changes with different ICC values. Obviously, as the ICC gets larger, the MDE increases too. With the given evaluation design, the MDE ranges between 0.28 and 0.33.

Figure 8. Relationship between minimum detectable effect and intracluster correlation coefficient

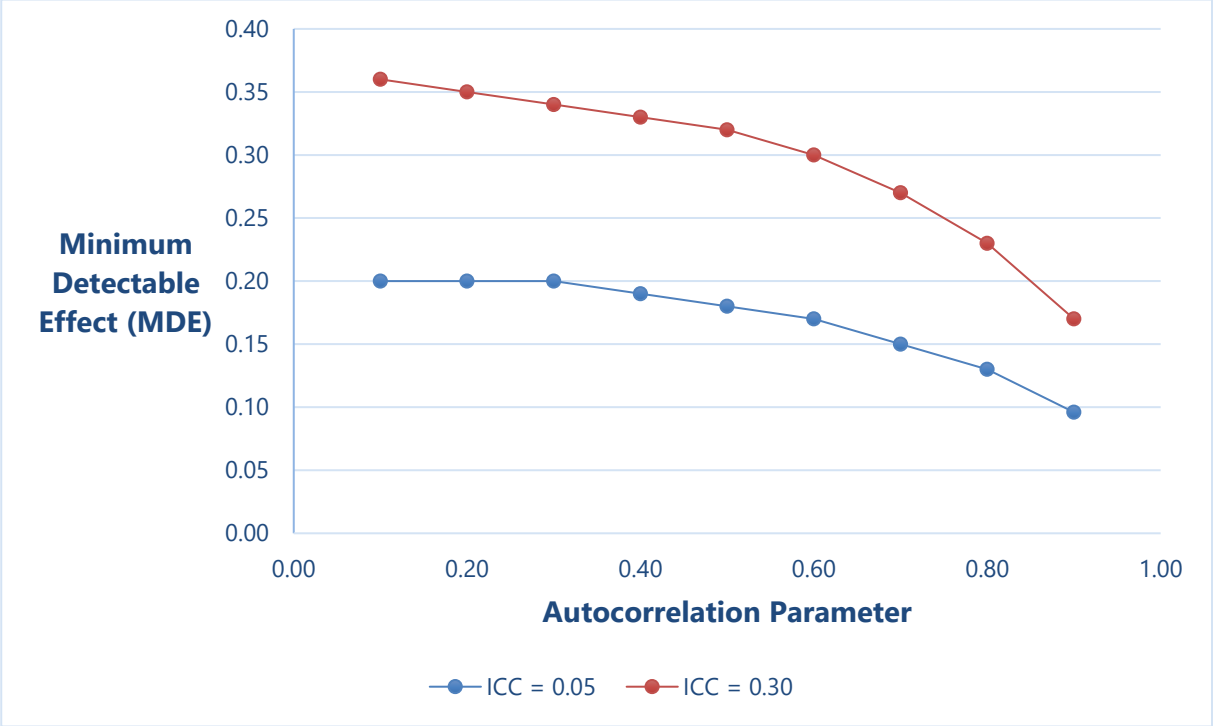


Note: Calculations done using the baseline data collected for the project in January/February 2023. Significance level α and power of the test β equal 0.05 and 0.8. We use 15 observations for each cluster, 56 clusters per treatment arm, no staggered treatment timing, 4 time periods, a longitudinal design with 1 pre- and 3 post-treatment periods, and AR(1) parameter error equal to 0.4. Power calculations carried out with R-shiny dashboard
Source: Schochet (2022).

[12] Intracluster correlation coefficients were calculated with Stata loneway command.

71. In Figure 9, we introduce variations in the autocorrelation parameter and calculate the corresponding MDE values. While keeping the ICC values constant at either 0.05 or 0.30, we modify the autocorrelation parameters, ranging from 0.1 to 0.9. Our calculations also show that an increase above 0.8 of the parameter governing the AR(1) process for the autocorrelated errors will lower considerably the MDEs, especially for those indicators with high levels of ICC.

Figure 9. Relationship between minimum detectable effect and autocorrelation parameter



Note: Calculations done using the baseline data collected for the project in January/February 2023. Significance level α and power of the test β equal 0.05 and 0.8. We use 15 observations for each cluster, 56 clusters per treatment arm, no staggered treatment timing, 4 time periods, a longitudinal design with 1 pre- and 3 post-treatment periods. Power calculations carried out with R-shiny dashboard
Source: Schochet (2022).

72. At the baseline we have interviewed 1 697 households, with 857 comparison households in 57 clusters and 840 households in 56 clusters. At the follow-up it will be unlikely to reach all the respondents initially sampled. The risk of substantial attrition is high, especially considering the possibility of displacement due to conflict or other shocks in the study area. Additionally, even more important than the level of attrition is that attrition in this setting is likely to be selective, threatening the internal validity of the estimates, for example if it is differential across treatment arms.

73. We will check whether non-response is correlated with the random assignment. If there is a statistically significant difference in non-response between the cash-for-work group and the comparison group, we will follow the procedure proposed by Kling *et al.* (2007). We will obtain lower bounds of the treatment effect by replacing missing observations in the treatment (control) arms by the corresponding arm's mean value minus (plus) 0.05, 0.10 and 0.20 standard deviations of the control group. Upper bounds of the treatment effects will be constructed in a symmetrical way.

74. Finally, in this study we will also adjust the p-values for the fact that we are testing the impact on several outcomes. We will calculate q-values using the Benjamini-Hochberg step-up method, which minimizes the false discovery rate (Benjamini and Hochberg, 1995; Benjamini and Yekutieli,

2001). The false discovery rate method entails that the M p-values of the i hypotheses are ordered from low to high and that the critical value of the p-value is then p(i) = α*i/M. Therefore, with six outcomes and hypotheses and a significance level (α) of 0.05, the critical p-value would be 0.0083 for the one with the lowest p-value (0.05* 1/6), which coincides with the most restrictive Bonferroni correction. For the second hypothesis, the critical p-value is 0.01666667 (0.05*2/6) and for the seventh it is 0.05 (0.05*6/6). Thus, correcting for the false discovery rate increases the MDEs, which will be comprised between 0.304 and 0.472 standard deviations in the experimental evaluation.

75. We will also look at each variable composing the outcome measures, to open the "black box" represented by the summary indices. The main advantage of using the Benjamini-Hochberg procedure is that it has more power to detect real differences with the same uncorrected p-value, especially if the number of measured parameters is large. Further, it is less conservative as it allows for correlation across test statistics, while other methods such as Bonferroni are based on the assumption of independence. This is unlikely to be the case, especially within the summary indices "family".

7. Conclusions

76. In this report we described the impact evaluation design of the cash-for-work component of the 2023 Humanitarian Response Plan in Somalia carried out by FAO. To evaluate the impact of the programme, we use a longitudinal, quasi-experimental design. The evaluation focuses on 14 districts in the regions of Somalia, Somaliland and Puntland. Within each district, four treatment villages and four comparison villages were selected. Given the impossibility of randomizing villages, the comparison group was selected with propensity score matching. To carry-out the matching exercise, we use a list of villages located in rural areas of targeted districts and where infrastructure such as feeder roads, contour bunds, or water catchments existed within a 6 km radius of the village and were not among those already rehabilitated by the cash-for-work programme in the 2023 project round.

77. The descriptive analysis of the baseline survey data shows surprisingly small imbalances across the treatment and comparison groups. We calculate five main programme outcomes as summary indexes following Anderson (2008): resilience, food security, income and livelihoods, agency, social cohesion. For all of them, we fail to observe statistically significant differences between the two evaluation arms. We look also at each indicator composing the indexes, and only for few indicators we observe a divergence in the groups' average. Participants in the cash-for-work treatment arm report a larger number of income sources, a larger number of months with members participating in wage labour, and a greater participation in social networks than the comparison group. These results were expected, since the baseline survey took place around three weeks after the public works activities started, though payments were not made.

78. Overall, two thirds of the sample households report adequate food consumption and two sources of labour and non-labour income. Further, agricultural activities mainly concern livestock, since 80 percent of the households hold at least one type of animal, especially small ruminants, while only 14 percent of the sample is engaged in crop production.

79. The findings from the descriptive analysis of sociodemographic characteristics point to a relatively successful targeting approach, which focused on both men and women with labour capacity. In fact, we observe a relatively large share of people of working age and a minor share of elderly people. Households lacking adult members are only 1 percent of the sample. Further, the number of adult household members with a moderate or severe disability is extremely low (0.173) and equivalent to one out of ten households reporting to have people with significant health issues.

80. The econometric analysis carried out on the baseline survey data provide some insights concerning the relationship between our subjective measure of resilience and its most immediate determinants. Three outcome indicators (food security, income and social cohesion) have the expected positive, albeit weak, correlation with the summary index of resilience. The agency index instead has a strong and negative correlation with resilience. After unpacking the agency index into the three main subcomponents (chance, powerful others and locus of control), we notice that the locus of control subindex is positively associated with resilience. In the context of fragility and protracted crises, this means that greater resilience is associated and determined not only by people's own actions, but also by other factors such as luck and the role played by leaders. Finally, we look at the relationship between resilience, climate and weather shocks. Although the long-run mean of rainfall precipitations in both the deyr and gu seasons are positively associated with household's resilience, weather shocks do not have any major association with resilience if occurred more than three months prior to the survey.

81. We conclude this report by providing a couple of reflections concerning the impact evaluation assessment:

i. Even though cash-for-work beneficiaries did not receive any payment, public works activities connected to the rehabilitation and maintenance of existing infrastructure had started three weeks prior to the baseline survey. This represents a limitation for the impact evaluation, since it may entail a behavioural response to the programme occurring before the baseline has occurred.

ii. The substantial equivalence between the two treatment arms, especially in terms of their demographic profile, is positive for the impact evaluation, but obviously it does not guarantee the parallel trend assumption is maintained. We do not have any pre-programme data that we can use to verify this hypothesis. So, the best approach we could take at the follow-up is to carry out a series of placebo tests, with either a fake treatment group or a fake outcome. Identifying a fake treatment group in the context of the cash-for-work evaluation is not easy, especially because we do not have access to additional sources of data in the treatment areas, such as a sample of ineligible households. However, one might believe that households that have low labour capacity should not be included in the programme, but if they are, the impact of the programme for them should be zero. By analysing the baseline survey data, we find that a share of households is either moderately (24 percent) or severely (8 percent) labour constrained. We define a household as severely labour constrained if there are no able-bodied members, while moderately labour constrained households have members fit-to-work, but dependency ratios greater than 3. We could use only the subsample of the households moderately and severely labour constrained in the treatment group as fake treatment group and look for a zero impact of the cash-for-work programme. Similarly, a placebo test can be performed also with a fake outcome. For instance, we may want to test the validity of the comparison group by estimating the impact of the cash-for-work on an outcome that we know is not affected by it, such as the number of children in primary or secondary school age, or the number of elderly or disabled members in the family. If our DiD estimation finds an impact on such variables, then it is likely that our comparison group is flawed.

iii. Quasi-experimental impact evaluations typically control for a range of characteristics that are unbalanced at the baseline or those that are good predictors of the outcomes. The choice of these indicators is, however, subjective. So, the question is: what controls should we include to: i) satisfy the conditional independence assumption and avoid the omitted variable bias?; ii) satisfy the overlap assumption? Conditional independence means that, dependent on a set of control variables, the potential outcome is independent of the treatment assignment. The more variables there are in the model, the more plausible it is that the conditional independence assumption is satisfied. On the other hand, the overlap assumption implies that there is always a positive probability that any given unit is treated or untreated (the propensity scores are far away from zero and one). The fewer variables there are in the model, the more comfortable we can be that the overlap assumption is satisfied. To summarize, the conflict is that the conditional independence assumption expects many variables in the model, but the overlap assumption expects few variables. We aim to resolve this conflict and to prevent model selection mistakes by adopting least absolute shrinkage and selection operator (LASSO) techniques. For instance, Farrell (2015) and Chernozhukov *et al.* (2018) proposed to combine LASSO methods with inverse probability weighting doubly robust estimators to estimate average treatment effects. We will conduct robustness checks for our main

estimates, by revising the LASSO method used to select the indicators for both the treatment and outcome equations.

References

Anderson, M.L. 2008. Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, *103*(484), 1481–1495. https://doi.org/10.1198/016214508000000841

Andrews, C. and Kryeziu, A. 2013. *Public Works and the Jobs Agenda: Pathways for Social Cohesion?* [Background paper for the World Development Report 2013]. The World Bank. https://www.kpsrl.org/sites/default/files/publications/files/wdr2013_bp_social_cohesion_and_public_works_job_creation.pdf

Angeles, G., de Hoop, J., Handa, S., Kilburn, K., Milazzo, A. and Peterman, A. 2019. Government of Malawi's unconditional cash transfer improves youth mental health. *Social Science & Medicine*, *225*, 108–119. https://doi.org/10.1016/j.socscimed.2019.01.037

Angrist, J. D. and Pischke, J.-S. 2009. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Attah, R., Barca, V., Kardan, A., MacAuslan, I., Merttens, F. and Pellerano, L. 2016. Can Social Protection Affect Psychosocial Wellbeing and Why Does This Matter? Lessons from Cash Transfers in Sub-Saharan Africa. *The Journal of Development Studies*, *52*(8), 1115–1131. https://doi.org/10.1080/00220388.2015.1134777

Babajanian, B. 2012. *Social protection and its contribution to social cohesion and state building*. Deutsche Gesellschaft für Internationale Zusammenarbeit. https://cdn.odi.org/media/documents/7759.pdf

Beierl, S. and Dodlova, M. 2022. Public Works Programmes and Cooperation for the Common Good: Evidence from Malawi. *The European Journal of Development Research*, *34*(3), 1264–1284. https://doi.org/10.1057/s41287-022-00525-1

Benjamini, Y. and Hochberg, Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Benjamini, Y. and Yekutieli, D. 2001. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, *29*(4). https://doi.org/10.1214/aos/1013699998

Cattaneo, M.D., Titiunik, R. and Vazquez-Bare, G. 2019. Power calculations for regression-discontinuity designs. *The Stata Journal: Promoting Communications on Statistics and Stata*, *19*(1), 210–245. https://doi.org/10.1177/1536867X19830919

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. B. Hansen, W. K. Newey, and J. M. Robins. 2018. Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal* 21: C1–C68. https://doi.org/10.1111/ectj.12097.

Conning, J. and Kevane, M. 2002. Community-Based Targeting Mechanisms for Social Safety Nets: A Critical Review. *World Development*, *30*(3), 375–394. https://doi.org/10.1016/S0305-750X(01)00119-X

Dell, M., Jones, B.F. and Olken, B.A. 2014. What Do We Learn from the Weather? The New Climate-Economy Literature. *Journal of Economic Literature*, *52*(3), 740–798. https://doi.org/10.1257/jel.52.3.740

Famine Early Warning Systems Network (FEWSNET) and Food Security and Nutrition Analysis Unit (FSNAU). 2022. *Famine (IPC Phase 5) would likely occur if food assistance plans do not materialize* (June 2022 to January 2023; Somalia Food Security Outlook). https://fews.net/sites/default/files/documents/reports/SOMALIA-FSO-202206-final.pdf

FAO. 2016. *RIMA-II. Resilience Index Measurement and Analysis—II*. Food and Agriculture Organization of the United Nations. https://www.fao.org/3/i5665e/i5665e.pdf

FAO. 2022a. The Importance of Ukraine and the Russian Federation for Global Agricultural Markets and the Risks Associated with the War in Ukraine. (Information Note). https://www.fao.org/3/cb9013en/cb9013en.pdf

FAO. 2022b. Providing Emergency Life-saving Food and Livelihood Support to Drought- affected Communities in Somalia. Food and Agriculture Organization of the United Nations.

Farrell, M. H. 2015. Robust inference on average treatment effects with possibly more covariates than observations. *Journal of Econometrics* 189: 1–23. https://doi.org/10.1016/j.jeconom.2015.06.017

Fisher, E., Attah, R., Barca, V., O'Brien, C., Brook, S., Holland, J., Kardan, A., Pavanello, S. and Pozarny, P. 2017. The Livelihood Impacts of Cash Transfers in Sub-Saharan Africa: Beneficiary Perspectives from Six Countries. *World Development*, *99*, 299–319. https://doi.org/10.1016/j.worlddev.2017.05.020

Haushofer, J. and Shapiro, J. 2016. The Short-term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya*. *The Quarterly Journal of Economics*, *131*(4), 1973–2042. https://doi.org/10.1093/qje/qjw025

Hu, Y. and Hoover, D. 2018. Simple Power and Sample Size Estimation for Non-Randomized Longitudinal Difference in Differences Studies. *Journal of Biometrics and Biostatistics*, *9*(5), 415.

Intergovernmental Panel on Climate Change (IPCC). 2012. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Group I and II of the Intergovernmental Panel on Climate Change (pp. 65–107). Cambridge University Press.

Janzen, S.A., Carter, M.R. and Ikegami, M. 2021. Can insurance alter poverty dynamics and reduce the cost of social protection in developing countries? *Journal of Risk and Insurance*, *88*(2), 293–324. https://doi.org/10.1111/jori.12322

Jones, L. and d'Errico, M. 2019. Whose resilience matters? Like-for-like comparison of objective and subjective evaluations of resilience. *World Development*, *124*, 104632. https://doi.org/10.1016/j.worlddev.2019.104632

Kelly, D.L., Kolstad, C.D. and Mitchell, G.T. 2005. Adjustment costs from environmental change. *Journal of Environmental Economics and Management*, *50*(3), 468–495. https://doi.org/10.1016/j.jeem.2005.02.003

Kennedy, G., Berardo, A., Papavero, C., Horjus, P., Ballard, T., Dop, M., Delbaere, J. and Brouwer, I.D. 2010. Proxy measures of household food consumption for food security assessment and surveillance: Comparison of the household dietary diversity and food consumption scores. *Public Health Nutrition*, *13*(12), 2010–2018. https://doi.org/10.1017/S136898001000145X

Kling, J.R., Liebman, J.B. and Katz, L.F. 2007. Experimental Analysis of Neighborhood Effects. *Econometrica*, *75*(1), 83–119. https://doi.org/10.1111/j.1468-0262.2007.00733.x

Kolstad, C. D. and Moore, F.C. 2020. Estimating the Economic Impacts of Climate Change Using Weather Observations. *Review of Environmental Economics and Policy*, *14*(1), 1–24. https://doi.org/10.1093/reep/rez024

Levenson, H. 1981. DIFFERENTIATING AMONG INTERNALITY, POWERFUL OTHERS, AND CHANCE. In *Research with the Locus of Control Construct* (pp. 15–63). Elsevier. https://doi.org/10.1016/B978-0-12-443201-7.50006-3

Maxwell, D. and Caldwell, R. 2008. *The Coping Strategies Index: Field Methods Manual* [Second edition]. https://www.fsnnetwork.org/sites/default/files/coping_strategies_tool.pdf

McDonald, J., H. 2008. Handbook of biological statistics. Sparky House Publishing Baltimore, Maryland.

McKenzie, D. 2012. Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, *99*(2), 210–221. https://doi.org/10.1016/j.jdeveco.2012.01.002

McKenzie, D. 2011. Power Calculations for Propensity Score Matching? *World Bank Blogs*. https://blogs.worldbank.org/impactevaluations/power-calculations-for-propensity-score-matching

Mendelsohn, R., Nordhaus, W.D. and Shaw, D. 1994. The Impact of Global Warming on Agriculture: A Ricardian Analysis. *The American Economic Review*, *84*(4), 753–771. JSTOR.

Moore, F.C. and Lobell, D.B. 2014. Adaptation potential of European agriculture in response to climate change. *Nature Climate Change*, *4*(7), 610–614. https://doi.org/10.1038/nclimate2228

Ohrnberger, J., Fichera, E., Sutton, M. and Anselmi, L. 2020. The worse the better? Quantile treatment effects of a conditional cash transfer programme on mental health. *Health Policy and Planning*, *35*(9), 1137–1149. https://doi.org/10.1093/heapol/czaa079

Pace, N., Sebastian, A., Daidone, S., Dela O Campos, A.P., Prifti, E. and Davis, B. 2022. Cash transfers' role in improving livelihood diversification strategies and well-being: Short- and medium-term evidence from Zimbabwe. *World Development*, *154*, 105874. https://doi.org/10.1016/j.worlddev.2022.105874

Pavanello, S., Watson, C., Onyango-Ouma, W. and Bukuluki, P. 2016. Effects of Cash Transfers on Community Interactions: Emerging Evidence. *The Journal of Development Studies*, *52*(8), 1147–1161. https://doi.org/10.1080/00220388.2015.1134774

Schochet, P.Z. 2022. Statistical Power for Estimating Treatment Effects Using Difference-in-Differences and Comparative Interrupted Time Series Estimators With Variation in Treatment Timing. *Journal of Educational and Behavioral Statistics*, *47*(4), 367–405. https://doi.org/10.3102/10769986211070625

Schwab, B., Janzen, S., Magnan, N.P. and Thompson, W.M. 2020. Constructing a summary index using the standardized inverse-covariance weighted average of indicators. *The Stata Journal: Promoting Communications on Statistics and Stata*, *20*(4), 952–964. https://doi.org/10.1177/1536867X20976325

Swindale, A. and Bilinsky, P. 2006. *Household Dietary Diversity Score (HDDS) for Measurement of Household Food Access: Indicator Guide*. https://www.fantaproject.org/sites/default/files/resources/HDDS_v2_Sep06_0.pdf

UCS, Santa Barbara. 2024. CHIRPS: Rainfall Estimates from Rain Gauge and Satellite Observations. Climate Hazard Center. https://www.chc.ucsb.edu/data/chirps

Valli, E., Peterman, A. and Hidrobo, M. 2019. Economic Transfers and Social Cohesion in a Refugee-Hosting Setting. *The Journal of Development Studies*, *55*(sup1), 128–146. https://doi.org/10.1080/00220388.2019.1687879

Wooldridge, J.M. 2010. Econometric analysis of cross section and panel data (2nd ed). MIT Press.

World Food Programme (WFP). 2008. Food consumption analysis. Calculation and use of the food consumption score in food security analysis. https://www.indikit.net/document/8-food-consumption-analysis

Zintl, T. and Loewe, M. 2022. More than the Sum of Its Parts: Donor-Sponsored Cash-for-Work Programmes and Social Cohesion in Jordanian Communities Hosting Syrian Refugees. *The European Journal of Development Research*, *34*(3), 1285–1307. https://doi.org/10.1057/s41287-022-00536-y

Appendix 1. Additional tables and figures

Appendix table 1. Topics in the survey questionnaires

Household survey	
Roster, time use and wage labour	
Land	
Crop production	
Fruits and vegetables	
Livestock holding	
Livestock by-products	
Agricultural assets	
Agricultural inputs expenses	
Non-farm enterprises	
Hired labour	
Non-timber forest products	
Transfers	
Decision-making	
Access to information	
Agency	
Resilience to shocks	
Food insecurity	
Social cohesion and trust	

Appendix table 2. Nine-capacities model questions

#	Resilience capacity	Question
Q1	Absorptive / coping capacity	If [SHOCK] were to occur in my area tomorrow, my household would be able to successfully cope with the threats posed by the [SHOCK]
Q2	Transformative capacity	If [SHOCK] were to occur in my area tomorrow, my household can change its primary income or source of livelihood if needed
Q3	Adaptive capacity	If the rate and intensity of [SHOCK] was to increase significantly in the next 5 years, my household would have the ability to successfully adapt to the changing threats posed by the [SHOCK]
Q4	Financial capital	If [SHOCK] were to occur in my area tomorrow, my household would have access to sufficient financial resources to ensure that we fully recover from the threats posed by the [SHOCK] within a year
Q5	Social capital	If [SHOCK] were to occur in my area tomorrow, my household would be able to draw on the support of family and friends to ensure that we fully recover from the threats posed by the [SHOCK]
Q6	Political capital	If [SHOCK] were to occur in my area tomorrow, my household can rely on support from the government and local authorities
Q7	Learning	My household has learned considerably from how we have dealt with past [SHOCK] events. This knowledge is crucial in successfully dealing with future [SHOCK] events
Q8	Anticipatory capacity	My household is fully prepared for any future [SHOCK] that may occur in my area
Q9	Early warning	If [SHOCK] were to occur in my area tomorrow, my household would have access to early-warning information to ensure that we are fully prepared for the threats posed by the [SHOCK]

Note: [SHOCK] refers to droughts, floods and pests/diseases.

Appendix table 3. Agency questions

#	Domain	Question
Q1	Chance	To a great extent my life is controlled by accidental happenings
Q2	Powerful others	I feel like what happens in my life is mostly determined by powerful people
Q3	Internal locus of control	When I make plans, I am almost certain to make them work
Q4	Chance	When I get what I want, it's usually because I'm lucky
Q5	Powerful others	In order to have my plans work, I make sure that they fit in with the desires of people who have power over me.
Q6	Internal locus of control	My life is determined by my own actions
Q7	Chance	It's not always wise for me to plan too far ahead because many things turn out to be a matter of good or bad fortune
Q8	Powerful others	People like me have very little chance of protecting personal interests when they conflict with those of strong pressure groups
Q9	Internal locus of control	I am usually able to protect my personal interests

Appendix table 4. Attitudes statements to construct social cohesion index

#	Statement
S1	In your village, everyone generally gets along well
S2	People in your village do not care about other people's needs.
S3	The young people in your village treat elderly with respect
S4	People in your village are not friendly

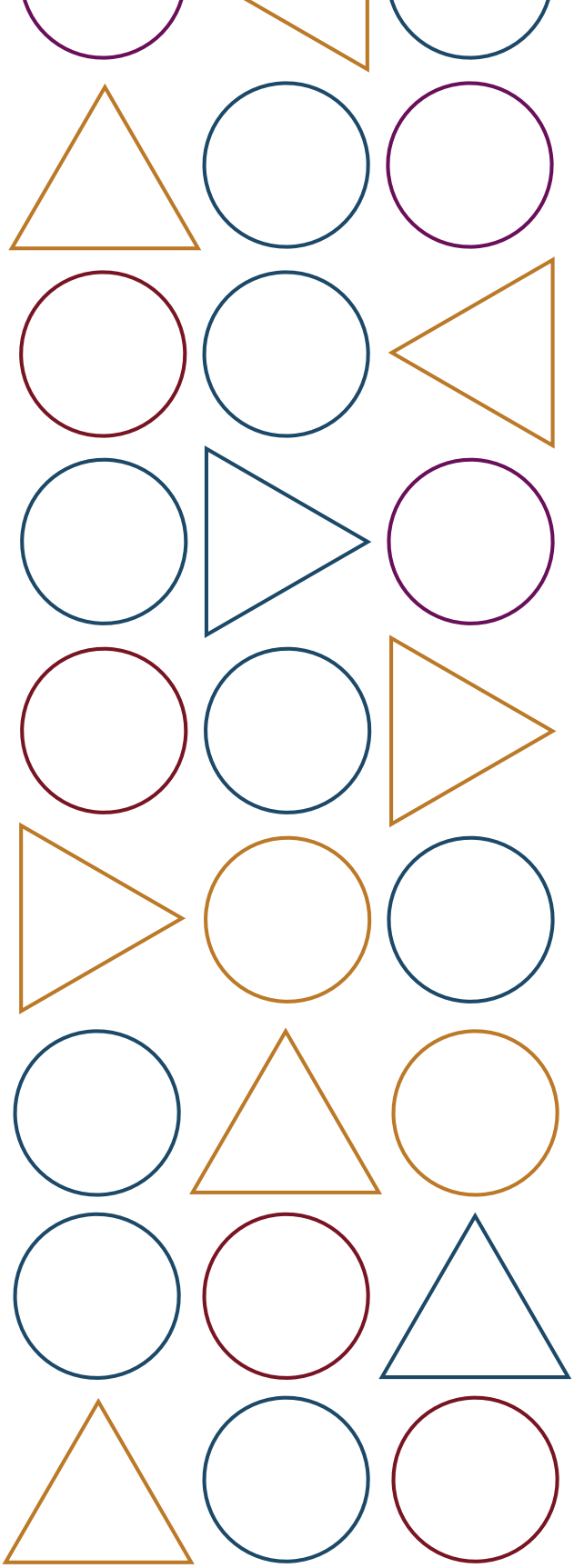
Notes: Strongly disagree=1, Disagree=2, Neutral=3, Agree=4, Strongly agree=5, not applicable/don't want to answer=99

Appendix table 5. Unadjusted and Holm- Sidak adjusted p-values for baseline comparisons

	All	Cash for work	Comparison	Difference	P-value	Sidak
# members in the HH	5.40	5.52	5.28	0.24	0.211	1.000
# HH members 0-17 yrs	3.06	3.20	2.92	0.28	0.048	0.913
# male HH members 18-59 yrs	1.15	1.12	1.19	-0.07	0.253	1.000
# female HH members 18-59 yrs	1.03	1.05	1.02	0.03	0.684	1.000
# HH members 60+ yrs	0.15	0.15	0.16	-0.01	0.842	1.000
Equivalence scale	3.36	3.40	3.32	0.08	0.487	1.000
# disabled HH members	0.17	0.17	0.18	-0.01	0.888	1.000
Households with only elderly (>59) and children (<18)	0.01	0.01	0.01	0.00	0.693	1.000
Share of dependents in HH	0.56	0.58	0.55	0.03	0.087	0.989
Female main respondent	0.37	0.39	0.36	0.03	0.436	1.000
Main respondent age	38.38	36.38	40.34	-3.96	0.000	0.000
Main respondent is married	0.89	0.89	0.90	-0.01	0.766	1.000
Main respondent yrs of education	1.15	1.11	1.19	-0.08	0.752	1.000
Summary index of food security	0.04	0.08	0.00	0.08	0.427	1.000
Food consumption score	47.72	48.16	47.28	0.88	0.797	1.000
Household dietary diversity score	5.30	5.34	5.25	0.10	0.744	1.000
Coping strategies index	21.85	21.55	22.14	-0.59	0.720	1.000
Summary index of income diversification	0.00	0.00	0.00	0.00	0.979	1.000
# income sources	2.08	2.31	1.86	0.45	0.001	0.033
# months food-retail business in operation	0.21	0.17	0.24	-0.07	0.516	1.000
# months non-food retail business in operation	0.07	0.09	0.05	0.04	0.283	1.000
# months manufacturing business in operation	0.02	0.03	0.00	0.03	0.159	1.000
# months services business in operation	0.13	0.13	0.13	0.00	0.986	1.000
# crops harvested after 2022 gu season	0.31	0.31	0.31	0.00	0.997	1.000

	All	Cash for work	Comparison	Difference	P-value	Sidak
# crops sold after 2022 gu season	0.12	0.13	0.12	0.01	0.833	1.000
# forms of transfers from public institutions received	0.94	1.05	0.83	0.22	0.126	0.998
# livestock income sources	0.52	0.60	0.45	0.16	0.118	0.998
# forms of private transfers received	0.59	0.62	0.55	0.07	0.518	1.000
# months last year adult members in wage labour	1.63	2.33	0.94	1.39	0.012	0.463
Summary index of agency	-0.01	-0.02	0.00	-0.02	0.844	1.000
Agency: chance subscale	53.92	54.85	53.01	1.83	0.366	1.000
Agency: powerful others subscale	61.50	61.65	61.34	0.32	0.873	1.000
Agency: internal locus of control subscale	59.41	58.40	60.39	-1.99	0.352	1.000
Agency score	58.31	58.35	58.27	0.08	0.936	1.000
Summary index of social cohesion	0.04	0.09	0.00	0.09	0.469	1.000
People's attitudes score	67.52	67.18	67.86	-0.68	0.589	1.000
People's help score	1.62	1.73	1.52	0.20	0.428	1.000
Trust score	56.32	55.42	57.20	-1.78	0.469	1.000
Group's participation score	0.57	0.68	0.46	0.23	0.029	0.775
Private transfers received	0.26	0.26	0.25	0.01	0.866	1.000
Private transfers made	0.34	0.34	0.34	0.00	0.947	1.000
Tropical Livestock Units	5.79	5.47	6.11	-0.64	0.504	1.000
Households owning livestock	0.80	0.80	0.80	0.01	0.865	1.000
# cattle	0.51	0.59	0.43	0.16	0.500	1.000
# sheep	10.16	8.90	11.39	-2.48	0.215	1.000
# goats	23.66	22.04	25.25	-3.21	0.529	1.000
# camels	1.87	1.77	1.97	-0.20	0.629	1.000
# donkeys	0.31	0.32	0.29	0.03	0.717	1.000
# chicken	0.53	0.59	0.47	0.12	0.513	1.000
No disability	0.89	0.90	0.89	0.00	0.863	1.000
Mild disability	0.03	0.03	0.03	0.00	0.757	1.000
Moderate disability	0.01	0.01	0.02	-0.01	0.223	1.000
Severe disability	0.06	0.06	0.06	0.00	0.864	1.000
Disability score	12.37	12.47	12.27	0.20	0.967	1.000

Note: Difference=mean difference between cash for work & comparison group. The total sample for all variables included is 1 697 observations.
Source: Authors' elaboration from survey data.



Office of Evaluation
evaluation@fao.org
www.fao.org/evaluation

Food and Agriculture Organization of the United Nations
Rome, Italy

CD2760EN/1/10.24