

# **Analysis plan of the impacts of the project “Gauging the impacts of Cash for Work emergency programming in Somalia.”**

Silvio Daidone, Fatimetou El Bah and Clemencia Cosentino

Food and Agriculture Organization of the United Nations (FAO)

July 20, 2023

## 1. Background

Somalia currently faces severe drought conditions caused by five consecutive poor rainy seasons. Most of the country experienced a hotter and drier than normal Jilaal<sup>1</sup> season prolonging the severe effects of the historic drought affecting pastoral communities over the past decade. The drought 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 depends heavily on food imports to mitigate the food deficit. However, the recent conflict in Ukraine has significantly worsened the situation, as Russia and Ukraine previously supplied over 90% of the country's wheat (FAO, 2022). The compounding effect of five consecutive seasons marked by below-average harvests, excessive livestock losses, and high global food prices has further aggravated the food security situation for the most vulnerable households in Somalia. The Famine Early Warning Systems Network and the Food Security and Nutritional Analysis Unit estimate that 6.5 million Somalis will face Crisis (IPC Phase 3) or worse acute food insecurity outcomes in 2023. Somalia currently receives large-scale humanitarian assistance to mitigate a worsening food security outcome in 2023.

FAO Somalia is contributing to the Humanitarian Response Plan (HRP) and the Drought Response and Famine Prevention Plan through the “Providing Emergency Life-saving Food and Livelihood Support to Drought-affected Communities in Somalia” project. This project seeks to improve the food security for the most drought-affected populations in rural Somalia (FAO 2022). The project is organized around five interventions:

1. Cash +
2. Unconditional Cash Transfers (UCT) and Cash-for-Work (CfW)
3. Transitional Cash and Livelihood Programme (TCLP)
4. Somalia Water and Land Information Management (SWALIM)
5. Desert Locust Control and Surveillance

Several donors fund this work. Although USAID's Bureau of Humanitarian Assistance (BHA) is FAO Somalia's primary Cash+ donor, other resource partners such as the Swedish International Development Cooperation Agency (Sida) and Germany's Federal Foreign Office provide resources for complementary agricultural livelihood inputs. FAO is evaluating the first two components—the short-term cash-based interventions. This report presents the design of the Cash-for-Work intervention (CfW).

Through its CfW intervention, FAO provides temporary employment opportunities in public projects, such as road construction, maintenance, and rehabilitation of public

---

<sup>1</sup> 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) followed the *jiilaal* dry season (December–March).

infrastructure. The primary objective of CfW programmes is to offer income support to vulnerable individuals, while facilitating the development of essential community assets. In 2022, FAO engaged 11,300 households across 21 districts in CfW 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. Furthermore, within the same CfW 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 (UCT) for the same duration as the CfW households, albeit for a smaller amount. This includes, for instance, households headed by ill family members, pregnant women, older persons and people living with disabilities.

## **2. Theory of change and hypotheses tested**

The Cash for Work (CfW) programme revolves around its dual objectives of 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.

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. Firstly, the CfW 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.

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 & Shapiro (2016) present compelling evidence of this causal link by investigating the impact of cash transfers on psycho-social 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. Comparable results are found in other countries, such as South Africa (Ohrnberger et al., 2020) and Malawi (Angeles et al., 2019) where national social cash transfer programs significantly improved mental health outcomes. Similarly, 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.

Lastly, the cash component of the CfW program 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. 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 program on social cohesion among Colombian refugees and poor Ecuadorians in urban areas of northern Ecuador. The study demonstrated that the program 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.

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

Firstly, through the incorporation of participatory elements in program 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 work programme is implemented participate in various aspects of programme implementation, including household identification, project selection, and project monitoring (Conning & 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 donor-

funded CfW programmes for Syrian refugees and their local neighbours in Jordan found that these interventions strengthened the sense of belonging and horizontal trust of participants and non-participants, refugees and locals, and in particular women (Zintl & 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 & Dodlova, 2022). Secondly, 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. Lastly, 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. Moreover, 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.

In studying the effect of the FAO Somalia CfW on households' resilience, and following the described theoretical framework, we formulate the following hypotheses (H):

H1: *CfW transfers will increase household resilience and reduce the adoption of negative coping-strategies.*

CfW 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 mean portions or frequency of food consumption. In addition, transfers will directly alleviate the credit and liquidity constraints faced by poor rural households.

H2: *CfW transfers will increase food security of beneficiary households.*

The most immediate impact of the CfW 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 fulfil their dietary needs. This, in turn, improves physical well-being, reduces days of work lost, and ultimately increase labour productivity.

H3: *CfW transfers will increase income diversification.*

CfW transfers induce diversification by allowing temporary labor market opportunities and enabling households to invest in income-generating activities that offer a higher return than subsistence agriculture.

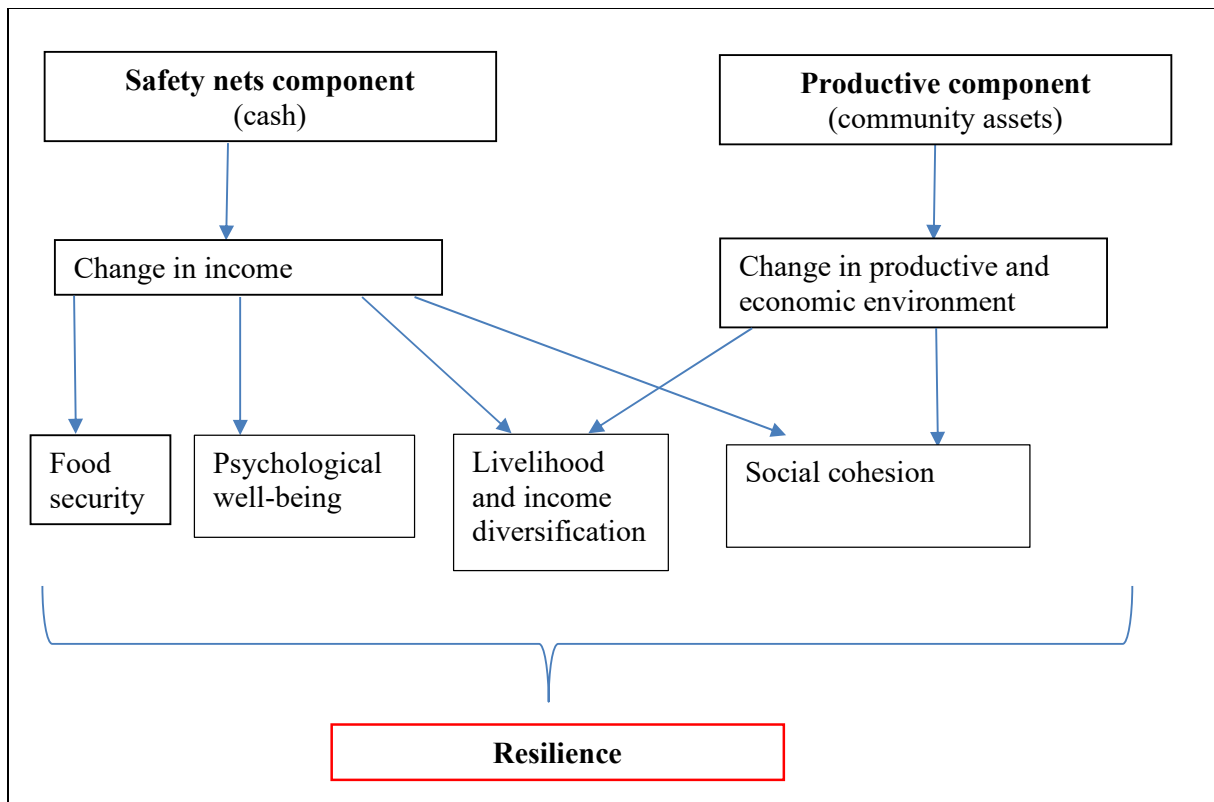
H4: *CfW will improve psychological well-being.*

CfW 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.

H5: *CfW will improve social cohesion.*

CfW 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 household’s resilience**



### 3. Study design

To evaluate the impact of the CfW programme, we will use a longitudinal, non-experimental design study. 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. The selection of villages for participation in public works was initially done based on specific eligibility criteria, with a key factor being the availability of able-bodied members within the household. The FAO Somalia office provided a list of villages involved in public works 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**

For its humanitarian response plan, FAO primarily targets rural communities and villages located more than 20 kilometers away from the main district town, specifically those that do not receive support from other humanitarian organizations. Given constraints on randomizing villages, the evaluation team used propensity score matching (with replacement) to create a matched comparison group of villages from the same districts where the CfW programme was implemented.

The model used to generate propensity scores to identify comparison villages included characteristics associated with the outcomes of interest, such as climatic conditions, land cover, distance from main district town, distance from water catchments/boreholes, markets, and paved road. The propensity score generated for each village in the sample, treated and not treated, helped identify those villages that could be included in the comparison group as they share similar characteristics with the treated villages in our sample. The final list of comparison villages included multiple potential replacements to account for cases where it was not feasible to reach a village or where the necessary conditions for inclusion in the sample were lacking.

Using list of comparison villages, along with their potential replacements, the data collection service provider contacted the chiefs of these villages to verify the presence of specific infrastructures, namely, feeder roads, contour bunds, or water catchments. If any of the mentioned infrastructures existed within a 6 km radius of the village and were not among those already rehabilitated by the CfW programme in the current project round, the village was confirmed to be eligible for the comparison group. If the required infrastructure was unavailable, the service provider proceeded to the next village on the list.

Village authorities or chiefs provided the approximate number of households in confirmed villages and a list of households with labor capacity as defined earlier. From this list, 15 households were selected for further interviews.

### **3.2. Data collection**

The CfW impact evaluation involves several rounds of data collection. These include baseline plus follow-ups at 6, 12, 24 and 36 months. The baseline data collection took place from January 26 to February 11, 2023. The fieldwork was conducted by a team consisting of one project coordinator and approximately 25 field enumerators.

The evaluation team originally aimed to gather a total of 1,680 observations, which were to be distributed as follows:

- 840 observations (household surveys administered) in 56 treatment villages (with 15 observations per village), covering 4 villages per district in 14 districts.
- 840 observations in 56 neighboring villages in the comparison group (also with 15 observations per village), also covering 4 villages per district in 14 districts.

The service provider surpassed this target by conducting a total of 1,697 interviews. This was achieved by including an additional comparison village in the district of Xudun.

Baseline data will enable us to:

- Produce and document descriptive characteristics—including demographics, initial living conditions and other circumstances—of both the CfW participants and comparison group prior to any cash transfers being disbursed.
- Measure changes over time against baseline and estimate impacts.
- Control for any baseline differences in the treatment and comparison groups at the start of the evaluation, making the evaluation results more reliable and credible.

### **3.3. Survey instruments**

Data for this evaluation will be collected through a household questionnaire, which will cover the following topics:

1. Household composition: data on household members, such as age, gender, and their relationship to the main respondent.
2. Land ownership and crop production: information about land ownership, tenure arrangements, and the household's engagement in crop cultivation.
3. Livestock possession: information about household's ownership and management of livestock.
4. Consumption: household's consumption patterns and expenditures on various goods and services.
5. Coping strategies: strategies employed by the household to cope with various challenges, including economic shocks or food insecurity.
6. Access to basic services: household's access to essential services, such as water supply and sanitation facilities.
7. Social cohesion: information on the dynamics of social relationships and support networks within the household and the community.
8. Resilience to shocks: household's resilience to weather-related shocks, such as droughts or floods.

The design of the household instrument was guided by three core principles:

1. The survey instrument included key questions that can be used in analysis—for example, to create indicators—and enable an assessment of the program against its stated objectives. These core indicators include food security, transfers, access to markets, and savings among other important factors.



2. Whenever possible, the survey includes questions used in similar surveys conducted in the country. This ensured that the questions were appropriate and relevant for local conditions, had been tested and successfully used in the past, and facilitated comparisons with national-level information (for benchmarking).
3. The survey instrument was designed to be of manageable length to prevent interviewer and respondent fatigue, thus promoting data quality.

## 4. Main outcomes

In this section we describe a list of indicators that corresponds to the goal and outcomes included in the project's log frame. This will allow us to assess the programme against its main stated objectives, namely: resilience, food security, income diversification, self-efficacy and social cohesion. Many of these outcomes are multidimensional concepts and many indicators can be used to measure them. For this reason, we decided to consider for the main outcome variables a summary index approach, which facilitates generalizations of findings about the program's effectiveness.

We calculate summary indexes by adopting the standardized weighted mean approach (Anderson, 2008), using the comparison group as the default reference group for standardizing.<sup>2</sup> These standardized summary indexes à la Anderson do not have a specific meaning as they merely reflect deviations from the comparison group and can be thus interpreted as effect sizes. To summarize, we compute the following index variables:

### 4.1. Resilience index

The resilience index is a comprehensive measure of 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 9 capacities model developed by Jones & 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 9 sub-statements using a five-point Likert scale (ranging from 'strongly disagree' to 'strongly agree') to capture the household perception of existing resilience capacities (Table 2). 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 minmax normalization, transforming the results in a score that ranges from 0 (not at all resilient) to 100 (fully resilient):

---

<sup>2</sup> See Schwab et al. (2020) for a detailed step-by-step guide to construct such summary indexes à la Anderson.

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

An alternate resilience score will also be computed using the Resilience Index Measurement 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. RIMA employs latent variables models to estimate resilience. Latent variable models assume that a) observed variables are manifestations of an underlying unobserved latent concept and b) other variables (correlates) construct and influence the latent factor(s), with a reciprocal effect. A Multiple Indicator Multiple Causes (MIMIC) model explains the relationship between observable variables and the unobservable variable by minimizing the distance between the sample covariance matrix and the covariance matrix predicted by the model. The observable variables are divided into correlates of the latent variable (they can be both endogenous and exogenous) and its indicators. The correlates are part of the structure of the model, while the indicators are measured. The MIMIC model assumes that the variables are measured as deviations from their means and that the error term does not correlate with the pillars (correlates), which in the RIMA modelling are access to basic services (ABS), assets (AST), social safety nets (SSN) and adaptive capacity (AC). A min-max scaling is used to transform the resilience capacity index value retrieved from the MIMIC model into a standardized index, ranging between 0 and 100.

#### **4.2. Food Security index**

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

FCS 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.

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

The CSI 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, to provide a quick, current status indicator of the extent of food insecurity and to measure or monitor the impact of food assistance programs (Maxwell & Caldwell, 2008)

#### **4.3. Income diversification index**

The income diversification index is a standardized weighted average of the (positively coded) number of income sources; the number of months in: a) a food retail enterprise, b) a non-food retail business, c) a processing enterprises (for instance a restaurant or carpentry), and d) a business in the services sector; the number of crops harvested, the number of crops sold; the number of livestock income sources (livestock heads sold and livestock by-products sold); the number of different public transfer sources received; the number of private transfer forms; and the number of months of adult wage labor.

#### **4.4. Self-efficacy index**

As a proxy for psychological well-being we construct a summary index of self-efficacy, using a reduced number of questions adopted for the Internality, Powerful Others and Chance (IPC) scale proposed by Levenson (1981). IPC consists of three subscales: 1. 'Internality' measures the extent to which a person believes their own actions influence their life; 2. 'Powerful Others' measures the extent to which an individual feels others influence their life; and 3. 'Chance' measures how much the individual feels luck or chance is driving their life (Table 3). For each of these sub-scales we generate a score, following the minmax normalization used for the three resilience capacity scores discussed in section 4.1, whereas the respondents provide an answer to 9 statements using a four-point Likert scale (Strongly disagree = 1, Disagree=2, Agree=3, Strongly agree = 4). The final summary index of self-efficacy is a standardized weighted average of the (positively coded) internal locus of control score, and the (negatively coded) chance score and powerful others score.

#### **4.5. Social cohesion index**

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, neighbors, 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 refer to people behavior (Table 4).

- iii) Help score. It indicates the help provided by the neighbors 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: a) cooking meals; b) cleaning the dwelling/clothes; c) fetching water and/or firewood; d) buying food.
- iv) Group score. It encompasses the main respondent's and their family member's participation in the following groups/associations: a) farmers; b) women support; c) youth; d) business association; e) religious; f) 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.

The four scores (trust, attitudes, help, and group) are constructed following the minmax approach already discussed in sections 4.1 and 4.4.

## 5. Estimation strategy

The CfW impact evaluation consists of a 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 CfW takes place and where FAO could rehabilitate at least one infrastructure within a six-kilometres radius, including water catchments, feeder roads, irrigation canals and contour bunds.

We will estimate the CfW 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:

$$\Delta Y_{ij} = \alpha + \beta CfW_j + \gamma X_{ij0} + \delta + \varepsilon_{ij}$$

Where  $\Delta Y_{ij}$  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 & Pischke, 2009).  $\delta$  is a vector of district dummies, while  $\beta$  represents the intention-to-treat (ITT) estimator of the programme impact. The outcomes of interest are resilience, food security, income diversification, self-efficacy, and social cohesion indexes.

For estimation, we will cluster standard errors at the village level, which is the unit of programme implementation, to account for the intra-cluster correlation and provide consistent estimates of standard errors. To account for possible imbalances across evaluation arms, we will combine the DiD model with propensity score matching (PSM). This approach of inverse probability weighting with regression adjustment (IPWRA) 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. Power and attrition**

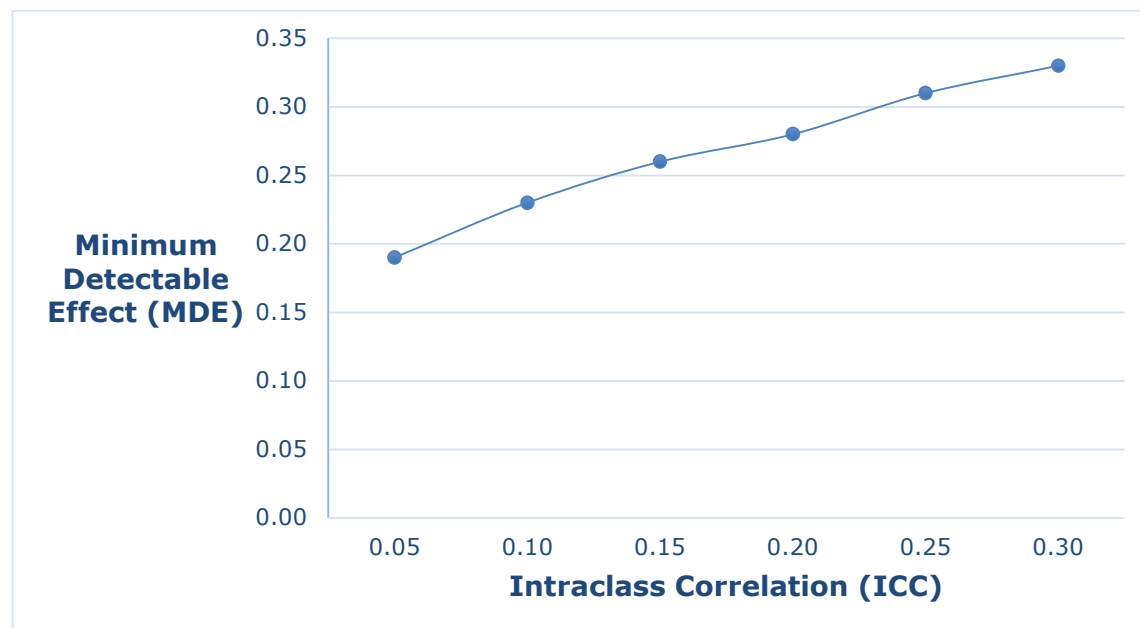
Although carrying-out power calculations for randomized control trials is relatively straightforward, doing so for quasi- and non-experimental designs requires more judgement. Recently, guidance became available for regression discontinuity designs (Cattaneo et al., 2019). For PSM and DiD designs, the 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% larger than in the pure experiment case. 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.

For DiD models, Hu & Hoover (2018) studied power / sample size estimation methods for non-randomized DiD designs. More recently, Schochet (2022) developed new closed-form variance expressions for power analyses for commonly used DID panel data estimators 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 cluster size, number of clusters and intra-cluster correlation (ICC) coefficient, other parameters must be considered in a DiD design. For the CfW 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. The literature lacks prior research that may suggest how this parameter might behave 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 1 or 2 years. Instead, income and expenditure measured at 6-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.

**Error! Reference source not found.** displays the outcomes of the minimum detectable effect (MDE) analysis, where the calculated values are based on a range of intra-class correlation coefficients (ICC) 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.<sup>3</sup> With the exception of the resilience index, for which we found an extreme ICC equal to 0.67, the other indicators were 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 2: Relationship between Minimum Detectable Effect and intra-cluster correlation coefficient**



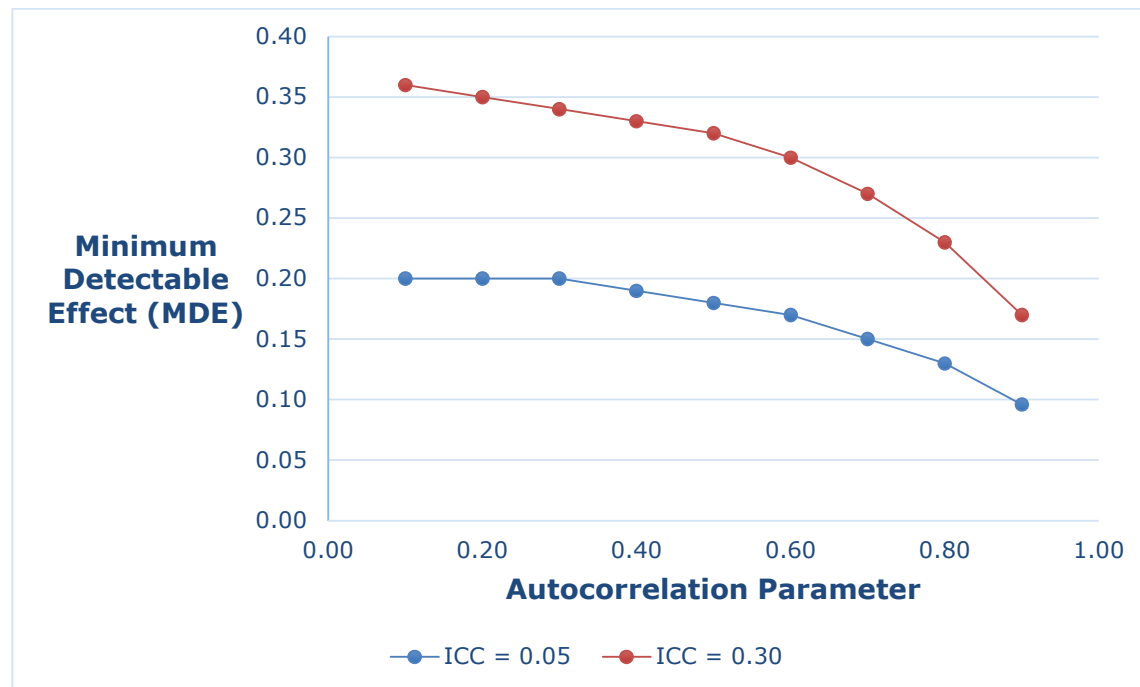
**Note:** Calculations done using the baseline data collected for the project in January/February 2023. Significance level  $\alpha$  and power of the test  $\beta$  equal 0.05 and 0.8. We use 15 observations for each cluster, 56 clusters per treatment arm, no staggered treatment timing, 4 time periods, 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 (Schochet, 2022).

In Figure 3, 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, spanning a range from 0.1 to 0.9. Our calculations show also that an increase above 0.8 of the parameter governing the AR(1) process for the auto correlated errors will lower considerably the MDEs, especially for those indicators with high levels of ICC.

<sup>3</sup> Intra-cluster correlation coefficients were calculated with Stata `loneway` command.



**Figure 3: 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  $\alpha$  and power of the test  $\beta$  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 (Schochet, 2022).

At baseline, we 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.

We will check whether non-response is correlated with the random assignment. If there is a statistically significant difference in non-response between the CfW 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.



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 & Hochberg, 1995; Benjamini & 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  $p(i) = \alpha * i / M$ . Therefore, with 6 outcomes and hypotheses and a significance level ( $\alpha$ ) 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.

We will also look at each variable composing the outcome measures, to open the “black box” represented by the summary indices. The main advantage for 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”.

## 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., & 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](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., & 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., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Attah, R., Barca, V., Kardan, A., MacAuslan, I., Merttens, F., & 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>
- Beierl, S., & 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., & 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., & 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., & 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>
- Conning, J., & 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](https://doi.org/10.1016/S0305-750X(01)00119-X)

- Famine Early Warning Systems Network (FEWSNET), & 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. (2022). *Providing Emergency Life-saving Food and Livelihood Support to Drought-affected Communities in Somalia*. Food and Agriculture Organization of the United Nations.
- Fisher, E., Attah, R., Barca, V., O'Brien, C., Brook, S., Holland, J., Kardan, A., Pavanello, S., & 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>
- Food and Agriculture Organization of the United Nations (FAO). (2022). *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>
- Haushofer, J., & Shapiro, J. (2016). The Short-term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya\*. *The Quarterly Journal of Economics*, 131(4), 1973–2042. <https://doi.org/10.1093/qje/qjw025>
- Hu, Y., & 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.
- Jones, L., & 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>
- Kennedy, G., Berardo, A., Papavero, C., Horjus, P., Ballard, T., Dop, M., Delbaere, J., & 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., & 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>

- 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., & Caldwell, R. (2008). *The Coping Strategies Index: Field Methods Manual* [Second edition]. [https://www.fsnnetwork.org/sites/default/files/coping\\_strategies\\_tool.pdf](https://www.fsnnetwork.org/sites/default/files/coping_strategies_tool.pdf)
- 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>
- Ohrnberger, J., Fichera, E., Sutton, M., & 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., & 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., & 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., & 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., & 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](https://www.fantaproject.org/sites/default/files/resources/HDDS_v2_Sep06_0.pdf)

- Valli, E., Peterman, A., & 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., & 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 – Additional tables and figures

**Table 1: Topics in the household survey**

---

Roster and wage labor
Land
Crop use
Livestock holding
Agricultural assets
Non-farm enterprises
Consumption
Coping strategies
Transfers
Access to basic services
Subjective resilience
Social Cohesion and trust
Self-efficacy

---

**Table 2: Nine-capacities model questions**

#	Resilience capacity	Question
Q1	Absorptive / coping capacity	If [SHOCK] was 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] was 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] was 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] was 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] was 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] was 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]

**Notes:** [SHOCK] refers to droughts, floods and pests/diseases. Strongly disagree = 1, Disagree=2, Neutral =3, Agree=4, Strongly agree = 5, not applicable/don't want to answer = 99





**Table 3: Self-efficacy 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 myself 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

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

**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