

# TUUNGANE

## Impact Evaluation Framework<sup>1</sup>

Macartan Humphreys  
Department of Political Science  
Columbia University  
mh2245@columbia.edu

Jodi Nelson  
Research and Evaluation  
International Rescue Committee  
jodi.nelson@theirc.org

July 2007

### ABSTRACT

We describe the research strategy to be used by IRC to evaluate the impact of a major post conflict development program in East Congo, TUUNGANE. The program is designed both to support economic recovery and to foster social cohesion and improve the quality of governance. The evaluation strategy is designed to measure whether these objectives are in fact met. To do so it uses a method of randomized intervention which allows for the creation of a set of control communities that are similar to the TUUNGANE communities in every respect except for the presence of the program. A large survey of a random sample of over 3000 respondents will be undertaken by IRC before the onset of programs and after program completion and allows for a comparison of differences in the changes over time between the two sets of communities. In addition to the main evaluation, the strategy described here also seeks to answer two key questions of program design. The first relates to women and the politics of affirmative action: does greater women's participation improve outcomes, do quotas on community councils to ensure greater participation work? The second addresses a central question in development economics: is the direct distribution of cash to be spent at the discretion of private citizens cheaper and more effective than more elaborate community managed public goods programs?

---

<sup>1</sup> This document draws in structure on a framework document co-produced by one of the authors for an evaluation of a related program in Indonesia (BRA-KDP / World Bank), also underway in 2007-2008. We are deeply indebted to Jana Frey, Charles Lor, Liz McBride, Eric Mvukiyeye, Brian Sage, and Adi Walker for sharing their insights and ideas regarding all aspects of the design of this evaluation strategy.

## Contents

1	INTRODUCTION.....	5
1.1	Background .....	5
1.2	Objectives of the TUUNGANE Program.....	5
1.3	Evaluation Strategy.....	6
2	EVALUATING TUUNGANE’S OVERALL IMPACT .....	10
2.1	Does Participation in TUUNGANE Improve Social Cohesion?.....	10
2.2	Does Participation in TUUNGANE improve good governance? .....	10
2.3	Does Participation in TUUNGANE foster socioeconomic recovery? .....	11
3	EVALUATING INDIVIDUAL COMPONENTS OF PROGRAM DESIGN .....	11
3.1	Does greater women’s participation in decision-making produce better outcomes for governance, social cohesion and welfare? .....	12
3.2	Is the direct delivery of cash more effective than village projects?.....	14
4	EVALUATING THE IMPACT OF CONTEXTUAL FACTORS .....	16
4.1	Hypotheses.....	16
5	UNDERSTANDING UNINTENDED CONSEQUENCES.....	18
5.1	Hypotheses.....	18
6	TOOLS AND DATA RESOURCES .....	19
6.1	Household and Chief Surveys [2007 - 2010] .....	19
6.2	Participation Data.....	19
6.3	Village Profile and Log .....	20
6.4	Project Records .....	20
6.5	Accounts.....	20
6.6	Events Data .....	20
7	EVALUATION TIMELINE AND OUTPUTS.....	20
7.1	Timeline.....	20
7.2	Output Description .....	21
8	ANNEX:.....	22
8.1	Randomized Allocation of Treatments.....	22
8.2	Power .....	22
8.3	Baseline Survey Instrument and Manual .....	25

## List of Hypotheses

- H 1 Individuals in communities that have taken part in TUUNGANE will exhibit higher levels of **acceptance of others** into their communities.
- H 2 Individuals in communities that have taken part in TUUNGANE will exhibit higher levels of **trust** in other members of their communities.
- H 3 TUUNGANE communities will be **more willing to contribute time and effort** individually to collective goods.
- H 4 TUUNGANE target **communities will be more likely to work together** to solve local development problems.
- H 5 Communities will be more **proactive in seeking support** from local government and NGOs for community initiatives and the private sector.
- H 6 Villages in TUUNGANE communities will have a greater propensity to **work collectively with other villages** to address development challenges.
- H 7 Individuals in TUUNGANE target communities will report a greater sense of a **right to take part** in local decisions.
- H 8 Individuals in TUUNGANE target communities will report a greater sense of **obligation to take part** in local decisions.
- H 9 Individuals in TUUNGANE target communities will report greater **knowledge** about local decision-making processes and outcomes [what will be tested will depend on TUUNGANE training].
- H 10 Individuals in TUUNGANE target communities will report an increased **willingness to hold traditional and political leaders accountable**.
- H 11 Individuals in TUUNGANE communities are more likely to believe that **local leaders should be elected** rather than selected through an alternative mechanism.
- H 12: Access to community utilities and infrastructure, including those not directly supported by TUUNGANE, will be greater in TUUNGANE target communities.
- H 13: Household Income and asset holdings will be greater in TUUNGANE target communities.
- H 14: Households will allocate a greater share of their time to productive activities in TUUNGANE target communities compared with control.
- H 15: Time devoted to productive activities not directly associated with TUUNGANE projects will increase.
- H 16: Communities with greater representation of women leaders produce higher gains in the three priority goals of cohesion, governance and welfare.
- H 17: Greater representation by women leaders produces higher levels and quality of participation in decision making at the village and community levels.
- H 18: Greater representation by women leaders on CDCs leads to a greater propensity to accept women as decision makers.
- H 19: Greater representation by women leaders leads to lower levels of financial irregularities in the use of project money.

- H 20: At the end of the project, individuals, and especially women, in areas with greater women's participation will express greater satisfaction with the choice of projects.
- H 21: [Cross Intervention Hypothesis] Public goods are more likely to be selected than cash in communities with greater women's participation (see 3.2).
- H 22: Communities that select cash transfers will exhibit subsequent extorted or higher levels of competition and conflict.
- H 23: Program gains in the areas of governance, cohesion, and welfare will not be as great in communities that select cash transfers.
- H 24: The use of cash transfers leads to private gains at least as great as the use of public goods.
- H 25: The gains from cash transfers are greatest in more severely conflict affected communities.
- H 26: Individuals in TUUNGANE villages that choose cash transfers will report a weaker sense of rights to take part in local decisions than villages that choose projects.
- H 27: Individuals in TUUNGANE villages that choose cash transfers will report weaker sense of obligation to take part in local decisions than villages that choose projects.
- H 28: [ENGAGEMENT] Individuals in TUUNGANE villages that choose cash transfers will report a weaker propensity to question the actions and decisions of local leaders, traditional chiefs and CDC leaders than villages that choose projects.
- H 29: Individuals in TUUNGANE villages that choose cash transfers will report less willingness to hold local leaders, traditional chiefs and CDC leaders accountable than villages that choose projects.
- H 30: [History of conflict] The impact of TUUNGANE does not depend on the level of conflict intensity.
- H 31: [Wealth] The impact of TUUNGANE does not depend on the level of village wealth.
- H 32: [Cohesion] Improvements in social cohesion and governance does not depend on prior levels of cohesion or decision making norms.
- H 33: [Population flows] The impact of TUUNGANE does not depend on the numbers of returnees and new migrants.
- H 34: [Consistency of Impact across TUUNGANE Regions] Controlling for contextual factors, program results are of equal quality across all four project areas.
- H 35: [Program Complements / Substitutes] Program impact will be weaker in communities where there has already been support at the village or community level from international organizations.
- H 36: **Migration:** The TUUNGANE program will not result in higher levels of in-migration to TUUNGANE target communities.
- H 37: **Jealousies:** The TUUNGANE program will not result in higher levels of discontent in control communities that are proximate to TUUNGANE communities.
- H 38: **Effects on other programs:** Participation in TUUNGANE will not result in inequities from the targeting of other programs disproportionately to areas in which TUUNGANE is already operating (other than through demand effects resulting from community initiatives).
- H 39: **Inflation:** The TUUNGANE program does not have adverse effects on local inflation.

## **1 Introduction**

### **1.1 Background**

The International Rescue Committee (IRC), with its partners CARE International and the International Foundation for Education and Self-Help (IFESH), is undertaking a major post conflict development initiative entitled TUUNGANE. Eastern Congo has been devastated by violent conflict for over a decade resulting in massive loss of life and declines in welfare. With poverty being both a result and a predictor of violent conflict there is a real fear that East Congo could be caught in a violence-poverty trap. The goal of the initiative is to contribute to a response to these concerns. By establishing and strengthening participatory local governance committees at village, community, and regional levels, TUUNGANE aims to build stability by fostering social cohesion, improving governance and strengthening citizens' relationships with local government. The conduit to achieve these ends is village- and community-level projects that contribute directly to socio-economic recovery.

The use of community driven reconstruction (CDR) projects in post conflict settings is a strategy that is increasingly used throughout the world. Major CDR projects have been implemented in Rwanda, Liberia, Sierra Leone, Indonesia and Afghanistan. They hold out the promise of creating major change from the grassroots up. Yet despite their growing popularity there is little rigorous research describing whether and how they are effective.<sup>2</sup> This document describes a simple strategy to respond to the question of program impact. For the purposes of the evaluation program impact is defined here as the difference the program makes for people in the areas of socioeconomic recovery, social cohesion and governance. It will be assessed using a rigorous evaluation design, involving data collection in randomly assigned TUUNGANE and non-TUUNGANE areas. Household and chief surveys together with qualitative and quantitative monitoring data will help to examine a series of hypotheses about the program's impact. Resources permitting, the final analysis will be strengthened by mid-term process evaluation to document implementation and additional research activities to investigate the effectiveness of two variations in program design in select areas.

In the sections that follow we describe the evaluation design, the primary and secondary hypotheses it seeks to examine and the related data that will be available to both implementing and funding partners at various times during the program's lifetime.

### **1.2 Objectives of the TUUNGANE Program**

TUUNGANE will work in approximately 280 communities comprising roughly 1,400 villages in four of the country's most war-affected areas: South Kivu Province, Maniema Province, and Nord Katanga and Haut Katanga Districts of Katanga Province. These communities represent an estimated

---

<sup>2</sup> For the case of Liberia a related evaluation of an IRC post conflict CDD program is also underway led by James Fearon and Jeremy Weinstein of Stanford University and Macartan Humphreys of Columbia University.

beneficiary population by project end of approximately 1,780,000 inhabitants. The program will deliver more than £12,300,000 worth of projects directly to the communities, or average block grants of over £37,000 per community and £1,250 per village, thus representing a contribution of £38 per household.<sup>3</sup> Should the program follow the trend of similar projects, shown to have an average economic rate of return of 63%,<sup>4</sup> TUUNGANE will put over £20 million into the local economy in the short run. If there is an average of two projects per village, TUUNGANE is similar in scale to 2,800 smaller initiatives targeting basic livelihoods for families through projects such as small scale rehabilitation and repair of communal infrastructure, village livestock or seed banks, and income generation for vulnerable groups.

TUUNGANE uses an approach called community-driven reconstruction (CDR). Guided by IRC's best practice manual, *IRC's Best Practice Approach to Community-Driven Reconstruction*,<sup>5</sup> TUUNGANE aims to mobilize communities to manage resources towards implementing their highest priority development projects in a way that builds both social cohesion and linkages with local government institutions. Through this process, TUUNGANE aims to contribute to increased stability in the region and instill practices that embody the principles of good governance.

The core objectives of TUUNGANE are to foster:

1. **Social cohesion:** The trust people have in each other and in the ability of local institutions to represent their interests.
2. **Good governance:** local governance institutions and adherence to principles that allow for transparent, inclusive decision-making, providing communities with the means to hold their leaders accountable, make collective decisions about development and resolve differences peacefully.
3. **Economic recovery:** Increased socio-economic activity, access to utilities and services and greater welfare levels.

### 1.3 Evaluation Strategy

Built into TUUNGANE is a robust evaluation and research program. The evaluation component is designed to answer three types of questions:

---

<sup>3</sup> Estimated at 5.5 people per household.

<sup>4</sup> A review of economic recovery rates within Afghanistan's National Solidarity CDR programmed is found in Morelli, Juan, "NSP Cost Effectiveness Study" (November 2005).

<sup>5</sup> IRC, "IRC's Best Practice Approach to Community-Driven Reconstruction" (2007). This living document stems from a study conducted by IRC's Post-Conflict Development Initiative (PCDI) analyzing its community driven reconstruction approaches in four countries. It aims to develop a coherent institutional best practice model for first generation CDR in protracted and post-conflict settings.

- First, to understand *whether* the program meets its objectives. This is being done by the first best method of randomized impact evaluation. Measures of outcomes on each of the major program objectives are collected before and after program implementation and in a set of comparable, randomly selected treatment and control areas.
- Second, to understand *what* works and *why* the TUUNGANE works (or does not). This is done by exploring the relative effectiveness of alternative CDR program designs, again using the first best method of randomized evaluation.
- Third, to understand *when* and *where* the TUUNGANE program is most effective. This is done by collecting data on contextual variables, notably prior levels of community cohesion, economic deprivation and the existence of other development projects in the region. By examining the extent to which program gains depend on these background variables, we can draw conclusions about the conditions under which the CDR design is most appropriate as a post conflict development strategy.

In all of these objectives the evaluation design pays particular attention to the situation of women, examining the extent to which women in particular benefit from the program, their changing role in decision making and the impact of increased women’s participation on social outcomes.

### **1.3.1 Random assignment of communities**

To examine whether the program contributes to the economic recovery, social cohesion and local governance of targeted communities in Eastern DRC the evaluation strategy relies on the random selection of project sites.<sup>6</sup> Random selection allows for the identification of a control group of villages that is identical to treatment communities in all respects other than the fact of participating in TUUNGANE. This approach has benefits from both ethical and evaluation perspectives.<sup>7</sup> From an evaluation perspective the random assignment means that differences in outcomes between areas that do and do not benefit from TUUNGANE can be unambiguously attributed to the program. Without random assignment this attribution is difficult since the criteria for selecting communities for a program may itself be related to outcomes leading one either to underestimate or to overestimate program effects. From an ethical perspective the process employed provides an unusually high degree of ex ante equity in the distribution of limited resources. In addition since the selection is done in a highly transparent manner this procedure is likely to prevent concerns that favoritism or discrimination is at play in the selection of sites.

---

<sup>6</sup> The TUUNGANE program works with villages and with “communities” that are formed by grouping multiple villages together; the communities have approximately 6000 individuals and often (though not always) correspond to *localités*. The grouping together of villages into communities is done before sampling and survey implementation to allow sampling to be stratified by community and to ensure that data can be collected on the same units in which TUUNGANE will be operating.

<sup>7</sup> See Duflo, Esther and Michael Kremer, “Use of randomization in the evaluation of development effectiveness,” (July 2003).

The process used to select treatment communities is one in which the NGO partners first identify a set of eligible *territoires* that contain a population that is twice as large in size as the ultimate beneficiary population.

The *territoires* are selected based on the following criteria:

*Positive:*

- Conflict-prone or war-torn due to ethnic tensions, political clashes, land and resource disputes, return issues, the presence of militias, and other triggers, including the number of returnees, IDPs, and former combatants;
- Significant poverty, malnourishment, physical destruction, dearth of infrastructure and social services, and general deprivation resulting from war, displacement, and overall neglect;
- Consortium members' experience in the area and the presence of current programs with compatible elements;

*Limiting:*

- Adequate security in order that staff may engage with populations;
- Underserved areas balanced against access considerations with potential impact on budget, timing, logistics, and other factors: access to sites will be particularly important;
- No existing or planned programs that duplicate CDR structures.

This pool of communities thus contains only communities that are in great need. The number of such communities that is identified is however deliberately structured to be larger than the number that can be directly assisted by the TUUNGANE program. The ultimate selection of communities from within this pool is done by a public lottery. This is done as follows.

Within each *chefferie*, representatives of all communities are invited to attend a meeting in which the set of TUUNGANE communities will be selected. Those that are selected fall into the category of "treated" communities, those, in equal number, that do not, fall into the category of "control" communities. In this way the selection process is both transparent—all stakeholders can see the process through which selection into the program is undertaken—and equitable—all communities in the wider area have an ex ante equal chance of participating in the program.

With this strategy for identifying treatment and control communities complete, data from a large baseline and follow up survey are used to generate measures on a set of outcomes that capture core components of cohesion, governance and welfare. Impact will ultimately be estimated by examining the "difference in differences"; that is, how the difference in before and after outcomes differs between treatment and control sites.

### 1.3.2 Evaluation hypotheses

For the major outcomes of interest, the core measures will come primarily from pre- and post-intervention household and chief surveys in treatment and control communities in 2007 and 2010. The measures are used to test a set of hypotheses about program effects derived from the logic of CDR program design. There are two distinct types of hypotheses that guide the data collection process.

“Primary hypotheses” relate directly to what the TUUNGANE program hopes to achieve. Tests of these hypotheses provide a basis for assessing the first question: *whether* the CDR program meets its objectives. The evaluation will answer this question for the work of all TUUNGANE’s implementing partners.

In addition to these primary hypotheses we identify a set of “secondary hypotheses.” These hypotheses serve a number of different roles. They do not provide the basis for evaluating the success of the TUUNGANE program or CDR more generally. Rather, they provide a structure to help learn broader lessons about what works, when, where and why. Some provide a basis for evaluating specific CDR strategies or for determining the conditions under which CDR approaches are most effective (for example: when there is a 50% quota for female representation on a CDC). Those secondary hypotheses that relate to the relative effectiveness of different approaches will be relevant to only IRC project areas.

The pages that follow detail the primary hypotheses and what they will tell us about the impact of TUUNGANE. This is followed by a list and description of the relevance of secondary hypotheses. Concluding sections of this document provide additional technical details related to the evaluation, including the power calculations and a timetable for the data that will be available over the course of TUUNGANE’s implementation.

## 2 Evaluating TUUNGANE's overall impact

### 2.1 Does Participation in TUUNGANE Improve Social Cohesion?

TUUNGANE seeks to accomplish two types of outcomes with respect to social cohesion. The primary set of outcomes we seek to examine relates to cohesion among members of local communities: does participation in TUUNGANE improve levels of acceptance and trust within a community? The second is cohesion *across* communities: does participation in TUUNGANE lead to better relations and more dense exchanges?

We examine six hypotheses relating TUUNGANE to social cohesion:

- H 1 Individuals in communities that have taken part in TUUNGANE will exhibit higher levels of **acceptance of others** into their communities.
- H 2 Individuals in communities that have taken part in TUUNGANE will exhibit higher levels of **trust** in other members of their communities.
- H 3 TUUNGANE communities will be **more willing to contribute time and effort** individually to collective goods.
- H 4 TUUNGANE target **communities will be more likely to work together** to solve local development problems.
- H 5 Communities will be more **proactive in seeking support** from local government and NGOs for community initiatives and the private sector.
- H 6 Villages in TUUNGANE communities will have a greater propensity to **work collectively with other villages** to address development challenges.

If we find that, when comparing TUUNGANE villages to those in the control group, perceptions of social tension are lower, and levels of trust and willingness to contribute to collective goods and solve local development problems are greater, we will interpret this as evidence of program success with respect to the promotion of social cohesion.

### 2.2 Does Participation in TUUNGANE improve good governance?

Here, we focus primarily on within community governance. We will examine both the rights and the obligations that individuals feel to participate in community political decisions. We focus especially on the roles that they see for themselves relative to political and traditional leaders, and their willingness to hold these leaders to account.

We examine five hypotheses:

- H 7 Individuals in TUUNGANE target communities will report a greater sense of a **right to take part** in local decisions.
- H 8 Individuals in TUUNGANE target communities will report a greater sense of **obligation to take part** in local decisions.
- H 9 Individuals in TUUNGANE target communities will report greater **knowledge** about local decision-making processes and outcomes [what will be tested will depend on TUUNGANE training].
- H 10 Individuals in TUUNGANE target communities will report an increased **willingness to hold traditional and political leaders accountable**.
- H 11 Individuals in TUUNGANE communities are more likely to believe that **local leaders should be elected** rather than selected through an alternative mechanism.

If we find that, when comparing TUUNGANE villages to those in the control group, individuals feel greater rights and responsibilities regarding local decision making, know more about it, are more likely to question their leaders, elect them and hold them accountable, we will interpret this as evidence of program success with respect to improving good governance.

### **2.3 Does Participation in TUUNGANE foster socioeconomic recovery?**

An important issue in evaluating TUUNGANE is to determine whether in fact the program advances socio-economic recovery. Key will be to examine whether welfare improvements derive exclusively from the direct expenditures resulting from participation in TUUNGANE, or whether further gains result from secondary activities catalyzed by, if not directly resulting from, TUUNGANE initiatives.

- H 12: Access to community utilities and infrastructure, including those not directly supported by TUUNGANE, will be greater in TUUNGANE target communities.
- H 13: Household Income and asset holdings will be greater in TUUNGANE target communities.
- H 14: Households will allocate a greater share of their time to productive activities in TUUNGANE target communities compared with control.
- H 15: Time devoted to productive activities not directly associated with TUUNGANE projects will increase.

If we find that households in TUUNGANE villages allocate more time to productive activities, have greater access to utilities and services than in non-TUUNGANE villages, we will interpret this as evidence of program success with respect to economic outcomes.

## **3 Evaluating Individual Components of Program Design**

As part of a strategy for measuring the efficacy of community-based reconstruction programs, it is also important to assess the contribution of particular elements of program design. Slightly varying

this design within IRC areas and testing a series of related hypotheses will allow TUUNGANE to understand the relative effectiveness of important CDR elements.

Known as “variation in treatment,” this component of the evaluation centers on two key questions:

- Does greater women’s participation in decision-making produce better outcomes for governance, social cohesion and welfare?
- Is the direct delivery of cash more effective than village projects?

The secondary hypotheses in this section will guide data collection and analysis in IRC areas where variation is implemented.

As before, the key evaluation strategy is to use random assignment of treatment. Since there are no strong reasons to believe in advance that one treatment is inherently superior to the other, the use of different treatments with different communities does not raise ethical concerns. As before the fact that the different treatments are randomly assigned will allow us to attribute differences in outcomes clearly to the program design and not to other features of the communities in the program.

### **3.1 Does greater women’s participation in decision-making produce better outcomes for governance, social cohesion and welfare?**

A central principle of CDR programs is that by extending participation, a broader set of views gets represented which, in turn, results in more beneficial outcomes. By involving not just traditional elites but also youth and women in decision making, for example, better quality decisions emerge and receive broader support from communities. Formally however, opening decision-making processes to a broader constituency does not guarantee that a broader group of people can actually take part. In response to this concern, many political processes build in a formal requirement for minimal participation by vulnerable groups. Such a requirement can ensure that these groups have access to a forum for decision making; it does not guarantee that they will in fact be influential.<sup>8</sup>

Arguments against using a quota system include a concern not to restrict the freedom of communities to appoint their representatives and in particular a concern to ensure that

---

<sup>8</sup> One of the few compelling studies undertaken in this area examined the different outcomes when leadership roles were assigned to men and to women according to a random protocol in India. This study found that there were indeed marked differences in outcomes, with communities that were headed by women more likely to invest in clean water and road infrastructure. In some areas female leadership was also associated with greater female participation. See Raghavendra Chattopadhyay and Esther Duflo. “Women as policy makers: evidence from a randomized policy experiment in India.” *Econometrica*, Vol. 72, No. 5 (September, 2004), 1409–1443.

representatives are drawn from among the most able members, irrespective of their gender. Proponents of quotas can counter that the capacity of women to improve decision making is not known unless they have a real opportunity to participate; quotas thus may be an important tool for community learning. In short, there are good arguments in favor and against the idea of introducing gender quotas. It remains an empirical question whether or not they in fact return the benefits attributed to them.

### **3.1.1 Variation in program implementation**

To answer this key question TUUNGANE will pilot the use of gender quotas in a sub-sample of IRC project communities. The approach used will be to structure the ten member Community Development Committees (CDCs) and Village Development Committees (VDCs) in *all* areas such that there are two individuals in each of the five positions of Chair, Treasurer and Secretary, Community Mobilizer; and Community Outreach Coordinator. These officers will be selected through a secret ballot in all areas. In one half of the communities however there will be a formal requirement that one of each position be occupied by a woman (thus 50% of the committee will be composed of women). Randomization will take place among IRC communities and the same treatment will be applied to all villages in a given community.

The voting procedure will take the following form: subject to nomination and seconding, individuals can stand for each of the positions in the in turn (Chair, Treasurer and Secretary, Community Mobilizer, and Community Outreach Coordinator). Each individual in the electorate then casts a ballot in support of one candidate for each of the five positions. In the pilot areas, a first vote is held to select from among the pool of female candidates for a given position; a second vote is then used to select from among the pool of male candidates. In the non-pilot areas there are again two separate votes for each position, with the top ranking candidate elected on each vote, but in this case both votes are open to all individuals in the community irrespective of gender. The same procedure is then followed for each of the other positions. After each vote, winning candidates are eliminated from the pool of potential ordinary members. The electorate for this procedure at the village level (VDC) is the set of all citizens (over the age of 18). At the community (CDC) level the electorate is the set of all VDC members.

Communities are allocated to the pilot approach subject through a randomized assignment at the community level – thus all VDCs are assigned to the same treatment within a single community.

### **3.1.2 Hypotheses**

Comparing outcomes across these two groups allows us to test the following six hypotheses.

- H 16: Communities with greater representation of women leaders produce higher gains in the three priority goals of cohesion, governance and welfare.
- H 17: Greater representation by women leaders produces higher levels and quality of participation in decision making at the village and community levels.
- H 18: Greater representation by women leaders on CDCs leads to a greater propensity to accept women as decision makers.
- H 19: Greater representation by women leaders leads to lower levels of financial irregularities in the use of project money.
- H 20: At the end of the project, individuals, and especially women, in areas with greater women's participation will express greater satisfaction with the choice of projects.
- H 21: [Cross Intervention Hypothesis] Public goods are more likely to be selected than cash in communities with greater women's participation (see 3.2).

If we fail to find evidence that the quality of participation improves or that substantive outcomes are better under the pilot arrangement, we will take this as evidence that quotas play no role in improving outcomes. If we fail to find evidence that program choices and implementation are unaffected in any way we will take this as evidence that quotas are irrelevant to the decision-making process. In such an eventuality, survey outcomes and other data will be used to assess whether this is because women's interests are sufficiently represented in the absence of quotas, or because *quotas* by themselves are insufficient to ensure genuine representation.

### **3.2 Is the direct delivery of cash more effective than village projects?**

In post conflict environments there is a pressing need for access to cash and credit to purchase basic necessities such as seeds and tools. By some accounts, private agents have the best information about what their basic needs are and are highly motivated to ensure that money allocated for their private benefit is well spent. This argument is especially compelling in areas where there are high levels of distrust and in which institutions are weak. Moreover, it is not uncommon for donors to question why a large share of project funding is spent on facilitation, rather than passed through to civilians directly.

Counterarguments suggest that participation in the production and management of local public goods produces tremendous gains; the practice of working together and making decisions jointly can foster the trust needed for communities to undertake joint actions to reach longer term goals. Indeed this notion is at the heart of the CDR approach. The distribution of private goods on the other hand may lead to increased competition and tension.

Disagreements over what is the optimal approach have been prominent in a number of developmental areas including debates over the optimal use of oil and gas revenues and discussions over the most effective way to engage ex-combatants in demobilization and reintegration programs.

In the absence of robust empirical evidence supporting one position or the other, it remains unclear whether and when recourse to cash transfers is optimal in post conflict situations.

### **3.2.1 Variation in program implementation**

To address this key question, TUUNGANE will pilot the use of a cash transfer option in a sample of communities. The pilot will be limited to decisions made at the village level (where total budget sizes are in the order of £1 per person or perhaps \$8 - \$10 per household) rather than to decisions made at the community (CDC) level (where decisions are on the order of £7 per person or perhaps \$50 - \$80 per household). As with the gender pilot, this variation will be limited to IRC areas.

The procedure is likely to work as follows. All villages will produce a listing of households which records the number of individuals presently over the age of 12. These lists are publicly endorsed at a village meeting through a no-objection procedure (a consensus rule). When villages choose how to allocate financing for village initiatives they face the broad set of options conventionally available in CDR programs. Pilot villages, however, will have the additional option of electing to distribute the funding available for the village projects directly to citizens in the form of cash. If this option is considered, then some per household allocation is to be recorded on the village list and this allocation becomes an option on the menu of possible projects. If the proposal is accepted the distribution is to be done by the VDCs subject to monitoring by IRC staff.

Communities are allocated to the pilot approach through a randomized assignment (not a public assignment).

### **3.2.2 Hypotheses**

Comparing outcomes across these two groups allows us to test the following hypotheses.

- H 22: Communities that select cash transfers will exhibit subsequent extorted or higher levels of competition and conflict.
- H 23: Program gains in the areas of governance, cohesion, and welfare will not be as great in communities that select cash transfers.
- H 24: The use of cash transfers leads to private gains at least as great as the use of public goods.
- H 25: The gains from cash transfers are greatest in more severely conflict affected communities.
- H 26: Individuals in TUUNGANE villages that choose cash transfers will report a weaker sense of rights to take part in local decisions than villages that choose projects.
- H 27: Individuals in TUUNGANE villages that choose cash transfers will report weaker sense of obligation to take part in local decisions than villages that choose projects.
- H 28: [ENGAGEMENT] Individuals in TUUNGANE villages that choose cash transfers will report a weaker propensity to question the actions and decisions of local leaders, traditional chiefs and CDC leaders than villages that choose projects.

H 29: Individuals in TUUNGANE villages that choose cash transfers will report less willingness to hold local leaders, traditional chiefs and CDC leaders accountable than villages that choose projects.

In addition to testing these hypotheses, the project can establish what cash transfers are used for. Cash transfers could, in principle, be aggregated up by communities and used for the implementation of community projects or they could be spent privately. In the former case a direct comparison of treatment and control villages can be undertaken to establish the economic benefits of TUUNGANE support for project implementation.

It is important to note that the ability to identify effects using variation in treatment of this form depends on the actual choices to opt into the treatment by villages. Villages are not *required* to take the cash option and the decision to do so may reflect such features as the level of distrust (or trust) they have in their communities, or more directly, the type of needs they face. These selection effects limit our ability to identify the effect of cash allocations to the extent that they result in a small share of the pilot groups electing to select the cash option. If all pilot communities select the cash option then we will be able to identify the joint effect of the cash transfer and the *option* to have cash transfer. If no pilot communities select the cash transfer then we will observe the effects of the option but not of the use of the deployment cash transfers (otherwise known as “intention to treat.”)<sup>9</sup>

#### **4 Evaluating the Impact of Contextual Factors**

The TUUNGANE program will provide assistance to 280 communities and over 1000 villages throughout East Congo. Villages differ from one another on many dimensions that might affect how the program unfolds and how projects are ultimately implemented. As part of the evaluation effort, we will also ask: In what environment are community-based approaches to reintegration most likely to work and under what conditions will they not work?

We expect to look at factors related to the external environment to figure under what circumstances the CDR approach is most likely to work. This is essential to determine the extent to which the lessons learned in DRC can be applied to other contexts. In addition, a focus on contextual factors allows us to control for these factors in examining the extent to which program implementation is different among TUUNGANE implementation sites.

##### **4.1 Hypotheses**

To examine contextual features the evaluation project will test the follow hypotheses.

---

<sup>9</sup> A related issue arises with respect to the gender quotas. Insofar as communities may select women representatives even in the absence of quotas the true treatment is the use of quotas rather than gender parity.

- H 30: [History of conflict] The impact of TUUNGANE does not depend on the level of conflict intensity.
- H 31: [Wealth] The impact of TUUNGANE does not depend on the level of village wealth.
- H 32: [Cohesion] Improvements in social cohesion and governance does not depend on prior levels of cohesion or decision making norms.
- H 33: [Population flows] The impact of TUUNGANE does not depend on the numbers of returnees and new migrants.
- H 34: [Consistency of Impact across TUUNGANE Regions] Controlling for contextual factors, program results are of equal quality across all four project areas.

These five hypotheses have been phrased negatively. For most of these however effects could run in either direction. If we do find a positive relation, for example, between conflict intensity and program effectiveness we will reject the null hypothesis of no effect. We interpret this as evidence in support of the claim that CDR approaches are especially effective in post conflict cases.

For these, and for all contextual hypotheses, two notes are in order. The first is that the identification strategy, while first best with respect to the program impacts, is only second best with respect to the contextual features themselves; it is always possible that other features, different from the contextual features that we examine, are what truly determine the variation in program impact.<sup>10</sup> The second is that if we draw lessons with respect to other contexts the lessons we can learn do not allow us to compare one program with another – for example CDR compared with another approach – but rather whether CDR is more effective in one place or another.

The final contextual hypothesis focuses especially on interactions with other programs.

- H 35: [Program Complements / Substitutes] Program impact will be weaker in communities where there has already been support at the village or community level from international organizations.

---

<sup>10</sup> The reason is that while we can randomize where the program takes place, we cannot randomize features about those places, such as their prior wealth levels.

## 5 Understanding Unintended Consequences

In addition to examining the *intended* consequences of TUUNGANE, the evaluation will explore a set of possible unintended consequences. All of our tests on the hypotheses listed herein will allow us to identify whether effects actually operate in the *opposite* direction to those expected. We will also seek to examine other potential unintended consequences.

### 5.1 Hypotheses

These final hypotheses are as follows.

- H 36: **Migration:** The TUUNGANE program will not result in higher levels of in-migration to TUUNGANE target communities.
- H 37: **Jealousies:** The TUUNGANE program will not result in higher levels of discontent in control communities that are proximate to TUUNGANE communities.
- H 38: **Effects on other programs:** Participation in TUUNGANE will not result in inequities from the targeting of other programs disproportionately to areas in which TUUNGANE is already operating (other than through demand effects resulting from community initiatives).
- H 39: **Inflation:** The TUUNGANE program does not have adverse effects on local inflation.

Again these hypotheses are given in a negative form; they reflect the assumptions of the CDR program design. Rejection of these hypotheses will be taken as evidence that these assumptions are invalid. The measures used to examine heterogeneity in program effect are in large part to be collected at the village level and through chief surveys (see appendices for the baseline questionnaire).

## **6 Tools and Data Resources**

A mix of tools and data resources will be employed as part of the TUUNGANE evaluation. In addition to the data that result from household and chief surveys the analysis will be strengthened with data on process and project implementation that will provide key information for examining variation in treatment and for examining the progress and/or obstacles of the program over its lifetime. For the larger questions of impact, however we will not have process data from the control villages and will rely instead on the larger surveys.

The following list of data resources to be collected by IRC and partners is illustrative, not final. At this stage of program implementation (Summer 2007), the final components include the household and chief surveys. We will wait until NGO partners have begun implementation to continue to discuss various options for gathering supportive data, as well as additional research activities depending on resources, time and logistics.

### **6.1 Household and Chief Surveys [2007 - 2010]<sup>11</sup>**

A quantitative survey of households will be conducted before and after project implementation. The first of these will be conducted in July – August 2007 (see schedule in Section 7.1); the second will be conducted three years later. The core purpose is to generate the primary quantitative source of data for measuring the overall impact of the program through the comparison of changes in “treatment” and “control” communities. It will include detailed measures of social cohesion, attitudes towards governance and welfare. It will also include measures designed to capture key contextual variables that can be used to understand when and where CDD projects are most effective. This survey will be administered to 5 respondents in 560 villages – one in each of the treatment and control communes. In addition a chief survey will be administered to the village chiefs of each of these 560 villages capturing key data on village characteristics as well as the attitudes of local political leaders. Finally, a smaller household survey is anticipated for Spring of 2008 to capture key outcomes of the cash transfer intervention including data on decision making and use of funds.

### **6.2 Participation Data**

To facilitate a substantive investigation of how participatory decision-making processes work in practice, we expect to use a protocol for recording detailed information about attendance and participation in village-level meetings. The goal of this data collection effort will be to establish how patterns of engagement in the decision-making process vary across groups in environments where committees differ in structure (given the gender quota) or where the stakes are different (because of the cash option). This will also serve to supplement conventionally quantitative output indicators (for example, the number of meetings or development plans) with more nuanced data related to the quality of participation in these meetings.

---

<sup>11</sup> Baseline household and chief survey tools are attached as Appendix 8.3.

### 6.3 Village Profile and Log

Facilitators will maintain a community profile document capturing key information used for the analyses in Section 4. These include data collected in the chief surveys (village background data, village political structures, local conflict management mechanisms, and village demographics and social structure). In addition, village logs, based on quarterly reports, will maintain information on new projects starting in communities, tensions arising, conflict events, and details of implementation difficulties.

### 6.4 Project Records

Project records will keep account of proposals submitted for discussion, projects funded, and, in cases where beneficiaries are targeted, information about the basic characteristics of those targeted. In addition regular data will be gathered through the TUUNGANE complaints mechanism regarding concerns raised about process and implementation in each village in order to provide outcome measures for both of the variation in treatment evaluations.

### 6.5 Accounts

In order to examine the impact of women's involvement on financial probity as well as to generate measures of variation in success the project will use reports on project-related conflicts, concerns over corruption or attempted extortion. An important source of data will be village accounts and auditing reports which will include information on prices paid for different types of goods and services.

### 6.6 Events Data

Depending on resources and access to data, TUUNGANE M&E staff will keep track of conflict events.

## 7 Evaluation timeline and outputs

### 7.1 Timeline

#	Activity	Timeline
1	Survey Implementation	July – August 2007
2	Randomization	August 2007
3	Program Implementation Begins	August 2007
4	Initial Baseline Report Submitted to IRC	30 September 2007
5	Final Baseline Report Submitted to DFID	30 October 2007
6	Program Monitoring	August 2007- 2010
7	Process Evaluation	August 2008 & 2009
8	Variation in treatment research – likely to include rapid surveys	Cash variation September 2008
9	Follow Up Survey	August 2010
10	Final analysis of program impact	Fall 2010

## 7.2 Output Description

**Baseline Report** – Based on data from both household and chief surveys, the baseline report will have three main sections: background characteristics of the sample frame communities; detailed information about the pre-intervention status of three main outcomes TUUNGANE seeks to improve; and an analysis of whether or not there are any systematic differences among treatment and control communities (randomly assigned by that time). This last section is useful for demonstrating whether communities selected for the CDR program are similar in all other ways to those that will be tracked as part of the control group.

**Program Monitoring** – The ongoing collection and analysis of program data by a team of M&E staff will provide TUUNGANE’s management and donor with a regular source of information about the program’s progress, activities and outputs completed. Additional qualitative data such as those mentioned in section 6 – to be confirmed with the TUUNGANE team this fall – will provide more targeted information on key areas of interest – the quality of participation in community meetings, people’s satisfaction with the decision-making process, the existence and resolution of local conflicts are all examples.

**Process Evaluation** – Depending on resources, one or two mid-term process evaluations will analyze the monitoring data and investigate any issues related to program implementation. This may include questions such as whether or not targeted populations are indeed part of the CDR process, whether the program is being implemented according to its design across the three NGO partners, and whether or not the appropriate staff and inputs are in place.

**Variation in treatment research** – The two variations in treatment that will be implemented in select IRC areas will be analyzed with complementary research activities, conducted and written up for TUUNGANE stakeholders.

**Final analysis of program impact** – Produced by the senior consultant to the project Macartan Humphreys, the final analysis of program impact will use all available data to estimate the “program effect” of TUUNGANE and answer both primary and secondary hypotheses described herein.

## 8 Annex:

### 8.1 Randomized Allocation of Treatments

There are two types of variation in treatment to be implemented at the community level:

1. Gender balance requirement
2. Public / Private Goods Option

We allocate these treatments in a balanced manner as follows.

**Table 1: Gender and Cash Treatments**

	Cash Allocations Permitted	Only Community Projects Permitted
Gender Balance Requirement	35 (18, 17)	35 (17, 18)
No Gender Balance Requirement	35 (18, 17)	35 (17, 18)

**Note:** Cell entries represent the number of total number of communities receiving each of the same combinations of treatments; the numbers in parentheses report the number with each of the two regions, Sud Kivu and Haut Katanga, receiving each of the same treatments.

To ensure balance, this division will be further stratified by Territoire. Further stratification may be possible depending on the ultimate pool of communities (information that will not be available until the main randomization) to ensure balance on other covariates gathered from baseline survey data and to minimize spillover effects.

### 8.2 Power

Power calculations are used to estimate the likelihood that we will be able to identify an effect of a given size, at a given confidence level, conditional on a given data structure and a given sample size. It is always possible that a true program effect exists but that a research strategy will nevertheless fail to identify it with statistical confidence. The likelihood that a true effect will be found is called the *power* of the test. Small effects are more difficult to identify than larger effects and reduce power. Larger sample sizes produce more power. Samples that have a lot of clustering have less power. Exact power calculations are difficult to undertake since they depend on unknowns, most importantly the structure of the data and the size of the true program effects. Nonetheless we can generate power estimates conditional on a set of assumptions regarding each of these features.

The targeted sample for the baseline survey is 2800 subjects with 5 subjects randomly selected from each of 560 clusters, with one cluster, corresponding to a village or sous-village, in each of the

treatment and control communities. In this section we describe the statistical power that we expect to gain from a sample this size given the structure of our data and compare this power to what could be achieved with smaller and larger sample sizes.

We assume a model of the following form:

$$Y_{ij} = \alpha + \beta T_j + u_j + \varepsilon_{ij}$$

where  $T_j$  is an indicator for whether or not an observation in village  $j$  receives treatment,  $u_j \sim N(0, \tau^2)$  is a village specific error and  $\varepsilon_{ij} \sim N(0, 1)$  is an individual specific error (we ignore community specific errors). In the analyses that follow we report the power associated with identifying a treatment effect for different values of  $\beta$  as  $\tau^2$  (the “between group variance”) varies from 0 to 2.<sup>12</sup> Since we have normalized the variance of the individual specific errors to 1 all other quantities are interpreted in units of this error.

Figure 1 shows the probability that we will be able to reject a null of no effect at the 95% level given a true effect of 1/8, 1/4 and 1/2 of a standard deviation for different levels of between group variance. The two panels examine the power for both the overall treatment/control comparisons and for the variation in treatment (within IRC treated communities) case. In each graph we assume 5 observations per community.

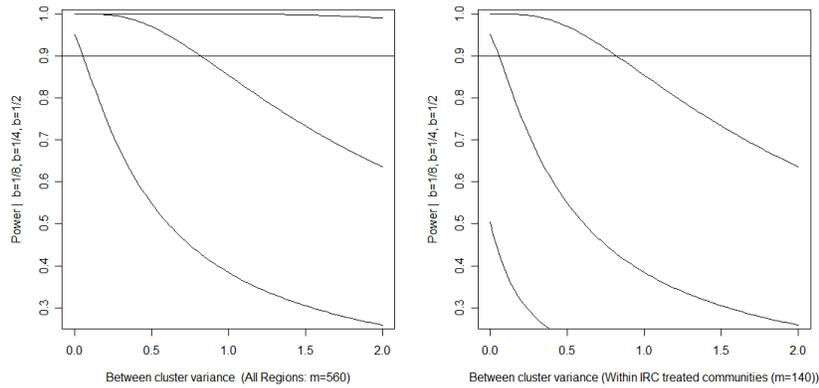
The left figure is for the case with only IRC communities showing that with 5 observations per site we have excellent power for the main analysis (left panel) for effect size ( $\beta$ ) of 1/2 or 1/4 of a standard deviation, even for between cluster variance of up to 1. We have reasonable power for effect sizes of 1/8 only if there is very weak clustering.

The right figure shows the power for variation in treatment. Note that for these power calculations the number of observations is only one quarter of the number for the full sample; a result of this is that if we have a given power for a given effect,  $\beta$ , in the full sample, we will have the same power only for effect  $2\beta$  in the small sample; this is what explains the similarity between these two graphs.

---

<sup>12</sup> In principle  $\tau^2$  is not bounded from above; high values are possible if cluster effects dominate individual effects. Much of the interpretation of what follows depends on an understanding of  $\tau^2$ . For one interpretation of  $\tau^2$  imagine an individual for whom a share  $\beta$  of the change in his welfare depends on a coin toss that affects him only and another share  $(1-\beta)$  depends on a coin toss that affects the whole village. In this case the individual specific variance is  $\beta^2/4$  and the town specific variance is  $(1-\beta)^2/4$ . The town specific variance as a fraction of the individual variance is  $\tau^2 = ((1-\beta)/\beta)^2$ . Thus we have  $\tau^2=1$  if  $\beta=1/2$  and so income is drawn equally from the two sources; we have lower values however when the individual’s income is determined more by the coin toss that affects him only: thus  $\tau^2=.25$  if  $\beta=2/3$  and  $\tau^2=.1$  if  $\beta \approx 3/4$ .

**Figure 1: Power and Effect Size (Individual Level Outcomes)**

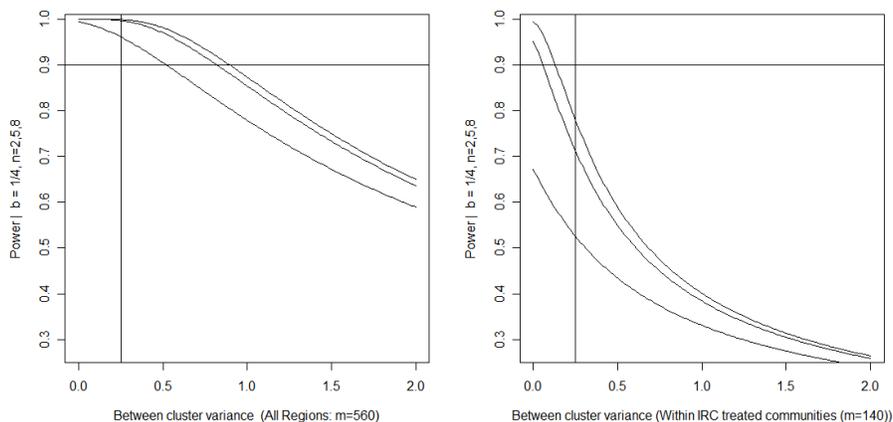


Substantively for the variation in treatment we have excellent power for an effect size of one half of a standard deviation, even for between cluster variance of up to 1, we have reasonable power for effect sizes of 1/4 only if there is very weak clustering, finally, power is very poor for effect size of 1/8 of a standard deviation.

*In summary, the statistical power associated with our survey strategy is likely to be sufficient to pick up even relatively small effects for the main analysis. Power is substantially weaker for the variation in treatment analysis but is still likely to be sufficient for identifying medium to large effects.*

In Figure 2 we show how power in the middle scenario, for effect size of 1/4, depends on the cluster size. To do so we examine cases with smaller and larger cluster sizes. The lower lines in these graphs show the case in which rather than selecting 5 observations per village for a total sample size of 2800 we select only 2, for a sample size of 1120. The upper lines show the case in which we sample 8 per village for a sample size of 4480. In both cases the middle lines are the same as before.

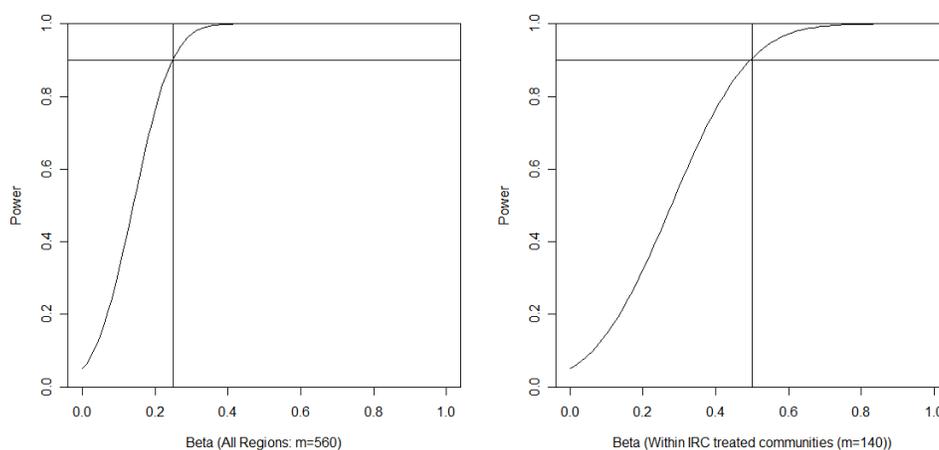
**Figure 2: Power and Cluster Size (Individual Level Outcomes)**



From this analysis we can see that the power gains from moving from 2 to 5 respondents per village are substantial under all scenarios, save the full sample case with clustering close to 0. Increasing beyond this to 8 respondents per cluster produces only modest gains in the full sample case. For the variation in treatment power gains are on the order of 5 to 10% for modest clustering. Increases like this are however expensive for the baseline before the set of IRC communities are known but can be envisioned for the follow up analysis.

The third and final set of graphs show power for “village level” outcomes as a function of true effect sizes.

**Figure 3: Power (Village Level Outcomes)**



**Assuming just one observation per village (that is, a village level outcome), then, as show in**

Figure 1, we have sufficient data to identify an effect of .25 of a (village level) standard deviation for the full sample treatment and control analysis. Our power is greatly reduced however for the within IRC (variation in treatment) analyses. As indicated above, power at the 90% level is reached only for effect size of approximately one half of a standard deviation.

### 8.3 Baseline Survey Instrument and Manual

The baseline survey instrument along with the chief survey, manual and codebook is attached as an annex to this document.