**PROPOSED PILOT EVALUATION DESIGN[1]**
**Community-based Conditional Cash Transfer Pilot**
**Tanzania Social Action Fund (TASAF)**

## 1. Description of CCT Pilot

The objectives of this pilot are to test how a conditional cash transfer (CCT) program could be implemented through a social fund using a community-driven development (CDD) approach. This CCT program, similar to other CCT programs, will provide grants to poor and vulnerable households contingent upon specific household actions: keeping children enrolled in and attending school and taking them to health centers on a regular basis. This CCT also stipulates that elderly persons visit health centers regularly, albeit less frequently than the young children. Recipients will receive between US$12 and $36 every other month, depending on the number of children and elderly in the household. These amounts were arrived at in consultation with Tanzania Social Action Fund (TASAF) representatives and are designed to make up 100% of the most recently available per capita food poverty line for adults (and 50% for children, presuming that children consume less than adults).[1]

This CCT pilot distinguishes itself in leveraging the central management capabilities of TASAF as well as the capacities of community organizations to deliver a CCT program. The Government of Tanzania (GoT) identified TASAF I-supported communities as the best places to pilot a community-based CCT program: these communities have previously received financial management training, they have successfully managed TASAF-funded sub-projects in their communities, and they have experience in monitoring and managing contractors. Community Management Committees (CMCs) in these communities will play several roles usually played by a centralized administration in CCTs: they will identify the recipient households, communicate the project conditionalities, monitor compliance with conditionalities, and manage the cash transfers.[2]

The GoT is interested in scaling up this program, and the training and experience received by TASAF I communities are being expanded from 40 districts and 2 islands in TASAF I to all of Tanzania's 123 districts and 2 islands in TASAF II. Thus, the lessons from this pilot will directly inform the potential scale-up. In addition, we will use data from the nationally representative Tanzania household budget survey (last administered in 2001) to identify how program households and villages compare with national averages to further inform scaling considerations.

This is both the first time that a social fund agency is being used to implement a CCT program in Africa and the first time that a CCT program is being delivered using a CDD approach. Specific objectives of the pilot are to (a) Develop operational modalities for

[1] Prepared by David Evans (AFTRL, World Bank) with input from Samantha de Silva and Julie Van Domelen (HDNSP, World Bank) and Myrtle Diachock (SDV, World Bank).

the community-driven delivery of a CCT program through a social fund operation; and (b) Test the effectiveness of the community-based CCT model and ensure that lessons from the pilot inform government policy on support to vulnerable families. Lessons from an impact evaluation of this pilot will have direct relevance for the portfolio of 13 social fund programs in the Africa region, totaling US$1 billion. This pilot will also inform any country which is considering a CCT program but where centralized administrative capacity is limited. It complements other efforts to test the effectiveness of cash transfer programs in Africa, such as the proposed evaluations of an unconditional cash transfer project in Zambia and a mixed conditional/unconditional project in Burkina Faso; but this project offers unique insights on using communities to identify target households and to monitor compliance with conditions. The quantitative impact evaluation is tied to a qualitative examination of how the program and the role of the community therein affect community dynamics and a process evaluation to illuminate the mechanisms by which the program has its impacts. As the first combined CDD-CCT effort, a high quality impact evaluation is needed for this pilot to inform future efforts not only in Tanzania, but also the future cash transfer programs globally.

## 2. Objective of the Evaluation and Outcome Indicators

The primary objective of this evaluation is to test the combined effectiveness of (a) a CCT program in Tanzania and (b) the CDD model of administering a CCT program. If either of these parts fails, then the CCT will be ineffective in improving outcomes for vulnerable households.[3] In order to assess how effective the program is in accomplishing improved conditions for vulnerable households, the impact evaluation will seek to answer the following research questions in this novel context with limited central administration capacity (as is the case in much of Sub-Saharan Africa):

- What is the impact of CCTs on health for vulnerable children and the elderly?
- What is their impact on education for vulnerable children?
- What is their impact on consumption for vulnerable children and the elderly?
- What is their impact on pre-existing informal systems of caring for vulnerable households?
- What is the impact of a community managed version of the CCT program on community dynamics?

The research question underlying most of the questions above is *whether this community-based model of conditional cash transfers is as effective at achieving health, education, and consumption gains as the centrally administered models used elsewhere.* Significant impacts in answering the questions above will give implicit affirmation to that question.

*Outcome Indicators*

In order to measure changes in the areas covered by the impact questions, relevant indicators will be measured for education, health, consumption, and transfers. For education, each child's current school enrollment, a measure of frequency of attendance, their current standard and their standard the previous year (to measure grade

progression), and end-of-year test scores will be measured. If enough households have children ages 5 to 8, we will estimate the impact on age at school entry, which can be relatively late in Tanzania. In health, information on clinic visits, recent episodes of disease or illnesses and the steps taken to treat them, the ability to perform daily activities (for the elderly, it can be very indicative of their health to ask if they can, for example, use the bathroom by themselves, walk a certain distance by themselves, and so on) will be gathered. For children, we will also gather information on height and weight to check for malnutrition: short-term malnutrition leads to reduced weight (for a given height), called "wasting," and longer-term malnutrition leads to reduced height (for a given age), called "stunting." (Over the time period in this study, we expect clearer impacts on weight than on height.)

For consumption, we will measure of the number of meals consumed as well as how much and what kinds of foods individuals consume each day, as the transfers are likely to directly affect household food consumption. Cash transfers are also likely to affect whatever systems already exist to transfer assistance to the poor. We will gather data on cash and in-kind transfers received by beneficiary households in the month previous to the baseline and then later to examine how the magnitude and nature of those transfers may have shifted in response to this external program. We will also ask about vulnerable households' savings to see whether the transfers allow the households to build a buffer against adverse shocks.

Household indicators will be supplemented with a short module of community indicators. We will gather information on the education and employment of community leaders as well as their effectiveness in mobilizing the community (e.g., by measuring the number of community meetings and the number of projects carried out by the community). This information will both indicate how the program affects community leadership capacity and demonstrate how the pre-existing community capacity (measured at baseline) affects the impacts of the program.

The evaluation will contain a qualitative component in order to understand the impact of the program on community dynamics, especially given the innovative broader role for the community in this project. The qualitative assessment will examine communities' perception of the targeting systems and outcomes; relationships between program beneficiaries and non-beneficiaries; household decision making; and other areas that come to light as the program progresses. The qualitative assessment will complement the quantitative analysis: whereas the quantitative analysis gauges whether the programs works or not, the qualitative assessment will shed additional light on the social dynamics behind the changes in countable outcomes. Methods to be used include semi-structured interviews with key informants as well as established participatory rural appraisal methods. (The latter will include a range of visualization methods for group-based analysis of social and environmental issues.)

We will also gather output and process indicators to help understand the causal chain between the program design and measured impacts: the program may have positive impacts, but understanding the mechanisms by which those impacts come about depends on a number of intermediate indicators. These are discussed in detail in Section 5.

## 3. Poverty and Vulnerability Targeting

The pilot aims to focus on the poorest and most vulnerable districts, villages, and households of Tanzania. This is achieved by confining the pilot to districts and villages that received a TASAF I intervention. Ranking of regions was undertaken during TASAF I to determine the number of districts that would be assigned to participate within each region. Regions were selected using various indicators of poverty (e.g., poverty level, food insecurity, primary school gross enrollment ratio). Within the regions, districts were prioritized using an index of relative poverty and deprivation constructed using data from the 1992 Income and Expenditure Survey. TASAF I this way targeted vulnerable districts, and this program builds on that by focusing on a collection of villages identified by TASAF I administrators as being among the poorest in the program.

At the household level, eligibility criteria for beneficiary households are based on household characteristics of the very poor that were defined by communities themselves through focus group discussions. The criteria are that the household be (a) very poor, (b) not receiving similar benefits in kind or cash from another program, and (c) home to an elderly person (60+) or an orphan or vulnerable child (OVC). "Very poor" was defined by stakeholders as a household meeting at least three of the following characteristics: (1) lack of a basic dwelling or shamba; (2) difficulty providing two meals per day; (3) no adult member has worked in the last month; (4) children with clothes/shoes in poor condition; (5) family does not own livestock; and (6) family does not own land.

## 4. Impact Evaluation Design

The project uses randomized assignment of the program at the village level, ultimately relying on comparing the changes in outcomes of beneficiary households in 40 randomly selected treatment villages over time to those of households that *would be* beneficiaries in 40 control villages. Impact evaluation communities will be drawn from Bagamoyo, Chamwino, and Kibaha districts. There are 80 villages within those three districts that have managed at least one TASAF-supported project and – in that context - have training and experience in financial management, monitoring, and implementation of small-scale infrastructure. We have met with representatives of the GoT and they have approved this plan; they will not be rolling out programs in the next two years that will specifically target the control villages. (In the case of some unforeseen contamination, an alternative strategy is to take advantage of the community ranking of vulnerable households and use a regression discontinuity approach to evaluating the impact of the cash transfers.)

In each community participating in the impact evaluation, we will hold meetings to ensure understanding among participant households and community leaders about the

purposes of the impact evaluation (i.e., to increase knowledge and inform future planning) and the reasons that the treatment cannot be universal (the program does not have resources to benefit all villages). Discussions with TASAF leaders suggest that even villages not benefiting from the cash transfers have received significant benefits from TASAF in the past and so are likely to participate in the necessary data collection exercises.

The selection of treatment and control households will follow the following process:

*Phase 1: Selection of program villages.* In this phase, the team will compile village-level information on the size of villages, the existence of the infrastructure necessary to accommodate the increase in demand (for example, in school enrollment or health clinic usage) that a CCT will induce, and the experience and quality of CMCs, as will be necessary to stratify the sample and to ensure that the villages are suited to the requirements of the CCT and to the enforcement of the program conditions. The sample of villages must have enough geographic heterogeneity to ensure that treatment and comparison villages will not be geographically adjacent after random assignment. This is important to avoid extensive migration or other confounding of the treatment and control villages. If some of the 80 villages do not satisfy these conditions, we will draw on available villages from the edge of adjacent districts to fill the sample without expanding operational costs.

*Phase 2: Identifying eligible households.* In this phase, the potential beneficiaries in all program villages (not yet divided into control and treatment communities) will be identified. CMCs and village councils will prepare ranked lists of households based on the criteria for vulnerable households which have been determined in discussions with TASAF communities *before* the villages are assigned as treatment and comparison villages. These lists will inform the selection of recipient households in treatment villages and of households for data collection in control villages. Extensive discussions with TASAF officials suggest that clear communication from the start that not all villages can participate will help to manage expectations to minimize tension that could arise if hypothetical beneficiaries never become actual beneficiaries.

*Phase 3: Selecting the treatment and control villages*. Once eligible households have been identified in all 80 program villages, 40 villages will be selected at random. Random selection will be stratified on known village characteristics (such as sub-district and village size) to ensure comparability between treatment and comparison villages.

*Phase 4: Selecting the treatment and control households.* The design team will use the total share of the eligible population across all selected communities to ensure proper coverage among all treatment communities. CMCs will receive a cap of how many households in the community can participate in the program based on a combination of village population and poverty map projections.

*Phase 5: Data collection.* Once all communities are assigned into the treatment or comparison groups, sampling for data collection can begin. Through power calculations,

we have identified the need to interview an average of 25 households per village.[4] In cases where participating households (i.e., households that *would* receive treatment, whether in a treatment or control community) do not exceed that number, the team will interview the full sample of target households. Alternatively, in communities with more participating households, the team will collect data on a random sample of households to achieve the program-wide average of 20 households per village. We will supplement this household survey with administrative data on child school attendance gathered from local schools and on health care from local clinics.

We plan to gather data on target households in both treatment and control villages at baseline and then again after one year and after two years. Administrative data will be gathered at intervals in the course of the two years. Qualitative data will be collected in parallel data rounds, from a sample of the target households as well as key community members (e.g., members of the CMCs) in the same villages as the quantitative data sample. Finally, each round of the household survey will be supplemented with the community module, administered to a community leader.

## 5. Process Evaluation

The previous sections have principally dealt with the impact evaluation, examining the impacts of the program on various outputs and outcomes: quantitative and qualitative measures of child, household, and community well-being. This project will also examine whether the CCT program can be effectively administered by CDD mechanisms through various elements of process evaluation.

*A.* *Quality of targeting*. In a selection of communities, we plan to carry out a household survey to examine whether the households that have been targeted for the CCT program (a) fulfill the conditions established for participations, and (b) fulfill more of these conditions than other households in the community. (If many households are eligible, the goal of the program is to target the more vulnerable households.) We will also use data from the national consumption survey and our own qualitative data to gauge whether the program indeed seems to be targeting the most vulnerable households.

*B.* *Fund management*. Audits will be performed at the end of the program to a selection of communities, and two possibilities are to (1) announce to all communities that there is a certain chance of being audited at the end of the program, or (2) give some communities a guarantee of being audited and give other communities only the possibility and then test whether that leads to differences in fund management.

We will also, in the follow-up household survey, inquire about the amounts that households received to learn whether they actually received the amount intended.

C. *Service availability*. The household surveys evaluating individual outcomes will also inquire about the availability of necessary services to fulfill the CCT conditions (e.g., were there enough places in school for students to attend?).

D. *Cost evaluation*. The process evaluation will include a detailed accounting of the costs of the program to allow analysis of whether it can be replicated elsewhere. Effectiveness is only half the story; costs are the other half.

E. *Documentation of process*. An important element of the process evaluation will be to document just how the program rolls out in actuality, identifying necessary changes to the planned roll-out so that scale-up and replication can learn from the pilot.

## 6. Timeline for Impact Evaluation

| **Activity** | **Timeline** |
|---|---|
| Begin process for contracting a survey firm and someone to manage the impact evaluation on the ground | October 2007 |
| Develop questionnaires for baseline survey | November – December 2007 |
| Hold sensitization meetings | November 2007 – January 2008 |
| Randomly select villages for treatment and comparison | January – February 2008 |
| Translate and pre-test questionnaires, train enumerators for baseline survey | January – February 2008 |
| Administer baseline questionnaire | March – April 2008 |
| Data entry | May – August 2008 |
| Repeat process for follow-up questionnaire #1 | March – April 2009 |
| Data entry | May – August 2009 |
| Repeat process for follow-up questionnaire #2 | March – April 2010 |
| Data entry | May – August 2010 |
| Dissemination workshop | Before December 2010 |

[1] Project dates for TASAF I are from Nov 2000 – June 2005; for TASAF II, they are from Nov 2005 – June 2010.

[2] Conditionality is "soft": households receive two warnings for non-compliance over eight months before a proportion of their benefits are withheld. At least one *un*conditional cash transfer program has showed significant impacts on physical and cognitive development: see Paxson and Schady, "Does Money Matter? The Effects of Cash Transfers on Child Health and Development in Rural Ecuador," World Bank Policy Research Working Paper 4226, May 2007.

[3] Ideally, we would test the effectiveness of each of these separately: however, due to the lack of a strong central administration to manage a CCT program in Tanzania, all participating communities will use the CDD model.

[4] With a total of 80 participating villages (40 treatment and 40 control) and an effect size of 0.20, we expect to need to interview 20 households per village in order to achieve 80% power. We recommend

interviewing 25 households per village since not every household will have vulnerable children: some few households may only have vulnerable elderly. This assumes 95% confidence levels for statistical significance and an intra-cluster correlation of 0.05. Evaluations of conditional cash transfer programs elsewhere have found effects of this size. For effects of this magnitude on health and education outcomes in a Nicaraguan CCT, see Rawlings & Rubio, "Evaluating the Impact of Conditional Cash Transfer Programs," *World Bank Research Observer*, 20(1):29-55, 2005, Table 6. For Mexico's PROGRESA program, see effect sizes on child height in Behrman and Hoddinott, "An Evaluation of the Impact of PROGRESA on Preschool Child Height," Food Consumption and Nutrition Division Discussion Paper, IFPRI, March 2001. See effect sizes on longer-term schooling outcomes in Behrman, Sengupta, and Todd, "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment in Rural Mexico," *Economic Development and Cultural Change*, 54: 237-275, 2005.