Uganda survey in the refugee-hosting districts

1. Overview

The objective of the survey is to generate panel evidence from the refugee and host households living in the refugee-hosting districts[1] of Uganda to monitor the implementation of the Refugee Response Plan (RRP).

2. First wave: 3 baselines Nov/Dec 2017 - April/May 2019

The first wave is composed of three separate surveys. The three baselines have been administrated in 2017, 2018, and 2019 as shown in table 1. Overall, a total sample of 6,236 households (including refugee and host) have been interviewed in 10 districts. The districts of Moyo, Lamwo, and Kyegegwa have been visited twice but different households have been interviewed.

- The objective of the 2017 data collection was to understand the food security and resilience condition of the households living in the Northern districts. The sample of the 2017 survey has been designed to be representative at the district and settlement level[2], with the refugee population being oversampled.
- The 2018 survey aimed at extending the 2017 survey in the Southwester districts. The host communities have been identified as the closest communities living in the same sub-county.
- Finally, the 2019 survey responds to a different objective. Indeed this is a baseline survey for the ACREI project covering both beneficiaries and control households in host and refugee communities. The households have been selected based on the impact evaluation sample design.

Table 1. Survey coverage by baselines

District	Settlement	# Refugee HHs	# Host HHs	Dates
Yumbe	Bidibidi	290	240	Nov / Dec 2017
Moyo	Palorinya	200	166	Nov / Dec 2017
Adjumani	Adjumani	200	153	Nov / Dec 2017
Lamwo	Palabeck	200	166	Nov / Dec 2017
Arua	Imvepi	208	199	Nov / Dec 2017
Arua	Rhinocamp	289	196	Nov / Dec 2017
Kiryandongo	Kiryandongo	325	202	Nov / Dec 2017
Kyegegwa	Kyaka II	202	158	Mar 2018
Kamwenge	Rwamwanja	193	152	Mar 2018
Isingirio	Nakivale	333	377	April/May 2019
Arua	Omugo (Rhinocamp extension)	231	169	April/May 2019
Kikuube	Kyangwali	150	150	April/May 2019
Kyegegwa	Kyaka II	217	165	April/May 2019
Lamwo	Palabeck	190	140	April/May 2019
Moyo	Palorinya	226	149	April/May 2019
Tot.		3,454	2,782	

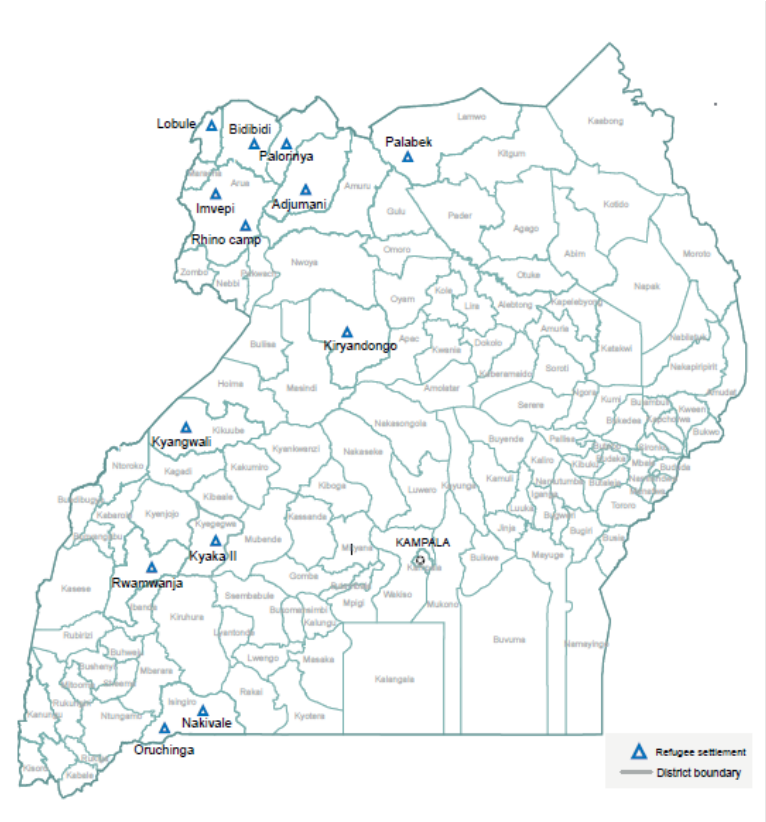
[1] Arua, Yumbe, Moyo, Adjumani, Lamwo, Kiryandongo, Kyegegwa, Kamwenge, Isingiro Kikuube and Koboko.
[2] Population weights can be calculated using UNHCR/OPM population data for refugee and host households.

Table 2. Survey coverage by district

District	Settlement	# Refugee HHs	# Host HHs
Yumbe	Bidibidi	290	240
Moyo	Palorinya	426	315
Adjumani	Adjumani	200	153
Lamwo	Palabeck	390	306
Arua	Imvepi	208	199
Arua	Rhinocamp	289	196
Arua	Omugo	231	169
Kiryandongo	Kiryandongo	325	202
Kyegegwa	Kyaka II	419	323
Kamwenge	Rwamwanja	193	152
Isingirio	Nakivale	333	377
Kikuube	Kyangwali	150	150
Tot.		3,454	2,782

The total sample by settlement and refugee/host is shown in table 2 while figure 1 shows the map of the survey coverage.

Figure 1 Map of survey coverage



Source: Elaboration by FAO Uganda, 2020

Sampling design

The current task is to perform a baseline study that would serve to collect indicators on food security, resilience capacity and well-being of both refugee population and host communities.

The sampling analysis benefits of previous RIMA studies implemented in the Karamoja region and South Sudan.

Cluster sampling is appropriate for this exercise; households are concentrated in settlements (refugees) or villages (host communities). Upon discussion with local people, country office, ngos, three clusters have been selected for sampling; 1) refugees and 2) host communities living separated; and 3) refugees and host communities living together. This will help understanding the relations put in place. An example of how and why such clusterization is relevant to this study comes from land allocation. Refugees' communities are provided with a piece of land for "backyard plot", which is within the small plot of land allocated for dwelling and cropping upon the registration of the refugees. However sometimes refugees and host communities live together in the same area; this can create tensions for the land use. In other cases, although the two communities are separated, some refuges can access land (upon renting or free donation) otherwise used from the host communities. Consequently, by having access to data on the three clusters will help a better understanding of the internal dynamics.

The primary sampling units (PSU) will be the single settlement or village. Households will be the Second Sampling Unit (SSU); they will be randomly selected from either a list of household provided by local authority or by walking through the village or settlement.

Sample Size Calculation based on the Resilience Capacity Index.

This survey will serve multiple purpose, investigating food security, livelihoods strategies, poverty, vulnerability and resilience of both host and refugees communities. Consequently, it is complex undertaking to design a proper sample that can serve all the purposes.
When a sampling strategy needs to satisfy multiple indicators (i.e. providing an adequate number of households that can be statistically significant for different variables), it is good practice to utilize the minimum requirements of the indicator with the largest required sampling size. In this case, the resilience capacity index (as estimated by RIMA, FAO) is adopted.
Otherwise, the entire sampling strategy is designed around the comparison between refugees and host communities.

In order to obtain the sample size the following information are required: expected impact of the projects; power of design (in order to reduce the probability of a Type II error - incorrectly accepting the null hypothesis); standard deviation of the indicator of interest; level of confidence (in order to reduce the probability of a Type I error – incorrectly rejecting the null hypothesis).
Based on previous experiences on similar exercises (see FAO, 2016) and on the occurrence of various projects, the sample size will be calculated in order to be able to detect at least an impact of 10%. In terms of statistical power the frequently adopted value of 0.90 will be chosen. In terms of level of significance, the frequently adopted value of 95% will be chosen. Other assumptions for the sample size estimation are:

- An expected null difference in resilience capacity between refuges and host communities at the beginning of our study;
- The possibility of detecting a possible increase of difference between the two groups after a short period.
- A correlation between the resilience capacity of the two groups that ranges between .4 and .5 (average correlation).

- Initial differences in the standard deviation between the two groups (i.e. we may expect more homogenous socio economic characteristics for the host communities and a more heterogeneous context for the refugees people).

Considering the inherently fluid nature of refugees' status, a 20% oversampling will be adopted in order to control for possible attrition in case of follow-up data collection. Intra-cluster correlation is also considered. An example of the trade-off between power and numerosity of sample (where it is assumed that the baseline value for RCI is 17% (with standard deviation of 0.3) and an increase which ranges between 7 to 13% is achieved) is reported in Figure 1.

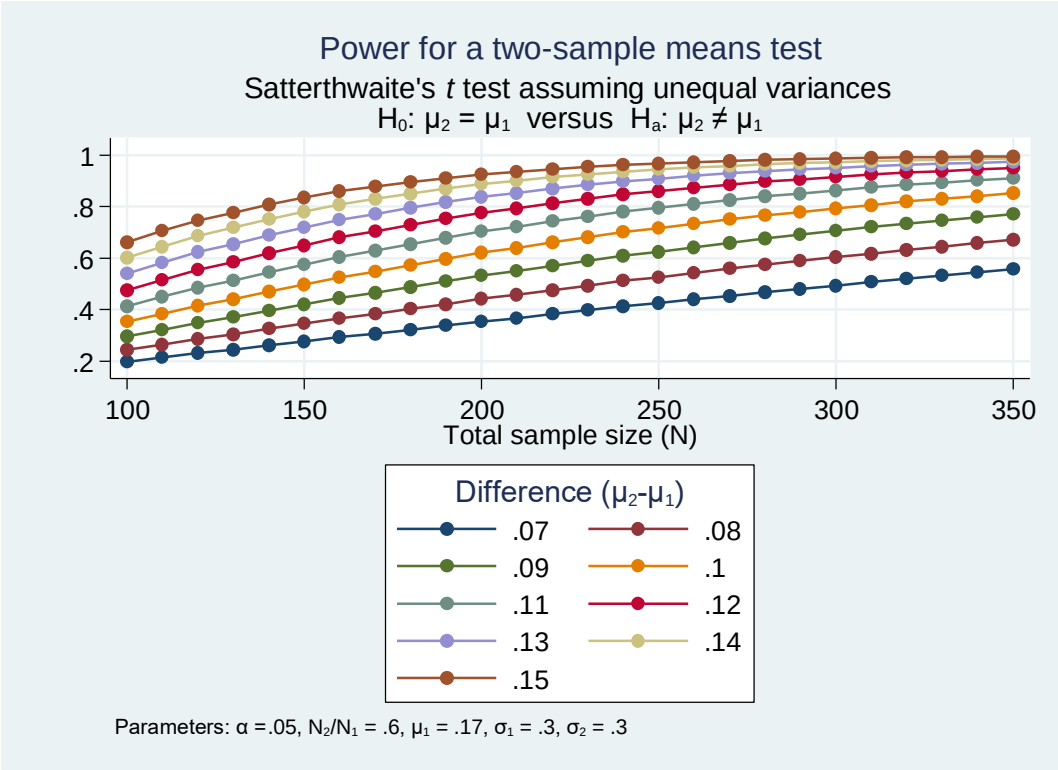
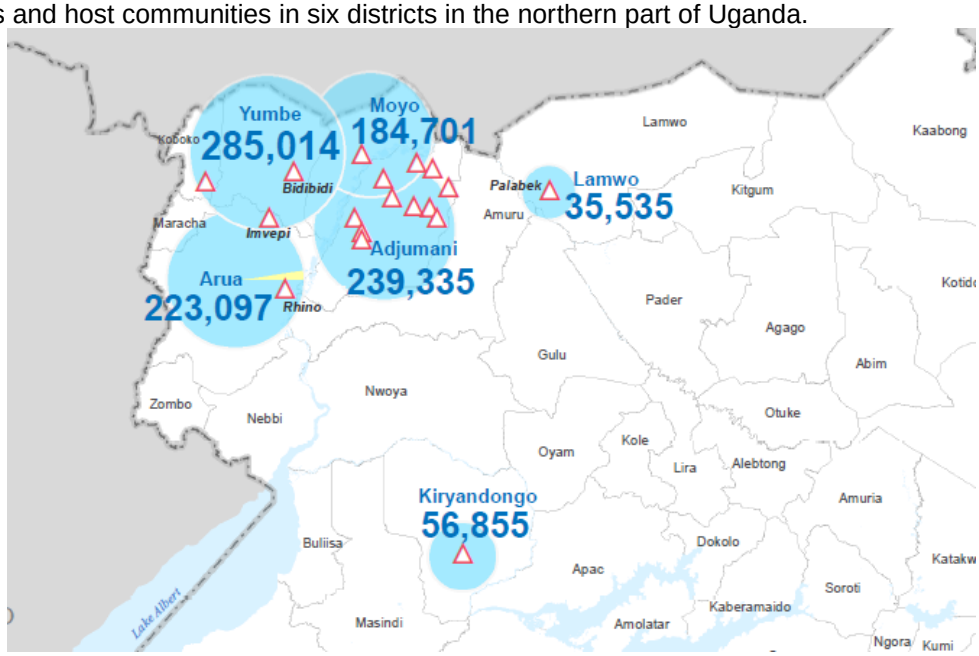


Figure 1: Sample size variation with power

A proper example of sample size is reported in table below:

alpha	power	N	delta	d0	da	ma1	ma2	sd_d	sd1	sd2	corr
0.05	0.8	406	0.1394	0	3	50	53	21.52	17	15	0.1
0.05	0.8	362	0.1478	0	3	50	53	20.3	17	15	0.2
0.05	0.8	317	0.1579	0	3	50	53	19	17	15	0.3
0.05	0.8	273	0.1704	0	3	50	53	17.61	17	15	0.4
0.05	0.8	228	0.1864	0	3	50	53	16.09	17	15	0.5
0.05	0.8	184	0.208	0	3	50	53	14.42	17	15	0.6
0.05	0.8	139	0.2394	0	3	50	53	12.53	17	15	0.7
0.05	0.8	95	0.2914	0	3	50	53	10.3	17	15	0.8

The Uganda Bureau of Statistics (UBOS) and the Office of Prime Minister (OPM) provided the sampling framework adopted. It is based on the available information on the composition of the refugees' settlements and host communities in six districts in the northern part of Uganda.



While more detailed information will be presented in a separated excel file, here below we report a table with the overall sample. The focus will be on the seven settlements were South Sudanese refugees are located. The three groups (clusters) are under the "refugees camps" column; host communities; and the two columns "refugees from mixed areas" and "hosts from mixed areas".
The representativeness of this sample is twofold: at settlement level (i.e. for the 7 settlements) and for the three communities (refugees; hosts; mixed).

Settlement	Refugees Camps	Host Community	Refugees from mixed areas	Hosts from mixed areas	Household per site
Bidibidi	*280*	*240*			520
Palorinya	*200*	*160*			360
Adjumani	*200*	*160*			360
Palabeck	*100*	*160*			260
Imvepi	*200*	*160*			360
Rhinocamp	*280*	*240*			520
Kiryandongo	*280*	*240*			520
Total	**1,300**	**1,120**			**2,900**