TRACING THE DEVELOPMENT OUTCOMES OF AGRICULTURAL INVESTMENTS IN SIX RESEARCH LOCATIONS IN TANZANIA AND UGANDA: QUESTIONNAIRE SURVEY DESIGN AND SAMPLE CHARACTERISATION

Helle Munk Ravnborg, Bernard Bashaasha, Rikke Brandt Broegaard, Michael Byaruhanga, Evelyne Lazaro, Festo Maro, Khamaldin Mutabazi, Teddy Nakanwagi and David Tumusiime

### Acknowledgements

Working Papers make DIIS researchers' and partners' work in progress available to readers prior to formal publication. They may include documentation which is not necessarily published elsewhere. DIIS Working Papers are published under the responsibility of the author alone.

This DIIS Working Paper has been prepared as part of the Agricultural Investors as Development Actors (AIDA) research programme, which is funded through the Danish Council for Development Research under the Danish Ministry of Foreign Affairs. The AIDA programme involves scholars from DIIS and Copenhagen University, both in Denmark, Makerere University in Uganda and Sokoine University of Agriculture in Tanzania.

**Helle Munk Ravnborg** Senior Researcher hmr@diis.dk **Bernard Bashaasha** Professor bashaasha@caes.mak.ac.ug **Rikke Brandt Broeggard** Associate Professor rikke.broegaard@ign.ku.dk Michael Byaruhanga PhD Candidate bmbyaruhanga@gmail.com **Evelvne Lazaro** Senior Research Fellow lazaroa@sua.ac.tz **Festo Maro** PhD Candidate marofesto@gmail.com Khamaldin Mutabazi Senior Lecturer khamaldin2011@gmail.com Teddy Nakanwagi PhD Candidate trizan2002@yahoo.com **David Tumusiime** Associate Professor tumusiime@caes.mak.ac.ug

### **DIIS WORKING PAPER 2021: 05**

DIIS · Danish Institute for International Studies Østbanegade 117, DK-2100 Copenhagen, Denmark Tel: +45 32 69 87 87 E-mail: diis@diis.dk www.diis.dk ISBN 97887-7236-044-7 (pdf) DIIS publications can be downloaded free of charge from www.diis.dk © Copenhagen 2021, the authors and DIIS

# TRACING THE DEVELOPMENT OUTCOMES OF AGRICULTURAL INVESTMENTS IN SIX RESEARCH LOCATIONS IN TANZANIA AND UGANDA: QUESTIONNAIRE SURVEY DESIGN AND SAMPLE CHARACTERISATION

Helle Munk Ravnborg, Bernard Bashaasha, Rikke Brandt Broegaard, Michael Byaruhanga, Evelyne Lazaro, Festo Maro, Khamaldin Mutabazi, Teddy Nakanwagi and David Tumusiime

 $\mathsf{DIIS} \cdot \mathsf{DANISH}$  INSTITUTE FOR INTERNATIONAL STUDIES

## **TABLE OF CONTENTS**

Abstract	3
Selection and delimitation of research locations	4
Sampling approach for drawing the six independent samples	11
Stage 1: geographically stratified, proportional random selection of villages	12
Stage 2: Proportional random sampling of individuals	15
Administering the questionnaire survey	20
Administering the questionnaire survey in Tanzania	20
Administering the questionnaire survey in Uganda	21
Data cleansing	23
Sample characterisation	24
Demographic characterisation of samples	24
Characterisation of the parcels about which information was provided	27
Spatial characterisation of samples	28
Density of foreign agricultural investments	31
Household poverty profile of the samples	35
Concluding remarks	43
References	44

### ABSTRACT

This paper describes a questionnaire survey conducted as part of the Agricultural Investors as Development Actors (henceforward AIDA) research programme to establish the development outcomes of selected foreign agricultural investments for people living in the vicinity of such investments in six research locations, three in Tanzania and three in Uganda.

The AIDA programme focuses on development outcomes in terms of

- employment, migration, food security and wider dynamic economic effects, such as access to technology, infrastructure and markets;
- land markets and perceived security of land tenure; and
- water access and security of tenure.

The AIDA research programme is being conducted by a team of researchers from Makerere University, Kampala, Uganda; Sokoine Agricultural University, Morogoro, Tanzania; Copenhagen University, Copenhagen, Denmark; and the Danish Institute for International Studies, Copenhagen, Denmark, under a grant from the Danish Development Research Council, Ministry of Foreign Affairs, Denmark.

In addition to establishing the presence and reach of development outcomes associated with foreign agricultural investments, the AIDA questionnaire survey aims to establish the extent to which development outcomes differ for different segments of populations living in the vicinity of foreign agricultural investments defined by, for example, age, gender, socio-economic status, occupation, origin, or distance from the investment location. The questionnaire survey was developed jointly by the AIDA research team, informed by joint exploratory fieldwork carried out in the six research locations during 2017. Interviews for the questionnaire survey were conducted from February to May 2019.

Following a brief description of the selection and delimitation of the six research locations, the paper describes the sampling approach used to draw six independent samples of approximately four hundred respondents each and the way the survey was administered in each of the six research locations. The paper concludes by describing the demographic, geographical and socio-economic characteristics of the six samples.

### **SELECTION AND DELIMITATION OF RESEARCH LOCATIONS**

As a starting point, the selection of research locations for the AIDA research programme was informed by the presence of foreign agricultural investments, more specifically Danish investments of this type, aided by the privileged access enjoyed by the AIDA programme by virtue of it being funded and coordinated from Denmark. The AIDA research programme focuses on such investments in sub-Saharan Africa, particularly in Tanzania and Uganda, which

- involve Danish private capital, whether solely or in combination with ODA or other forms of 'soft' capital;
- involve the acquisition of land rights, e.g. through purchase, lease or rent, either personally or as part of a company, or imply direct control over land use, e.g. through contract farming that obliges land-rights holders to conduct their productive activities in a specific way on a permanent or a temporary basis; and
- engage in primary production, i.e. in plant production (annual as well as perennial crops, such as coffee, tea, jatropha, etc. whether for food, fodder or fuel), or in animal production (livestock as well as poultry, but not aquaculture) as a commercial activity.

Informed by a mapping of Danish agricultural investments in sub-Saharan Africa (Broegaard and Ravnborg, in preparation), six agricultural investments were selected which meet the above criteria, three in Uganda and three in Tanzania. In the following, we refer to these investments as the 'AIDA investments'. In addition to meeting the criteria described above, in each country the investments are located in zones that are mutually different with respect to, for example, agro-ecological conditions and agrarian structure, while at the same time allowing cross-country comparisons with respect to factors such as crop types, type of management (farm size, type of employment of local population as permanent and/or temporary employees, suppliers, outgrowers, contract farmers etc.) and the type of financing used (private individual investment, private institutional investment or some sort of blended public and private financing) (Table 1).

Tanzania	Uganda
Investment I – production of flowers, small	Investment IV – small scale, coffee production,
scale, labour-intensive, small private company	labour-intensive, private company with little or
with support from development finance	no support from development finance
Investment II – large scale, cereal production,	Investment V – large scale, cereal production,
company owned by capital fund supplied with	private company with support from foreign
finance from pension schemes and	capital fund
development finance	
Investment III – medium scale, coffee farm,	Investment VI – medium scale farm run in
private company with support from	partnership between Ugandan and Danish
development finance	owners. Primary crop is maize, and operations
	are financed primarily from own funds

### Table 1. The six AIDA investments\*

\* For the purposes of this paper, the names of the investments are not relevant and have therefore been omitted.

The six research locations (Map 1) were selected and delimited with reference to these AIDA investments.





Taking the location of each of these six investments as the centre, six research locations have been delimited as the 'functional territories' for these investments, taking into account functions such as labour recruitment, marketing of produce, the area for potential investment expansion and the areas from which water is drawn

and to which it is discharged. Input guiding the delimitation of these research locations was obtained through field interviews conducted by the AIDA Tanzania and Uganda research teams respectively during October and November 2017. Although some of the selected AIDA investments, as well as clusters of other agricultural investments located around them, were found to attract labour from other parts of the country, in most cases such labourers were found to establish themselves around investments or clusters of them, at least during part of the year. In addition, other effects, e.g. with respect to water use, knowledge-sharing etc., appeared to congregate relatively close to the investment centre, within a radius of approximately 25 km. This coincides with the finding of Deininger and Xia regarding Mozambique, namely that "most spillover effects are quite localised, i.e. establishment of large farms seems to have no impact on smallholders if it is farther than 50 km" (Deininger and Xia, 2016: 237).

On this basis, the six research locations were delimited through the following three step-procedure:

- 1. First, for each investment a 25 km radius circle was drawn around the AIDA investment centre.<sup>1</sup> Only the part of the area falling within the district where the AIDA investment centre was located was considered in making the following steps.
- 2. Second, to align the research locations with existing sub-district administrative boundaries and at the same time ensure that the research location covers areas located both within and outside the 25 km radius from the AIDA investment centre, each of the six research locations was delimited by the boundaries of the wards (in Tanzania) or parishes (in Uganda) that are fully or partly contained within the 25 km radius.<sup>2</sup> Wards and parishes which only have a minor part of their area within the 25 km radius from the investment centre were not included as part of the research location.
- 3. Third, for urban wards and parishes falling fully or partly within the 25 km radius, interviews with key informants, e.g. at bus stops, market places etc., were conducted to establish whether people from the urban ward or parish would occasionally orient themselves towards the area around selected investment when searching for employment, goods etc. If that turned out not to be the case, the urban ward or parish would be omitted from the research location.

The resulting research locations are shown as the coloured areas in Maps 2-7, while Table 2 lists the wards and parishes which form part of the six research locations. As shown in the maps, the research locations vary greatly in terms of geographical

<sup>1</sup> Due to a combination of low population density in the vicinity of Investment V and a GPS error in fixing the exact location of the investment centre, a 50 km radius circle (rather than a 25 km radius circle) was drawn around Investment V as the first step towards delimiting this research location.

<sup>2</sup> In the case of Investment IV, which is located in a densely populated area where the parishes (known as the LC2 level in Uganda) are geographically small, delimitation of the research location was done with reference to the sub-county level (known as the LC3 level) instead.

size, just as the population density varies greatly both within and among the research locations (see also Table 3).

In addition to agricultural investments from Denmark, each research location comprises a number of other foreign agricultural investments with equal potential to affect local dynamics. For each research location, these investments were identified through interviews with key informants, particularly district authorities and the owners and/or managers of foreign agricultural investments. The locations of these investments are also indicated in the research location maps (Maps 2-7).













Map 5. Kanungu research location, Uganda







Map 7. Nakasongola research location, Uganda



#### Table 2. Delimitation of research locations

Wards (Tanzania) and parishes (Uganda) (sub-counties in the case of Kyeshero, Kanungu) to be included in research locations

Wards and parishes that are fully or partly within a 25 km radius from the investment location (50 km radius in the case of Nwoya research location)

Wards and parishes marked in bold indicate where the investment is located; wards and parishes marked in light grey indicate those that were not included in the research location due to their not forming part of the potential functional territory of the selected farm

		Tanzan	ia		
Karatu resea	arch location	Iringa resear	ch location	ation Njombe research location	
Fully within 25 km radius	Partly within 25 km radius	Fully within 25 km radius	Partly within 25 km radius	Fully within 25 km radius	Partly within 25 km radius
<ul> <li>Ganako</li> <li>Qurus</li> <li>Oldean</li> <li>Rhotia</li> <li>Daa</li> </ul>	<ul> <li>Endamarariek</li> <li>Mbulumbulu</li> </ul>	<ul> <li><i>Ifunda</i></li> <li>Mgama</li> <li>Lumuli</li> </ul>	<ul><li>Mseke</li><li>Wasa</li><li>Maboga</li></ul>	<ul> <li>Uwemba</li> <li>Utalingolo</li> <li>Lugenge</li> <li>Luponde</li> <li>Yakobi</li> </ul>	<ul> <li>Njombe mjini</li> <li>Mjimwema</li> <li>Ramadhani</li> <li>Kifanya</li> <li>Iwungilo</li> <li>Makowo</li> <li>Matola</li> </ul>
		Ugand	la		
Kanungu rese	earch location	Nwoya reseat	ch location	Nakasongola re	esearch location
Fully within 25 km radius	Partly within 25 km radius	Fully within 50 km radius	Partly within 50 km radius	Fully within 25 km radius	Partly within 25 km radius
<ul><li><i>Kayonza</i></li><li>Mpungu</li></ul>	<ul><li>Kihiihi</li><li>Nyakinoni</li></ul>	<ul><li><i>Panokrach</i></li><li>Alero Kal</li></ul>	<ul><li>Pangur</li><li>Amar</li></ul>	<ul><li>Katuugo</li><li>Kyabutaika</li></ul>	<ul><li>Namaasa</li><li>Kageri</li></ul>

<ul> <li>Kinaaba</li> </ul>	<ul> <li>Nyamirama</li> </ul>	<ul> <li>Bwobonam</li> <li>Pa</li> </ul>	bit • Kyeyindula	<ul> <li>Kamunina</li> </ul>
<ul> <li>Kanungu TC</li> </ul>	<ul> <li>Rutenga</li> </ul>	<ul> <li>Panyabono</li> </ul>	<ul> <li>Kabaale</li> </ul>	<ul> <li>Wampiti</li> </ul>
<ul> <li>Rugyeyo</li> </ul>	0	<ul> <li>Akago</li> </ul>	<ul> <li>Kakooge TC</li> </ul>	<ul> <li>Kyambogo</li> </ul>
<ul> <li>Kirima</li> </ul>		<ul> <li>Ceke</li> </ul>	Central	
<ul> <li>Butogota</li> </ul>		<ul> <li>Ogom</li> </ul>	<ul> <li>Kakooge TC</li> </ul>	
<ul> <li>Kanyantorogo</li> </ul>		<ul> <li>Paibwor</li> </ul>	North	
<ul> <li>Kambuga</li> </ul>		<ul> <li>Pabali</li> </ul>	<ul> <li>KIBIRA</li> </ul>	
<ul> <li>Kambuga TC</li> </ul>		<ul> <li>Todora</li> </ul>	<ul> <li>Nakasongola</li> </ul>	
<ul> <li>Katete</li> </ul>		<ul> <li>Ywaya</li> </ul>	TC Central	
<ul> <li>Kihihi TC</li> </ul>		<ul> <li>Coorom</li> </ul>	<ul> <li>Nakasongola</li> </ul>	
<ul> <li>Nyanga</li> </ul>		<ul> <li>Labyei</li> </ul>	TC West	
		<ul> <li>Latoro</li> </ul>	<ul> <li>Nakasongola</li> </ul>	
		<ul> <li>Paromo</li> </ul>	TC East	
		<ul> <li>Pawatomero</li> </ul>	<ul> <li>Kiwongoire</li> </ul>	
		<ul> <li>Patira</li> </ul>	<ul> <li>Wabigalo</li> </ul>	
		<ul> <li>Pangora</li> </ul>	<ul> <li>Sikye</li> </ul>	
		5	<ul> <li>Ssasira</li> </ul>	
			<ul> <li>Kyankonwa</li> </ul>	

#### Table 3. Research location area and population

Area (km2) and population density (persons/km2) by research location

Research location	Area (km <sup>2</sup> )	Population density (persons/km <sup>2</sup> )			
		Average Minimum per		Maximum per	
		ward/parish/sub-		ward/parish/sub-	
			county	county	
Karatu	1,121.6	106.2	43.9	237.0	
Iringa	2,073.7	34.6	20.2	66.9	
Njombe	2,670.6	25.5	10.6	54.5	
Kanungu	1,328.6	187.0	57.0	975.7	
Nwoya	3,632.7	29.5	7.0	199.4	
Nakasongola	1,204.4	48.0	12.8	1,028.9	

Sources (population data): United Republic of Tanzania (2013) and Uganda Bureau of Statistics (2016).

### SAMPLING APPROACH FOR DRAWING THE SIX INDEPENDENT SAMPLES

A number of the development outcomes that the AIDA questionnaire survey is designed to trace, such as employment practices and preferences, are specific to the individual, while others, such as access to water for domestic purposes, to a larger extent relate to the household. Therefore, the AIDA questionnaire is designed so that certain sections of the interview seek to solicit information specifically with respect to the individual respondent, while other sections solicited information with respect to the household as a whole. For sampling purposes, this implies that the sampling process was designed to draw random samples of individuals from each of the six research locations.

Due to the limited availability of data, e.g. with respect to population size, composition and the names of village and sub-village inhabitants,<sup>3</sup> to allow direct sampling, sampling frames had to be established through field visits as part of the

<sup>3</sup> Much of such data has to be produced through visits to the relevant places.

sampling process. Thus, the six samples – one sample for each research location – were drawn through a multi-stage random sampling procedure.

As the first step, following the process described below, a sample of twenty villages was drawn for each research location (stage 1). For each of the selected communities, sampling frames – i.e. comprehensive lists of individuals – were developed to enable the random sampling of individuals (stage 2) as respondents for the questionnaire survey.

### Stage 1: geographically stratified, proportional random selection of villages

Based on the results of previous questionnaire survey research work,<sup>4</sup> the variation among households belonging to different villages with respect to characteristics such as poverty level and origin or ethnicity can be assumed to be greater than the variation among households belonging to the same village. Thus, to ensure a geographical spread of each of the six samples and that each of the latter reflects the geographical distribution of the population, the number of villages to be selected from each ward or parish forming part of the research location was determined with reference to the proportion that the population of each ward or parish contributes to the total population of the research location (see Tables 4a-f). The desired number of villages from each ward or parish was subsequently randomly selected from a complete list of villages for each ward or parish.

Ward	Population	Ward	Number of	Adjusted
	(inhabitants)	population in	villages to be	number
		proportion (%)	selected per	villages to be
		of total research	ward	selected per
		location		ward
		population		
Ganako	15,481	13.0	2.6	3
Qurus	15,919	13.4	2.7	3
Oldean	6,870	5.8	1.2	1
Rhotia	24,268	20.4	4.1	4
Daa	9,868	8.3	1.7	2
Endamarariek	24,996	21.0	4.2	4
Mbulumbulu	21,764	18.3	3.7	3
Total	119,166	100.0	20.0	20

### Table 4a-f. Sample composition

Number of villages to be selected from each ward or parish by research location

b. Iringa research location				
Ward	Population (inhabitants)	Ward population in proportion (%) of total research	Number of villages to be selected per ward	Adjusted number villages to be selected per ward

4 Analysis conducted on the datasets reported in Ravnborg et al. 2004 and 2013.

5 In Ganako ward, it was only possible to select two villages. Thus, in order to have a total sample of two villages from the Ganako research location, four villages were randomly selected from Mbulumbulu ward.

		location		
		population		
Ifunda	12,199	17.0	3.4	3
Mgama	12,561	17.5	3.5	4
Lumuli	7,852	10.9	2.2	2
Mseke	15,868	22.1	4.4	4
Wasa	10,595	14.8	3.0	3
Maboga	12,642	17.6	3.5	4
Total	71,717	100.0	20.0	20

c. Njombe research location				
Ward	Population (inhabitants)	Ward population in proportion (%) of total research location population	Number of villages to be selected per ward	Adjusted number villages to be selected per ward
Uwemba	8,900	13.0	2.6	3
Utalingolo	4,536	6.6	1.3	1
Lugenge	5,843	8.6	1.7	2
Kifanya	9,011	13.2	2.6	3
Luponde	9,372	13.7	2.7	3
Yakobi	5,660	8.3	1.7	2
Iwungilo	8,419	12.3	2.5	2
Makowo	4,213	6.2	1.2	1
Matola	12,262	18.0	3.6	3
Total	68,216	100.0	20.0	20

### d. Kanungu research location

Sub-county (LC3)	Population (inhabitants)	Sub-county population in proportion (%) of total research location population	Number of villages to be selected per sub-county	Adjusted number villages to be selected per sub-county
Kayonza	27,588	11.1	2.2	2
Mpungu	11,946	4.8	1.0	1
Kinaaba	8,478	3.4	0.7	1
Kanungu TC	15,056	6.1	1.2	1
Rugyeyo	20,731	8.4	1.7	2
Kirima	17,655	7.1	1.4	1
Butogota	10,315	4.2	0.8	1
Kanyantorogo	18,949	7.6	1.5	2
Kambuga & Kambuga TC	27,813	11.2	2.2	2
Katete	7,276	2.9	0.6	1
Kihihi TC	19,812	8.0	1.6	2
Nyanga	7,428	3.0	0.6	1
Kihiihi	17,967	7.2	1.4	1
Nyakinoni	8,743	3.5	0.7	1
Nyamirama	14,544	5.9	1.2	1
Rutenga	13,578	5.5	1.1	0
Total	247,879	100.0	20.0	20

e. Nwoya research location <sup>6</sup>				
Parish (LC2)	Population	Parish	Number of	Adjusted
	(inhabitants)	population in	villages to be	number
		proportion (%)	selected per	villages to be
		of total research	parish	selected per
		location		parish
		population		
Panokrach	9,244	9.4	1.9	2
Alero Kal	8,562	8.7	1.7	2
Bwobonam	4,971	5.0	1.0	1
Panyabono	5,366	5.4	1.1	1
Akago, Pabali & Labyei	6,364	6.5	1.3	1
Ceke	5,069	5.1	1.0	1
Ogom	4,349	4.4	0.9	1
Paibwor	7,798	7.9	1.6	2
Pangur	5,868	6.0	1.2	1
Todora	4,312	4.4	0.9	1
Ywaya	3,957	4.0	0.8	1
Amar	4,119	4.2	0.8	1
Coorom	4,941	5.0	1.0	1
Latoro	12,839	13.0	2.6	2
Paromo	3,791	3.8	0.8	1
Pawatomero	7,013	7.1	1.4	1
Total	98,563	100.0	20.0	20

### f. Nakasongola research location

Parish (LC2)	Population (inhabitants)	Parish population in	Number of villages to be	Adjusted number
		proportion (%)	selected per	villages to be
		of total research	parish	selected per
		location		parish
		population		
Kyabutayika	2,575	4.5	0.89	1
Katuugo	5,200	9.0	1.80	1
Kyambogo	5,557	9.6	1.92	2
Kyankonwa	2,993	5.2	1.04	1
Kyeyindula	3,575	6.2	1.24	1
Kabaale	1,990	3.4	0.69	1
Kakooge TC Central	2,457	4.3	0.85	1
Kakooge TC North	3,203	5.5	1.11	1
Kibira	1,326	2.3	0.46	0
Nakasongola TC Central	5,255	9.1	1.82	2
Nakasongola TC West	1,827	3.2	0.63	1
Kiwongoire	1,383	2.4	0.48	0
Wabigalo	2,526	4.4	0.87	1
Sikye	1,997	3.5	0.69	1
Namaasa	2,821	4.9	0.98	1
Nakasongola TC East	2,750	4.8	0.95	1
Kageri	2,721	4.7	0.94	1
Kamunina	2,028	3.5	0.70	1
Ssasira	2,355	4.1	0.82	1
Wampiti	3,221	5.6	1.12	1
Total	57,760	100.0	20.0	20

6 Unfortunately, Pabit parish was not included in the sampling process.

In order to protect the identity of the respondents, the names of the selected villages have been pseudonymised using a three-digit code:

Village code = Research location code (1-6) \* 100 + village number<sup>7</sup>

### Stage 2: Proportional random sampling of individuals

# Determining the number of respondents to be selected from each selected village

For each research location, the aim was to draw a sample of four hundred individuals.<sup>8</sup> As for the number of villages to be selected from each ward or parish, <sup>9</sup> the number of individuals to be selected from each ward or parish, and subsequently from each selected village, was determined in proportion to population size. First the number of individuals to be sampled from each ward or parish was determined based on the proportion that the population of the ward or parish constitutes of the total population of the research location. Where only one village was selected, the desired number of individuals were sampled from the selected village. Where two or more villages were selected from a ward or parish, the number of individuals to be selected from each selected village was established in proportion to the relative share that the population of the village constitute of the total population of the selected village strom that ward or parish.

### Box 1.

# Determining the number of individuals to be sampled in each village: a hypothetical example

Suppose that three villages are randomly sampled from ward X and that ward X is home to 8 percent of the total population of the research location. That implies that 8 percent of the 400 individuals to be sampled, i.e. 32 individuals, should be sampled from the three villages combined. Now, suppose that one of the three sampled villages, village V1, has a population twice the size of the remaining two sampled villages (villages V2 and V3). In order to ensure that all individuals 18 years of age or above in the three sampled villages have an equal chance of being selected, 16 individuals should be selected from village V1, while 8 individuals should be selected from each of villages V2 and V3.

<sup>&</sup>lt;sup>7</sup> The village number is a random number assigned to each village by AIDA staff. Unlike the village code, the village number does not carry information about the research location where the village is located.

<sup>&</sup>lt;sup>8</sup> In order to ensure a 95% probability sample with a 5% confidence interval, i.e. to be 95% confident that a given proportion, say of respondents taking casual employment, in the research location lies within 5% of our sample estimate, the required sample size was determined to be 384 individuals (Krejcie and Morgan, 1970; here quoted from Bernard, 1994). Thus, it was decided to sample 400 individuals in each of the six research locations.

<sup>&</sup>lt;sup>9</sup> Wards in the case of the three Tanzanian research locations, parishes in the case of two (Nwoya and Nakasongola) of the Ugandan research locations, and sub-counties in the case of the third research location (Kanungu) in Uganda.

Tables 5 a - b<sup>10</sup> give the number of respondents per village in the three research locations in Tanzania and Uganda respectively.

### Table 5a. Sample composition, Tanzania

Number of respondents to be selected per village by research location

Research location	Ward	Ward code	Ward popula- tion share of research location popula- tion (%)	Number of individ- uals to be sampled from ward	Village code	Village population share	Number of individuals to be selected from village
Karatu	Ganako	1	13.0	52	101 102	29.4 70.6	15 37
	Qurus	2	13.4	53	102 103 104 105	42.3 40.3 17.4	23 22 9
	Oldean	3	5.8	23	106	100.0	23
	Rhotia	4	20.4	81	107 108 109	30.4 21.9 22.9	25 18 19
	Daa	5	8.3	33	110 111 112	24.8 41.8 58.2	20 14 19
	Endamarariek	6	21.0	84	113 114 115 116	26.9 32.6 12.5 28.0	23 27 10 23
	Mbulumbulu	7	18.3	73	117 117 118 119 120	18.0 32.3 26.2 23.6	13 24 19 17
Iringa	Ifunda	1	17.0	68	201 202 203	66.0 19.9 14.1	44 14 10
	Mgama	2	17.5	70	204 205 206 207	17.6 38.5 20.9 23.1	12 27 15 16
	Lumuli	3	10.9	44	208	60.7 39.3	27
	Mseke	4	22.1	89	210 210 211 212 213	40.4 18.9 22.5 18.1	36 17 20 16
	Wasa	5	14.8	59	214 215 216	39.4 28.7 31.9	23 17 19
	Maboga	6	17.6	71	217 218 219	34.6 30.6 17.5	24 22 12

 $^{10}$  In accordance with GDPR regulations, the key table linking village code to village name is kept separately from all other data through encryption.

Njombe Uw Uta Luş Kif	vemba alingolo 2 genge 3	1 2 3	13.0 6.6	52	301 302 303	79.6 13.9 6.5	42 7 3
Uta Luş Kif	alingolo 2 genge 5	23	6.6		302 303	13.9 6.5	7
Uta Luş Kif	alingolo 2 genge 3	2	6.6		303	6.5	3
Uta Lug Kif	alingolo 2 genge 3	2 3	6.6	07			5
Luş	genge 3	3		27	304	100.0	27
Kif			8.6	34	305	57.0	19
Kif					306	43.0	15
	anya 4	4	13.2	53	307	62.7	33
					308	19.5	10
Ι					309	17.9	9
Luj	ponde !	5	13.7	55	310	27.9	15
					311	37.0	20
					312	35.1	19
Yal	kobi (	6	8.3	33	313	64.0	21
					314	36.0	12
Iwi	ungilo	7	12.3	49	315	43.7	22
	-				316	56.3	28
Ma	ikowo 8	8	6.2	25	317	100.0	25
Ma	itola	9	18.0	72	318	32.8	24
					319	23.0	17

**Table 5b. Sample composition, Uganda**Number of respondents to be selected per village by research location

Research location	Sub-county (LC3)	Parish (LC2)	Pa- rish	Parish popu-	Num- ber of	Villa- ge	Village popu-	Num- ber of
			coue	share	duale	coue	share	duale
				of re-	to be		of re-	to be
				search	sam-		spon-	select-
				loca-	pled		dents	ed
				tion	from		to be	from
				popu-	parish		select-	village
				lation	r		ed	
				(%)			from	
				( )			parish	
Kanungu	Kayonza	Karangara	1	3.8	15	401	100.0	15
C C	Kayonza	Mukono	2	8.0	32	402	100.0	32
	Mpungu	Buremba	3	5.1	20	403	100.0	20
	Kinaaba	Mukirwa	4	3.6	14	404	100.0	14
	Kanungu	Northern	5	6.4	26	405	100.0	26
	TC							
	Rugyeyo	Kashojwa	6	6.2	25	406	100.0	25
	Rugyeyo	Nyaruram- bi	7	2.7	11	407	100.0	11
	Kirima	Rutugunda	8	7.5	30	408	100.0	30
	Butogota	Western	9	4.4	18	409	100.0	18
	Kanyan- torogo	Burema	10	3.4	14	410	100.0	14
	Kanyan- torogo	Nyamigoye	11	4.7	19	411	100.0	19
	Kambuga & Kambuga	Burongi	12	7.3	29	412	100.0	29
	IC							

	Kambuga & Kambuga TC	Nyaru- gunda	13	4.5	18	413	100.0	18
	Katete	Kishuro	14	3.1	12	414	100.0	12
	Kihihi TC	Nyaka- tunguru	15	5.3	21	415	100.0	21
	Kihihi TC	Rwanga	16	3.1	13	416	100.0	13
	Nyanga	Nkunda	17	3.2	13	417	100.0	13
	Nvakinoni	Nvakinoni	18	3.7	15	418	100.0	15
	Kihihi	Rushoroza	19	7.7	31	419	100.0	31
	Nyami- rama	Mashaku	20	6.2	25	420	100.0	25
Nwoya	Alero	Panokrach	1	9.4	39	501 502	46.6 53.4	18 21
		Alero Kal	2	8.7	36	503	42.4	15
		111010 1111	-	011	00	504	57.6	21
		Bwobonam	3	5.0	21	505	100.0	21
		Panyabono	4	5.4	23	506	100.0	23
		Paibwor	5	7.9	33	507	68.6	23
		i dio wor	U	7.5	00	508	31.4	10
		Pangur	6	6.0	25	509	100.0	25
	Anaka	Ceke	7	5.1	21	510	100.0	21
	Town	Ogom	8	4.4	18	511	100.0	18
	Council	Labvei	9	2.5	10	512	100.0	10
	Anaka	Todora	10	4 4	18	513	100.0	18
	Pavira	Ywava	11	4.0	17	514	100.0	17
	Koch-	Amar	12	4.2	17	515	100.0	17
	Goma	Coorom	13	5.0	21	516	100.0	21
	Purongo	Latoro	14	13.0	54	517	94.4	51
	i urongo	Lucoro		10.0	01	518	5.7	3
		Paromo	15	3.9	16	519	100.0	16
		Pawa-	16	7.1	30	520	100.0	30
		tomero						
Naka-	Kakooge	Katuugo	1	9.0	38	601	100.0	38
songola	0	Kyabuta-	2	4.6	19	602	100.0	19
		yika						
		Kyambogo	3	9.6	40	603	46.7	19
		, ,				604	53.3	22
		Kyankon- wa	4	5.2	22	605	100.0	22
		Kyeyindula	5	6.2	26	606	100.0	26
	Kakooge	Kabaale	6	3.5	14	607	100.0	14
	TC	Kakooge TC Central	7	4.3	18	608	100.0	18
		Kakooge TC North	8	5.6	23	609	100.0	23
	Nabiswera	Namaasa	9	4.9	20	610	100.0	20
	Nakason-	Nakason-	10	9.1	38	611	52.2	20
	gola TC	gola TC						
	0	Central				612	47.8	18
		Nakason-	11	4.8	20	613	100.0	20
		gola TC East						
		Nakason- gola TC West	12	3.2	13	614	100.0	13
			•			·····	••••	

Wabi- nyonyi	Kageri	13	4.7	20	615	100.0	20
5 5	Kamunina	14	3.5	15	616	100.0	15
	Ssasira	15	4.1	17	617	100.0	17
	Wabigalo	16	4.4	18	618	100.0	18
	Wampiti	17	5.6	23	619	100.0	23
	Sikye	18	3.5	15	620	100.0	15

### Establishing the village-level sampling frame

In order to draw a random sample of individuals eighteen years or above, a comprehensive list of individuals living in the selected villages was developed. However, for large villages, i.e. villages with more than approximately seventy households, the task of developing such lists is not only time-consuming, it also entails a greater risk that some individuals may be omitted, such as the members of newly established households or newcomers to the village, than in smaller villages. Therefore, in larger villages, key informants were asked to list all sub-villages or neighbourhoods, after which half of these sub-villages were randomly selected and comprehensive lists of individuals were developed for them.

In Tanzania, the village leadership in many villages keeps a population register which includes the names of the household heads, with year of birth and gender, and lists other members of the household under the categories of husband or wife, children, relatives, friends, year of birth and gender. Labour migrants are listed under the category of 'friends'. Thus, in order to develop the sampling frame in villages where such population registers existed, the village leadership was asked to produce an updated copy. In villages without an existing population register or where the register was very old, key informants were asked to draw up a new one on the sub-community level in order to ensure accuracy.

In Uganda, village-level population registries were found to exist in only a few of the sampled villages. Thus, in Uganda, in all the sampled villages (LC1s) the chairperson was contacted, provided with a notebook and asked to draw up a population register of all village inhabitants eighteen years or above by name, age and gender, and clustered by household.

For each sample village, each entry in the population registry was numbered. Based on a list of random numbers generated for each village,<sup>11</sup> the desired number of respondents (see Tables 2 a & b, above) and a small buffer of possible replacements was selected, the latter to be used if a selected respondent declined the invitation to participate in an interview or had left the village on a long-term basis. As a consequence of this sampling procedure, the sample may contain two or more respondents belonging to the same household.

As for the selected village, the identity of the selected individuals was pseudonymised, using a six-digit code where the first three digits correspond to the

<sup>&</sup>lt;sup>11</sup> The list was generated using excel (=randbetween(1-n), where 'n' corresponds to the total number of individuals eighteen years or older contained in the population registry for the particular village.

village code, described above, the other three to the number assigned to each individual in the population registry developed for each selected village, viz.:

```
Respondent code = Village code * 1000 + individual number
```

In accordance with GDPR regulations, the key table linking the respondent code to the individual and village of residence is kept separately from all other data through encryption.

### ADMINISTERING THE QUESTIONNAIRE SURVEY

### Administering the questionnaire survey in Tanzania

In Tanzania, the questionnaire survey was administered using a printed version of the questionnaire, translated into Kiswahili. A core team of three persons, consisting of the AIDA PhD candidate supported by two research assistants, was responsible for supervising administration of the questionnaire survey in the three research locations, including the identification and training of local enumerators.

In each research location, twenty enumerators were recruited, one from each of the sampled villages, to undertake the interviews corresponding to their respective villages and to record the data within a period of ten to twelve days. Prior to undertaking the interviews, a joint two-day training session was held for the local enumerators in each of the research locations in order to introduce the purpose of the survey, create understanding and familiarity with the questionnaire format, and instruct the enumerators in ethical issues relating to consent and data protection. Also, senior AIDA research team members participated in conducting these training sessions.

All enumerators were requested to provide their consent to observe to only interview respondents, who had provided their free, prior and informed consent. Thus, enumerators were instructed to inform potential respondents about the purpose of the interview, that data would be pseudonymised, that data would be handled confidentially, that raw data would never be shared publicly except in the form of anonymised aggregate data, and that as respondents they would have the right to know what data the AIDA team stores on the basis of the interview. Enumerators were also requested to record the consent or lack of consent given by potential respondents in a specific consent form. In cases where a sampled respondent declined to participate in the survey or where a sampled person would be out of the village for a prolonged period of time during the survey period, enumerators were instructed to contact the AIDA research team members, who would then provide the enumerator with the name of a person, randomly selected, as a replacement. Table 6 below reports on the number of consents, declines and replacements per research location.

During the interview period, the AIDA core team was available to support the enumerators and provide quality control and feedback on completed interviews

and questionnaire formats the enumerators handed in. In cases where inconsistencies were encountered in the completed questionnaires or where sections had been skipped unintentionally, enumerators were asked to go back to respondents to cross-check or fill in the missing sections.

For each questionnaire interview, enumerators were asked to record the start time, end time and duration of the interview. However, due to a confusion between the Kiswahili and English languages in how they record time,<sup>12</sup> even within the same interview, legible time records are only available for half (n=805 respondents)<sup>13</sup> of the interviews.<sup>14</sup> On this basis, the average duration of the interview was one hour and 39 minutes<sup>15</sup> with no significant difference between the research locations.<sup>16</sup>

Upon completion, questionnaires were handed over to a team of data-entry clerks recruited from Sokoine University of Agriculture. These clerks were trained for two days in how to enter data using the password-protected template developed by the AIDA project, and like the enumerators they were requested to provide their consent with respect to handling the data they obtained confidentially and hand it over to the AIDA-Tanzania team staff.

### Administering the questionnaire survey in Uganda

In Uganda, the questionnaire was transferred into CAPI (Computer Assisted Personal Interviewing) software, and interviewing and data recording took place using tablets. A total of forty enumerators were recruited from Makerere University, Kampala. A two-day training workshop was conducted in Kampala in order for the enumerators to acquaint themselves with the objectives and content of the questionnaire survey. They were also instructed in using the CAPI software and in the ethical issues associated with the survey, including the issue of consent. In addition, an in-field training session was held in Luwero district (not part of any of the research locations) for the enumerators to familiarise themselves further with the questionnaire and the CAPI software and to clarify any uncertain issues. As in Tanzania, the enumerators were requested to provide their consent to observe to only interview respondents from whom they had obtained free, prior and informed consent. In Uganda, all the respondents who were sampled and approached agreed to participate in the survey. Table 6 below reports on the number of consents and replacements for each of the three Ugandan research locations.

Enumerators were divided into three teams, one per research location. The teams were sent to the research locations in a cascading sequence to allow the two AIDA PhD candidates to accompany the teams during their initial days interviewing in each of the research locations and thus clarify possible issues of doubt or

<sup>&</sup>lt;sup>12</sup> In Kiswahili, 1 o'clock means the first hour of the day, i.e. 7 am English language time.

<sup>&</sup>lt;sup>13</sup> In these cases, the duration based on the start and end time as recorded in the questionnaire corresponds with the duration stated in the questionnaire.

<sup>&</sup>lt;sup>14</sup> These are distributed with 235 respondents in the Karatu research location, 314 respondents in the Njombe research location and 256 respondents in the Iringa research location.

<sup>&</sup>lt;sup>15</sup> Overall, the median (50 percentile) was 90 minutes, while the 5 and 95 percentiles were 50 minutes and 174 minutes, respectively.

<sup>&</sup>lt;sup>16</sup> *p*=0.083; significance of F, ANOVA.

misunderstandings in the field. In addition to the two AIDA PhD candidates, each team also had a number of appointed supervisors who, on a daily basis, checked through a randomly selected subset of the data collected by each enumerator to check for consistency and errors. Moreover, after three days of data collection, each team took a day off, and the enumerators with their supervisors went through their datasets to check for consistency in the way the questions were asked and answered. Enumerators exchanged experiences and discussed the observations they had made to guide how data would be collected going forward. At the end of the data collection exercise, all the enumerators gathered at Makerere University for data cleansing.

Interviewing took place between March 27 and April 18, 2019. In the Nakasongola research location, two of the selected villages were found to have very small or no sedentary populations, and neither the pre-selected respondents nor the replacements were available for interviews, as they had left the village in search of grazing for their animals elsewhere. Also, in the Nwoya research location, a number of the pre-selected respondents had temporarily left their villages to work their fields in other locations (see Table 6).

In Uganda, legible data on interview start and end times, and thus on the duration of the interview, were recorded for a little less than half the respondents (n=508 respondents), unfortunately not evenly distributed among the three research locations.<sup>17</sup> On this basis, the average duration of the interview was 1 hour and 2 minutes, ranging from 57 minutes in the Kanungu research location to 1 hour and 12 minutes in the Nakasongola research location and 1 hour and 24 minutes in the Navoya research location.<sup>18</sup>

Number of respo	Number of respondents sampled, declined consents and replacements during interviewing process								
Research	Ward	Number of	Number	Number	Final s	ample			
location	(Tanzanian	respondents	of	of					
	research	to be	sampled	replace-	Number	Number			
	locations)/sub-	sampled	respond-	ments**	of	of			
	county	(see Table 5	ents	included	respond-	respond-			
	(Ugandan	above)	declin-	in the	ents	ents			
	research		ing	sample	interview-	sampled			
	locations)		consent*		ed	as % of			
						sampl-			
						ing			
						target			
Karatu	Ganako	52	-	2	52	100.0			
	Qurus	53	-	13	54	101.9			
	Oldean	23	-	2	23	100.0			
	Rhotia	81	-	7	81	100.0			
	Daa	33	2	5	32	97.0			
	Endamarariek	84	-	12	82	97.6			
	Mbulumbulu	73	-	10	73	100.0			

#### Table 6. Final sample composition

<sup>17</sup> These are distributed with 340 respondents in the Kanungu research location, 147 respondents in the Nakasongola research location and only 21 respondents in the Nwoya research location.

<sup>&</sup>lt;sup>18</sup> Overall, the median was 54 minutes, with the 5 and 95 percentiles being 22 minutes and 117 minutes respectively.

Iringa	Ifunda	68	1	12	70	102.9
Ū.	Mgama	70	_	16	71	101.4
	Lumuli	44	_	20	44	100.0
	Mseke	89	_	14	89	100.0
	Wasa	59	2	18	60	101.7
	Maboga	71	_	15	71	100.0
Njombe	Uwemba	52	4	15	52	100.0
,	Utalingolo	27	-	8	27	100.0
	Lugenge	34	-	11	34	100.0
	Kifanya	53	2	11	48	90.6
	Luponde	55	-	5	54	98.2
	Yakobi	33	2	6	33	100.0
	Iwungilo	49	2	2	50	102.0
	Makowo	25	_	2	26	104.0
	Matola	72	_	3	73	101.4
Kanungu	Kavonza	47	_	9	62	131.9
0	Mpungu	20	_	1	13	65.0
	Kinaaba	14	_	2	12	85.7
	Kanungu TC	26	_	1	16	61.5
	Rugvevo	36	_	7	43	119.4
	Kirima	30	_	4	21	70.0
	Butogota	18	_	1	14	77.8
	Kanvantorogo	33	_	5	47	142.4
	Kambuga &	47	_	12	57	121.3
	Kambuga TC					
	Katete	12	_	3	12	100.0
	Kihihi TC	34	_	11	47	138.2
	Nvanga	13	_	2	8	61.5
	Nvakinoni	15	_	2	8	53.3
	Kihihi	31	_	5	27	87.1
	Nvamirama	25	_	1	12	48.0
Nwova	Alero &	177	_	48	171	96.6
	Lungulu					
	Anaka Town	50	_	14	46	92.0
	Council					
	Anaka Pavira	35	_	10	35	100.0
	Koch-Goma	38	_	8	37	97.4
	Purongo	100	_	53	99	99.0
Nakasongola	Kakooge	145	_	12	143	98.6
i takusongolu	Kakooge TC	56	_	17	57	101.8
	Nabiswera	20	_	_	16	80.0
	Nakasongola TC	71	_	36	72	101 4
	Wahinyonyi	108	_	16	106	98.1
	, abiliyofiyi	100	—	10	100	70.1

\* In Uganda, none of the sampled respondents declined to participate in the survey.

\*\* Number of replacements includes replacements of individuals who declined to participate in the survey, as well as of individuals who were out of the village for prolonged periods during the survey period.

### **Data cleansing**

Data cleansing, i.e. the identification and correction of incorrectly entered or inconsistent data, was undertaken in SPSS, using simple frequencies and

formulating logical conditions. In Tanzania, inconsistencies in the database were compared to the paper records of the interviews and were corrected accordingly.

### SAMPLE CHARACTERISATION

The six independent samples consist of approximately 400 respondents, ranging from 389 respondents in the Nwoya research location to 405 respondents in the Iringa research location.

### Demographic characterisation of samples

### Respondent's position within the household

Around half of the respondents (52.0 per cent) in the three Tanzanian samples and 60.6 per cent of those in the three Ugandan samples combined declared they were the heads of their households (Table 7).

According to the UN,<sup>19</sup> the average household size in Tanzania is 4.9 persons, including 2.9 children under the age of 15 years. This implies that an average household would have a little less than two persons 18 years or above, of whom one would be expected to be the household head. In Uganda the average household size is 4.7 persons, including 3.3 children under the age of the 15, implying that an average household would have a little less than 1.4 persons 18 years or above, of whom one would be expected to be the household head. Viewed on this basis, it seems that our sampling strategy of ensuring that all community members 18 years and above have an equal chance of being sampled, irrespective of their position within the household, has been successful.

Research location	Re	Relation to head of household					
	Head of	Spouse of	Son or	Other	respondents		
	household	head of	daughter				
		household	of				
			household				
			head				
Karatu (n=397	47.1	29.0	19.6	4.3	100.0		
respondents)							
Iringa (n=405	46.2	22.2	27.4	4.2	100.0		
respondents)							
Njombe (n=401	62.6	21.2	13.2	3.0	100.0		
respondents)							
Tanzania	52.0	24.1	20.1	3.8	100.0		
research							

Table 7. Respondent's	position within th	e household b	v research location

Percent respondents per relation to head of household, by research location

19

https://www.un.org/en/development/desa/population/publications/pdf/ageing/household\_size\_and\_comp osition\_around\_the\_world\_2017\_data\_booklet.pdf, pp. 12-13.

locations***					
(N=1 203					
(11-1,203					
respondents)					
Kanungu	49.6	30.8	17.0	2.5	100.0
(n=399					
respondents)	<b>—</b> —				
Nwoya (n=389	63.8	27.2	7.7	1.3	100.0
respondents)	<b>—</b>				
Nakasongola	68.8	17.8	11.2	2.3	100.0
(n=394					
respondents)					
Uganda	60.6	25.3	12.0	2.0	100.0
research					
locations***					
(N=1,182					
respondents)					

\*\*\* Significant correlation between the research locations with respect to the distribution of respondents according to relation of respondent to head of household; the significance of Pearson Chi-Square is < 0.001.

Overall, for the three Tanzanian samples, 30.1 per cent of respondents declaring they were household heads were women, ranging from 19.8 per cent in the Karatu sample to 36.9 per cent in the Iringa sample (not shown in Table). Similar proportions of female household heads were found for the Ugandan locations, where overall 30.9 per cent of respondents declaring they were heads of household were women, ranging from 26.3 per cent in the Kanungu sample to 36.8 per cent in the Nwoya sample.

### **Gender composition**

Overall, there is an almost equal distribution between men and women among the respondents in the Tanzanian and the Ugandan samples respectively (Table 8). However, there are significant differences, particularly among the three Ugandan research locations, with 54.0 per cent of the adult population in the Nwoya location being women, while in the Nakasongola location, which in recent decades has been a recipient of population from other parts of Uganda, 54.8 per cent of the population are men.

#### Table 8. Gender of respondents by research location

Percent respondents per gender, by research location

Research location	Gender of	Gender of respondent		
	Male	Female		
Karatu (n=397	54.7	45.3	100.0	
respondents)				
Iringa (n=405	47.9	52.1	100.0	
respondents				
Njombe (n=401	50.4	49.6	100.0	
respondents)				
Tanzania research	51.0	49.0	100.0	
locations <sup>ns</sup>				
(N=1,203				
respondents)				

Kanungu (n=399	47.9	52.1	100.0	
respondents)				
Nwoya (n=389	46.0	54.0	100.0	
respondents)				
Nakasongola	54.8	45.2	100.0	
(n=394				
respondents)				
Uganda research	49.6	50.4	100.0	
locations* (N=1,182				
respondents)				

<sup>ns</sup> No significant correlation between the research locations with respect to the distribution of respondents according to gender (p=0.154; Pearson Chi-Square test).

\* Significant correlation between the research locations with respect to the distribution of respondents according to gender at the 0.05 level (*p*=0.034; Pearson Chi-Square test).

### Age of respondent

With an average age of 38.7 years, the average respondent in the three Ugandan samples was slightly younger than the average respondent in the three Tanzanian samples, at 41.4 years (Table 9).<sup>20</sup> Thus, for the Tanzanian samples, 13.8 per cent of the respondents in the Njombe research location were below 25 years of age, while this was the case for 20.8 per cent of the respondents in the Iringa research location. In Uganda, the Kanungu research location had the oldest population, with 18.5 per cent of the respondents being below 25 years of age, compared to the Nakasongola research location, where 24.4 per cent of the respondents were below 25 years of age.

#### Table 9. Age of respondents by research location

Percent respondents per age category, by research location

Research location	Age of respondent All				
	<25	25-40	41-55	>55	respondents
	years	years	years	years	
Karatu (n=397	17.1	32.2	29.0	21.7	100.0
respondents)		_			
Iringa (n=403	20.8	31.5	23.3	24.3	100.0
respondents; 2					
respondents					
missing)					
Njombe (n=398	13.8	37.9	29.4	18.8	100.0
respondents; 3					
respondents					
missing)					
Tanzania	17.3	33.9	27.2	21.6	100.0
research					
locations*					
(N=1,198					
respondents; 5					
respondents					
missing)					
Kanungu (n=399	18.5	40.6	23.6	17.3	100.0
respondents)					

<sup>20</sup> In Tanzania, exact age was declared for three quarters of the respondents only (n=896).

Nwoya (n=388		22.9	40.2	21.9	14.9	100.0
respondents; 1						
respondent						
missing)						
Nakasongola		24.4	39.1	19.8	16.8	100.0
(n=394						
respondents)	_ = -					
Uganda research		21.9	40.0	21.8	16.3	100.0
locations <sup>ns</sup>						
(N=1,181						
respondents; 1						
respondent						
missing)						
<sup>ns</sup> No significant correl	ation between th	he research lo	ocations with	respect to the	distribution of	of respondents

according to age; the significance of Pearson Chi-Square is 0.485.

\* Significant correlation between the research locations with respect to the distribution of respondents according to age; the significance of Pearson Chi-Square is < 0.05.

### Characterisation of the parcels about which information was provided

Among the aspects of the potential development outcomes that are associated with foreign agricultural investments, some may be meaningfully examined with reference to an individual, while others are better examined through inquiries with reference to a particular parcel of land. This applies, for instance, to aspects that relate to security of land tenure, which may be perceived to be different for different parcels accessed by the same individual. Thus, part of the questionnaire is designed to solicit information with respect to specific parcels. All respondents with access to land were asked to respond to the questions in this section of the questionnaire with reference to the parcel they consider their most important. In addition, respondents who are not heads of households were also asked to provide information with reference to the parcel which is considered the most important for the household as a whole.

In Tanzania, 977 of the 1,203 respondents reported having access to a piece of land that is used for agricultural purposes and for which they participate in decisionmaking. As part of the questionnaire interview, they provided information on the parcel they considered most important to them. Of these respondents, 589 declared they were the head of their households. In by far the majority of these cases (522), the respondent considered the parcel most important to him or herself also to be the parcel that was most important to the household as a whole, while in the remaining 67 cases, the respondent identified another parcel to be the most important parcel to the household as a whole. In addition to the 977 parcels reported to be most important to the respondent (irrespective of household position), information was obtained with respect to 294 parcels which were identified by the respondents as the most important to the respondent. Thus, for the three Tanzanian samples combined, information was obtained with respect to 1,271 parcels (see Table 10).

In Uganda, 1,108 of the 1,182 respondents reported having access to a piece of land which is used for agricultural purposes and for which they are fully or partly responsible. As part of the questionnaire interview, they provided information on

the parcel they considered most important to them. Of these respondents, 681 respondents declared they were the head of their households. In around half of these cases (347), the respondent considered the parcel most important to him- or herself also to be the parcel most important to the household as a whole. In the other half of the cases (334), the respondent identified another parcel to be the most important parcel to the household as a whole. In addition to the 1,108 parcels reported as most important to the respondent, information was obtained with respect to 595 parcels which were identified by the respondents as the most important parcel to the household as a whole but not the parcel identified as most important to the respondent. Thus, for the three Ugandan research locations combined, information was obtained with respect to 1,703 parcels (see Table 10).

Table 10. Parcels identified in the questionnaire survey as most important to the respondent and as most
important to the household as a whole, by research location

Number of parcels per category, by research location
--

	Number of	Number of	Number of	Number of	Total
	parcels	parcels	respondents	respondents	number of
	reported as	reported as	reporting on	who did not	parcels
	most	most	two parcels	report on	reported as
	important to	important to	– one as the	any parcel	most
	the	the	most		important
	respondent	household (if	important to		either to
		different	respondent,		respondent
		from the	and one as		or to
		parcel	the most		household
		reported as	important to		
		most	the		
		important to	household		
		the			
		respondent)			
Karatu (n=397	291	136	81	51	427
respondents)					
Iringa (n=405	349	66	26	16	415
respondents					
Njombe (n=401	337	92	31	3	429
respondents)					
Tanzania research	977	294	138	70	1,271
locations (N=1,203					
respondents)					
Kanungu (n=399	364	310	289	14	674
respondents)					
Nwoya (n=389	378	124	123	10	502
respondents)					
Nakasongola	366	161	155	22	527
(n=394					
respondents)					
Uganda research	1,108	595	567	46	1,703
locations (N=1,182					
respondents)					

#### **Spatial characterisation of samples**

In both Tanzania and Uganda, the three research locations were chosen for their differences in a number of respects, including their agro-ecological conditions and

agrarian structures. In effect, in both countries the research locations reveal different patterns with respect to population densities and the size of administrative units such as villages and wards or sub-counties (see Maps 2-7 and Table 3, above). This is also evident when examining the spatial distribution of the respondents in each of the six samples.

As described above, the six research locations were delimited using the chosen investment as the geographical centre of the research location. While administering the questionnaire survey, the geographical position of the respondent's residence was recorded. In Uganda, locational data were recorded for all respondents except one, while in Tanzania locational data were recorded for 795 out of the total of 1,203. Based on these data, Tables 11 and 12 and Figure 1 summarise the spatial distribution of the six samples relative to each research location centre.

# Table 11. Distribution of respondents according to distance between their residence and the research location centre, by research location

Research location	Distance from residence of respondent to All						All
			distance				
		< 1 km	1-4.9	5-9.9	10-24.9	=> 25	ranges
			km	km	km	km	
Karatu (n=228		_	29.8	19.3	50.9	-	100.0
respondents; 169							
cases missing)							
Iringa (n=312		-	_	15.7	69.2	15.1	100.0
respondents; 93							
cases missing)							
Njombe (n=255		_	20.4	5.1	70.2	4.3	100.0
respondents; 146							
cases missing)							
Tanzania research		-	15.1	13.3	64.3	7.3	100.0
locations*** (n=795							
respondents; 408							
cases missing)							
Kanungu (n=399		3.0	9.0	16.3	70.9	0.8	100.0
respondents)							
Nwoya (n=388		-	3.4	1.0	2.3	93.3	100.0
respondents; 1							
case missing)							
Nakasongola		0.3	4.3	0.0	88.6	6.9	100.0
(n=394							
respondents)	■						
Uganda research		1.1	5.6	5.8	54.3	33.2	100.0
locations***							
(n=1,181							
respondents; 1							
case missing)							

Percent respondents per distance range, by research location

\*\*\* Significant correlation between research location and distance at 0.001 level (Pearson Chi-Square)

With average distances from the respondent's home to the research location centre of 9.6 kilometres and 36.0 kilometres respectively, the Karatu and Nwoya samples represent the most spatially concentrated and spatially dispersed samples. While

significantly different from the remaining samples, the Iringa and the Nakasongola samples on the one hand, and the Njombe and the Kanungu samples on the other constitute two homogenous subsets in terms of spatial distribution relative to their respective research location centres. Overall, for the six samples, 75 percent of respondents' homes were located within a radius of 24.4 kilometres from the research location centre, ranging from within 14.4 kilometres for the Karatu sample to 41.9 kilometres for the Nwoya sample.

Research location	Distance between respondent home and research location centre				
	Median (50 percentile)	25 percentile	75	Average	Standard
Karatu (n=228 respondents; 169 cases missing) <sup>a</sup>	10.1	3.8	14.4	9.6	5.3
Iringa (n=312 respondents; 93 cases missing) <sup>b</sup>	17.4	12.4	23.7	17.8	6.4
Njombe (n=255 respondents; 146 cases missing) <sup>c</sup>	16.1	9.3	17.9	14.4	6.8
Tanzania research locations (n=795 respondents; 408 cases missing)	14.4	9.3	19.2	14.3	7.1
Kanungu (n=399 respondents) <sup>c</sup>	13.9	8.5	17.2	13.2	6.0
Nwoya (n=388 respondents; 1 case missing) <sup>a</sup>	35.9	32.1	41.9	36.0	9.9
Nakasongola (n=394 respondents) <sup>b</sup>	18.7	14.5	21.1	17.8	5.3
Uganda research locations (n=1,181 respondents; 1 case missing)	19.0	14.2	31.8	22.2	12.2

 Table 12. Distances between respondents' homes and research location centres (kilometres), with median, 25 and 75 percentiles, average and standard deviation, by research location

<sup>a</sup> Distribution of distances from respondents' homes to research location centres is significantly different from the distribution in the remaining research locations at 0.001 level (one-way analysis of variance, Scheffe test).

<sup>b</sup> The Iringa and the Nakasongola research locations constitute a homogenous sub-set with respect to the distribution of distances from respondents' homes to research location centres, each being significantly different from the remaining research locations at 0.001 level (one-way analysis of variance, Scheffe test). c The Njombe and the Kanungu research locations constitute a homogenous sub-set with respect to the distribution of distances from respondents' homes to research location centres, each being significantly different from the remaining research locations at 0.001 level (one-way analysis of variance, Scheffe test).

## Figure 1. Distance between respondents' homes and research location centres (kilometres), with median, 25 and 75 percentiles, average and standard deviation, by research location

*Box represents the 25 and 75 percentiles, line inside box represents median, whiskers represent 10 and 90 percentiles, and dots represent outliers.* 



### Density of foreign agricultural investments

Based on an inventory of the foreign agricultural farms present in the research locations, an index was generated of the exposure of each respondent to foreign agricultural farms. The index is based on the location of the foreign investment headquarters and the location the house of the respondent as proxies for the location of the foreign agricultural farms and the respondent, respectively. Because the inventory of foreign agricultural farms focused on farms located within the research location, this provides a more comprehensive indication of the proximity and thus the potential exposure to foreign agricultural farms for respondents who are closer to the research location centre than for those having their homes nearer the periphery of the research location. Therefore, it was decided only to compute the exposure index for a sub-set of respondents within each of the six samples. The subsets consist of respondents living within a distance from the research location centre corresponding to the median distance to the research location centre, i.e. half of the respondents for whom locational data was obtained living closest to the research location centre. For each respondent included in the sub-set, the distance to each of the identified foreign agricultural investments was tabulated. The following tables report on the exposure to foreign agricultural farms of the respondents included in the sub-set.

Table 13 shows the distribution of respondents according to different distance ranges from respondent home to foreign agricultural farms, while Table 14 shows the average number of foreign agricultural farms within a given distance range from the respondents.

Overall, the two tables show that in the Tanzanian research locations a higher proportion of respondents are exposed to foreign agricultural investments with respect to both the number and proximity of investments compared to respondents in the Ugandan research locations. Among the Ugandan research locations, Nwoya research location stands out as the location with a relatively low proportion of respondents being exposed to foreign agricultural investments in the immediate vicinity of their residence while being exposed to a relatively high number of foreign agricultural foreign investments located further away from their residence.

# Table 13. Presence of foreign agricultural investments in the vicinity of respondents' residence, by research location

Percent respondents<sup>21</sup> having one or more foreign agricultural investments within their vicinity, by distance range, per research location; only respondents having their residence within the median distance to the research location centre are included (N=997 respondents).

Research location	Distance from respondent's residence to foreign					o foreign	All
			agricu	ultural inv	estment		distance
		<1 km	1-4.9	5-9.9	10-24.9	=> 25	ranges
			km	km	km	km	
Karatu (n=119		20.2	69.7	100.0	100.0	93.3	100.0
respondents)							
Iringa (n=157		22.3	65.6	100.0	100.0	32.5	100.0
respondents)							
Njombe (n=130		3.8	66.2	63.1	83.1	100.0	100.0
respondents)							
Tanzania research		15.8	67.0	88.2	94.6	71.9	100.0
locations (n=406							
respondents)							
Kanungu (n=200		10.5	31.5	78.5	68.5	0.0	100.0
respondents)							
Nwoya (n=194		1.5	8.8	10.3	100.0	89.7	100.0
respondents)							
Nakasongola		3.0	21.3	71.1	100.0	56.3	100.0
(n=197 respondents)							
Uganda research		5.1	20.6	53.6	89.3	48.2	100.0
locations (n=591							
respondents)							

\*\*\* Significant correlation between research location and distance at 0.001 level (Pearson Chi-Square)

<sup>&</sup>lt;sup>21</sup> Because there is more than one foreign agricultural investment in each research location, a respondent's residence may be placed at different distances to the foreign agricultural investments. Therefore, percentages do not add up to 100.

# Table 14. Density of foreign agricultural farms according to distance range for respondents having their residences within the median distance of the research location centre, by research location

Average number of foreign agricultural investments by distance range from respondent's residence, by research location; only respondents having their residence within the median distance to the research location centre are included (N=997 respondents).

Research location		Distance from respondent's residence					All
		< 1 km <sup>a</sup>	1-4.9	5-9.9	10-24.9	=> 25	distance
			km <sup>b</sup>	km۲	km <sup>d</sup>	km <sup>e</sup>	ranges
Karatu (n=119		0.3	3.2	7.9	7.7	0.9	20.0
respondents)							
Iringa (n=157		0.2	1.3	3.9	5.2	0.4	11.0
respondents)							
Njombe (n=130		0.1	2.1	4.7	5.0	3.2	15.0
respondents)							
Tanzania research		0.2	2.1	5.3	5.9	1.4	14.9
locations (n=406							
respondents)							
Kanungu (n=200		0.1	0.7	1.8	2.4	0.0	5.0
respondents)				_			
Nwoya (n=194		0.0	0.2	0.3	4.8	5.7	11.0
respondents)	■						
Nakasongola (n=197		0.0	0.2	0.9	3.9	1.0	6.0
respondents)							
Uganda research		0.1	0.4	1.0	3.7	2.2	7.3
locations (n=591							
respondents)							

<sup>a</sup> The Nwoya, Nakasongola, Njombe and Kanungu research locations constitute a homogenous sub-set that is significantly different from the sub-set of Kanungu, Karatu and Iringa research locations with respect to the distribution of number of foreign farms within a distance of 1 km from respondent's home at 0.001 level of significance (one-way analysis of variance, Scheffe test).

<sup>b</sup> Three homogeneous sub-sets of research locations were identified with respect to the distribution of the number of foreign farms within a distance of between 1-5 km from respondent's home at 0.001 level of significance (one-way analysis of variance, Scheffe test): Nakasongola and Nwoya research locations as one sub-set; Nakasongola and Kanungu research locations as a second sub-set; and the three Tanzanian research locations as a third sub-set.

<sup>c</sup> Three mutually exclusive homogenous sub-sets of research locations were identified with respect to the distribution of the number of foreign farms within a distance of 5-10 km from respondent home at 0.001 level of significance (one-way analysis of variance, Scheffe test), namely Nwoya research location as the first sub-set; Njombe, Nakasongola and Kanungu research locations as the second sub-set; and Karatu and Iringa research locations as the third sub-set.

<sup>d</sup> Three mutually exclusive homogenous sub-sets of research locations were identified with respect to the distribution of the number of foreign farms within a distance of 10-25 km from respondent's home at 0.001 level of significance (one-way analysis of variance, Scheffe test), namely Kanungu research location as the first sub-set, Njombe research location as the second sub-set and Karatu, Iringa, Nwoya and Nakasongola research locations as the third sub-set.

<sup>e</sup> Four mutually exclusive homogenous sub-sets of research locations were identified with respect to the distribution of the number of foreign farms within a distance of more than 25 km from respondent's home at 0.001 level of significance (one-way analysis of variance, Scheffe test), namely Nwoya, Karatu and Njombe research locations as one sub-set, and Nakasongola, Iringa and Kanungu research locations each constituting a sub-set.

Assuming that the potential impact of the presence of foreign agricultural farms is greater the closer the foreign agricultural farm is to the respondent, a foreign farm exposure index was computed for each respondent so that foreign agricultural farms located within a distance of 10 kilometres from the residence of the respondent were assigned a weight of 3, farms located within a distance of between 10 and 25 kilometres from the residence of the respondent were assigned at a distance of 25 kilometres or above from the residence of the respondent were assigned a weight of 1.<sup>22</sup> The results are reported in Table 15 and Figure 2 below. Judged on the basis of proximity to and the number of foreign agricultural investments in their vicinity, the results confirm the picture in Tables 13 and 14 above and show that respondents in the three Tanzanian research locations are on average significantly more exposed to foreign agricultural investments in the three Ugandan research locations, with an average exposure index of 36.0 across the Tanzanian research locations as compared to 13.9 across the Ugandan research locations. Among the Tanzanian research locations, the population of the Karatu location is most exposed to foreign agricultural investments (average exposure index of 50.5), while among the Ugandan research locations, the Nwoya location is where the exposure to foreign agricultural investments is highest (average exposure index of 16.9).

Table 15. Exposure to foreign agricultural investments (exposure index) for respondents with residences within the median value of the distance to their research location centre, by research location

Research location	Exposure index to foreign agricultural farms					
	Median (50	25	75	Average	Standard	
	percentile)	percentile	percentile		deviation	
Karatu (n=119	49.0	49.0	51.0	50.5	2.9	
respondents) <sup>a</sup>						
Iringa (n=157	27.0	25.0	29.0	27.0	2.5	
respondents) <sup>a</sup>						
Njombe (n=130	37.0	28.0	38.0	33.7	5.4	
respondents) <sup>a</sup>						
Tanzania research	30.0	27.0	49.0	36.0	10.4	
locations (n=406						
respondents)						
Kanungu (n=200	13.0	11.0	15.0	12.7	1.9	
respondents) <sup>b</sup>						
Nwoya (n=194	16.0	14.0	17.3	16.9	4.0	
respondents) <sup>a</sup>						
Nakasongola (n=197	13.0	10.0	13.0	12.2	1.7	
respondents) <sup>b</sup>						
Uganda research	13.0	12.0	15.0	13.9	3.5	
locations (n=591						
respondents)						

Median, 25 and 75 percentiles, average and standard deviation, by research location

<sup>*a*</sup> Significantly different from all other research locations at 0.001 significance level (*p*=0.000; one-way analysis of variance, Scheffe test).

<sup>b</sup> The Kanungu and the Nakasongola research locations constitute a homogenous sub-set (p = 0.874) while significantly different from all other research locations at 0.001 significance level (p = 0.000; one-way analysis of variance, Scheffe test).

<sup>&</sup>lt;sup>22</sup> Different formulas were tested for computing the exposure index, all yielding similar overall patterns within and among the research locations.

## Figure 2. Exposure to foreign agricultural farms of respondents with residences within the median distance from their research locations centre, by research location

Box represents the 25 and 75 percentiles, line inside box represents median, whiskers represent 10 and 90 percentiles, and dots represent outliers



### Household poverty profile of the samples

The questionnaire survey was developed to provide the data necessary to replicate and recalculate the household poverty measure previously generated for several locations in Uganda and Tanzania (Ravnborg et al., 2004; 2013; Boesen et al., 2006), thereby establishing the poverty level of the household to which the individual respondent belongs. Thus, in addition to soliciting information about the respondent, the questionnaire also solicited information with respect to features of the household as a whole.

Following the approach described by Ravnborg and colleagues (Ravnborg et al., 1999), the household poverty measure was developed on the basis of explorations of local perceptions of household well-being and poverty undertaken through the use of well-being rankings that were translated into a set of household poverty indicators and combined into a household poverty index. The household poverty index was subsequently validated by assessing the extent to which the index captures and is correlated with the initial household well-being and poverty rankings.

Many of the aspects mentioned by informants during these explorations of local perceptions of household well-being and poverty were used to characterise the

well-being of households, irrespective of the perceived level. This applies to aspects such as housing quality and sources of non-agricultural income. Other aspects tended to be emphasised particularly to describe households enjoying a higher level of well-being, e.g. being able to hire in labour to undertake agricultural tasks, or to identify households enduring a low level of well-being, e.g. not being able to deal with health issues, should they occur. The set of household poverty indicators developed, seeks to reflect this by defining three levels for some indicators, for example, distinguishing high, medium and low levels of housing quality, while distinguishing only two levels for other indicators, e.g. the ability to hire in labourers. For more detail of the methodology, see Ravnborg et al. 1999, 2004.

Table 16 provides a summary of the household poverty indicators.

### Table 16. Household poverty indicators

Scoring system for household poverty indicators constituting the household poverty index; the higher the score, the higher the level of household poverty

Indicator	Score	Description
	33	Household has access to more than five acres of land
	67	Household has access to one to five acres of land
ILAND	100	Household does not have access to land or has access to less than one acre of land
INONAG	33	Somebody in the household has a 'high entry cost' non-agricultural source of income, like being a professional, being permanently employed, owning a shop or business (trading, e.g. timber, agricultural produce, charcoal, transport), renting out rooms, etc.
	67	Somebody in the household has a non-agricultural source of income like tailoring, craft-making, brewing beer, providing transport by driving <i>boda boda</i> , making and selling food etc., or the household receives remittances from family members working elsewhere, while nobody is engaged in high-entry cost activities
	100	Nobody in the household is engaged in any of the above non-agricultural sources of income
	33	Neither respondent nor household head (if different from respondent) works as a casual agricultural labourer at neighbouring local farms
ILABOUR	67	Respondent and/or household head (if different from respondent) work as casual agricultural labourers at neighbouring local farms, but do so during a maximum of three months a year, and at a maximum of three days a week
-	100	Household head (whether the respondent or someone else) works as casual agricultural labourers at neighbouring local farms more than three months per year, or three months or less per year, but then four days or more per week
	33	Somebody in the household has cattle or oxen, possibly together with other animals
IANIMAL	67	Nobody in the household has cattle, but they have other animals (goats, sheep, pigs, chicken, turkeys or rabbits)
	100	Nobody in the household has animals, not even chicken
IHIRE	33	Hire labourers for at least two of the following agricultural tasks: land clearing, ploughing, planting, weeding, harvesting or post-harvest processing
	67	Do not hire labourers or hire labourers for one agricultural task only
IFOOD	33	Have not experienced a period of food shortage within the last year

Indicator	Score	Description
		Have experienced a period of food shortage within the last year which
	67	lasted less than two months or which lasted longer but the only response
	07	was to eat less meat, using farm products rather than buying so much or
		buying food, or the husband working more as a day-labourer
	100	Have experienced a period of food shortage within the last year which
	100	lasted two months or more
	22	Eat rice at least once a week; fry food at least once a week; and eat meat at
		least once a month
IDIET	67	Eat rice less than once a week, or fry food only occasionally, or eat meat
	07	less than once a week
	100	Eat rice less than once a month and eat meat less than once a month
	22	Have houses with walls of bricks or plastered walls and iron or tile roofs,
	33	and which are well maintained
	(7	Have houses which may have iron/tile roofs or brick or plastered walls, but
IHOUSING	67	not both conditions at once, or have both but are in need of maintenance
	100	Have houses with walls made of old tins, banana or mud; or roofs that are
		grass thatched or made of polythene papers, banana fibre, old tins; or have
		houses that are in need of major repairs
		Nobody in the household had suffered from malaria, TB, HIV/AIDS,
	67	anaemia or chest-related diseases during the past year, or had done so, but
		the household had consulted the clinic using its own money without the
		need to borrow money from relatives, neighbours, etc.
IHEALTH		Somebody in the household had suffered from malaria, TB, HIV/AIDS,
		anaemia or chest-related diseases during the past year, but either the clinic
	100	had not been consulted due to a lack of money, or the clinic had been
		consulted with money borrowed from relatives, neighbours, or made
		available through the sale of land or other assets
		Have or have had children at secondary school or higher, and do not have
	33	children (girls or boys), including orphans, between six and twelve years
		who are not attending school
ISCHOOL	67	Do not have children, including orphans, between six and twelve years
	07	who are not attending school
	100	Have children (including orphans) between six and twelve years who are
	100	not attending school
	67	Household head is male or a married or co-habiting woman
IMAKITAL	100	Household head is a widow or single or divorced woman

Figure 3 shows the distribution of the respondents according to the options or scores defined along each of these indicators for the six research locations combined.



Number of respondents by score, by household poverty indicator



The household poverty index is computed as the average of the scores assigned to the households to which the respondent belongs according to these indicators. Figure 4 shows the distribution of respondents according to the household poverty index, while Figure 5 shows the distribution of respondents according to the household poverty index by research location. Based on an examination of the combination of scores along the household poverty index, threshold values along the household poverty index in order to define three household poverty categories (see Figure 4).

## Figure 4. Distribution of respondents according to household poverty index (N=2,385)

Number of respondents; red lines indicate limits between the household poverty categories 'nonpoor' (=<58.0); 'less poor' (58.0 – 68.0); and 'poorest' (>=68.0)



# Figure 5. Distribution of respondents according to household poverty index by research location<sup>23</sup> (N=2,385)

Box represents the 25 and 75 percentiles, line inside box represent median, whiskers represent 10 and 90 percentiles and dots represent outliers



 $^{23}$  The Nwoya research location is significantly different from all other research locations, while Njombe, Nakasongola and Iringa research locations constitute a homogenous sub-set (p = 0.200), Karatu and Kanungu another (p=0.547), all mutually significantly different at 0.01 level (p = 0.000; one-way analysis of variance, Scheffe test). In order to examine how each of the eleven household poverty indicators contribute to the overall household poverty index, i.e. what could be perceived as the 'internal logic' of the household poverty index, Figure 6 shows the average index values for each of the options for the eleven household poverty indicators. The figure shows that, as intended, the indicators for marital status and ability to deal with health issues primarily serve as indicators of the highest level of poverty (the deviation from the global mean is larger for the households receiving a score of '100' than that for the households receiving a score of '67'), while the indicator for hiring labourers serves as an indicator of a low level of poverty. Also the indicator for working as a day-labourer for neighbouring local farmers (ILABOUR) serves as an indicator of medium to high levels of poverty more than an indicator of a low level of poverty. The remaining indicators serve to differentiate between all three levels of poverty.



Figure 6. Average household poverty index value by options (score) for the eleven household poverty indicators for the Tanzania and Ugandan research locations combined (N=2,385 respondents)

In order to explore further the contribution of each of the eleven household poverty indicators, a multiple correspondence analysis<sup>24</sup> was conducted. This identified the mutual correlation between multiple categorial variables, in our case the eleven household poverty indicators and their levels, in our case the categories or options identified for each of the eleven indicators, through a process of dimension reduction. Using the iterative alternating least-square technique, numerical values

<sup>&</sup>lt;sup>24</sup> Multiple correspondence analysis (MCA) resembles principal component analysis, but for categorial variables.

(object scores) are assigned to the cases (objects) and categories so that objects within the same category are close together and objects in different categories are far apart.

Examining the first two dimensions, the analysis shows that the first dimension primarily captures the variance represented by the indicators for housing quality, the ability to hire in agricultural labourers and to ensure food security and an appreciated diet, while the second dimension captures the variance represented by the indicators for land availability and animal ownership (see Figure 7). Both dimensions appear to capture variance stemming from the indicators for having non-agricultural sources of income and day-labouring for neighbouring local farmers, while the indicators for the ability to provide education for one's children and in particular the gender and marital status of the household head and the ability to deal with health problems only contribute to the two dimensions to a limited extent.<sup>25</sup> As shown in Figure 8, the household poverty indicator scores indicating the highest level of poverty correspond to higher values along dimension 1, while a more mixed pattern applies along dimension 2. This pattern is further confirmed in Figure 9, which plots the respondents with their respective scores separately for the six research locations along the two dimensions computed through the multiple correspondence analysis. Common across the six research locations, the plots show that respondents belonging to the households categorised as the poorest tend to concentrate along the right-hand end of dimension 1, while being widely dispersed along dimension 2. As dimension 2 primarily reflects sources of income, i.e. land and animal ownership, engagement in non-agricultural sources of income and in day-labouring for neighbouring local farmers, this patterns suggest that there is no single combination of such sources of income that predicts household poverty level. Focusing on respondents belonging to households categorised as non-poor, they tend to be located towards the left-hand end of dimension 1, and in most of the research locations to be slightly more spread along dimension 1 than is the case for the poorest households. While still being spread, the non-poor and less poor households tend to achieve higher scores along dimension 2 than the poorest households. Once again, this suggests that no particular combination of sources of income can predict a households' level of poverty.

<sup>&</sup>lt;sup>25</sup> It should be noted, however, that the model captures the total variation to only a limited degree, given that Cronbach's Alpha for the first two dimensions is 0.592 and 0.305 respectively, while as a rule of thumb only values above 0.7 suggest an acceptable level of internal consistency of the model (in this case the data reduction represented by the computed dimension). This suggests that the variance represented by the eleven household poverty indicators can only partly be reduced to a limited number of dimensions.

Figure 7. Correlation between household poverty indicators (N=2,385 respondents/households)

Figure 8. Correlation between household poverty indicator options (N=2,385 respondents/households)

Discrimination measures resulting from multiple correspondence analysis

*Category quantifications (centroids) resulting from multiple correspondence analysis* 



Figure 9. Distribution of respondents by household poverty level along dimensions 126 and 2,27 resulting from multiple correspondence analysis among household poverty indicators



- <sup>26</sup> Significant correlation between dimension 1 score and household poverty level at 0.05 level (One-way ANOVA, Scheffe test) with means for all three poverty levels being mutually significantly different in all six research locations.
- <sup>27</sup> Mean score for respondents belonging to households categorised as the poorest being significantly different from the mean score for respondents belonging to households in the other two household poverty categories in Karatu and Nakasongola research locations at 0.05 level (One-way ANOVA, Scheffe test).



Finally, Table 17 shows the distribution of respondents according to household poverty category by research location. With more than half of the respondents in the Nwoya research location (55.8%) belonging to households categorised as the poorest, Nwoya is the research location with the most widespread poverty of the six research locations (see also Figure 5 above). At the other end of the scale, Njombe, Iringa and Nakasongola research locations are where a relatively higher share of respondents belong to households categorised as non-poor, ranging from 40.6 percent of respondents in Njombe research location to 35.5 percent and 30.4 percent in Nakasongola and Iringa research locations respectively.

# **Table 17. Distribution of respondents according to household poverty category by research location** Percent respondents per household poverty category, by research location

Research location	House	Household poverty category		
	Non-poor	Less poor	Poorest	categories
Karatu (n=397	24.4	37.8	37.8	100.0
respondents)				
Iringa (n=405	30.4	51.4	18.3	100.0
respondents)	<b>—</b>			
Njombe (n=401	40.6	44.9	14.5	100.0
respondents)	<b>—</b>			
Tanzania research	34.0	42.9	23.1	100.0
locations***				

(N=1,203 respondents)				
Kanungu (n=399 respondents)	 22.1	34.1	43.9	100.0
Nwoya (n=389 respondents)	 13.9	30.3	55.8	100.0
Nakasongola (n=394 respondents)	35.5	40.1	24.4	100.0
Uganda research locations*** (N=1,182 respondents)	23.9	34.9	41.3	100.0

\*\*\* Significant correlation between research locations and the distribution of respondents according to household poverty level; the significance of Pearson Chi-Square is < 0.001.

### **CONCLUDING REMARKS**

This paper has described the objectives, sampling approach and administration of a questionnaire survey conducted as part of the AIDA research programme in Tanzania and Uganda. It has also provided an overall characterisation of the resulting samples of respondents in the six research locations from which the samples were drawn. It has also described the process of constructing an exposure index assessing the exposure of respondents and their households to foreign agricultural investments, as well as a household poverty index which provides the basis for the classification of households according to their levels of poverty. The paper should therefore constitute a reference document for further analysis to be conducted as part of the AIDA research programme as well as a source of methodological inspiration of similar research efforts.

### REFERENCES

Bernard, H. Russell. 1994. Research methods in anthropology: Qualitative and quantitative approaches. 2nd ed. Sage Publications, California.

Boesen, J.; Raben, K.; Akello, Z.; Kidoido, M. & Lopa, D. 2006. "Poor people in environmental management in Uganda and Tanzania". DIIS Working Paper, 2006: 19. Copenhagen: Danish Institute for International Studies.

Broegaard, R.B. & Ravnborg; H.M. In preparation. "Dreams and motivations: Insights from a mapping of Danish agricultural investors in sub-Saharan Africa" (provisional title).

Deininger, K. and Xia, F. 2016 "Quantifying Spillover Effects from Large Landbased Investment: The Case of Mozambique." World Development, 2016, 87:C, 227-241.

Ravnborg, Helle Munk, with the collaboration of R.M. Escolán, M.P. Guerrero, M.A. Méndez, F. Mendoza, E.M. de Paez and F. Motta. 1999. "Developing regional poverty profiles based on local perceptions" (also available in Spanish). CIAT publication no. 291, 56 pp. Cali: Centro Internacional de Agricultura Tropical.

Ravnborg, H.M.; Boesen, J.; Sørensen, A.M.; Akello, Z.; Bashasha, B.; Kasozi, S.; Kidoido, M. and Wabukawo. V.2004. "Gendered District Poverty Profiles and Poverty Monitoring, Kabarole, Masaka, Pallisa, Rakai and Tororo Districts, Uganda". DIIS Working Paper 2004: 1. Copenhagen: Danish Institute for International Studies.

Ravnborg, H.M.; B. Bashaasha; R.H. Pedersen; R. Spichiger and A. Turinawe. 2013. "Land Tenure under Transition: Tenure Security, Land Institutions and Economic Activity in Uganda." DIIS Working Paper 2013:03. Copenhagen: Danish Institute for International Studies.

United Republic of Tanzania-URT. 2013. "2012 Population and Housing Census-Population Distribution by Region, District, Ward and Village". Dar es Salaam and Zanzibar: National Bureau of Statistics, Ministry of Finance and Office of Chief Government Statistician, President's Office, Finance, Economy and Development Planning.

Uganda Bureau of Statistics. 2016. "The National Population and Housing Census 2014". Kampala: Uganda Bureau of Statistics.