Components of Income Aggregate:

"National Panel Survey- Tanzania 2009[1]"

Prepared for the Rural Income Generating Activities (RIGA) Project[2]

of the Agricultural Development Economics Division,

Food and Agriculture Organization

January, 2012

This document provides the survey-specific details associated with the income aggregate construction. For more information about the RIGA project, please refer to http://www.fao.org/es/esa/riga. For additional detail regarding the overall RIGA income aggregate construction approach, please refer to Carletto, et al (2007), "Rural Income Generating Activities Study: Methodological note on the construction of income aggregates," found on the RIGA website.

The Tanzania National Panel Survey (NPS-1) was carried out for twelve months from October 2008 to October 2009[3]. The survey is the first wave of a panel for which the follow-up data collection was to take place in 2010. It collected data using Household, Agricultural and Community questionnaires and obtained information at the individual, household, plot, business and community levels.

The sample for NPS-1 was drawn using a multi-stage stratified random sampling procedure from three sampling frames: (1) the 2002 Population and Housing Census; (2) the 2007 Household Budget Survey; and (3) the 2002 National Sample Census of Agriculture. The full sample contained 3280 households from 410 Enumeration Areas (EAs), each containing 8 households, and it includes a subsample from the 2007 HBS.

The survey was sampled to be nationally representative. In order to obtain nationally representative statistics from the NPS data, it is necessary to apply the sampling weights provided in the data. The sampling weights variable in the original data is called "HH_WEIGHT"; it is renamed to "WEIGHT" in the RIGA datasets.

[1] The information provided in this document relies substantially upon the Basic Information Document (BID), provided with the NPS-1 data.

[2] The RIGA Project is a collaboration between FAO, the World Bank and American University in Washington, D.C. Original data can be obtained from the World Bank's Living Standards Measurement Study by visiting the LSMS website at: http://www.worldbank.org/lsms.

[3] Source: 2009 Tanzania BID.

In the original datasets, the various household-level modules of the NPS-1 data households can be linked by the variable HHID. Agricultural module datasets can be linked either with the HHID variable or by combining HHID with the unique plot identifier, PLOTNUM. The variable HHID is renamed to "HH" for the final RIGA datasets.

In the original datasets, "LOCALITY" is the variable that identifies whether households are urban, rural, or "mixed" which refers to a peri-urban area. The variable "RURAL" distinguishes urban from rural households in a definition that groups the "mixed" households into their corresponding location. There are 2,057 rural household and 1,201 urban households in the dataset. In the do files, "RURAL" is recoded to "URBAN" in order to use the same variable name across different RIGA surveys.

Regarding income from different sources, revenues and costs were disaggregated when such information was available. The disaggregated sources for each income component are summarized in output variables column of Table 1. The net variables and the data files included in the final total income aggregate (Income.dta) are in **bold**. **Unless otherwise noted, all variables included in the aggregate income variable are net of costs**.

An average household size in Tanzania is 5.2 persons[4]. All money amounts are in Tanzania Schillings (TZS). In 2009, the official exchange rate[5] was TZS 1,320= $1.0. **The income aggregates are calculated at the household level and all aggregates are annualized.**

Comments

- In all sections, the raw data undergoes a transformation (it is annualized, aggregated, taken from person – household level, etc) before a check for outliers takes place.
- The industry codes used for classifying wage employment follow the United Nations International Standard Industrial Classification (ISIC) standards. Given the survey classification of each employed household member by industry, the employment sectors include: Agriculture and fishing, Mining, Manufacturing, Electricity and utilities, Construction, Commerce, Finance insurance and real state, Services and Unknown.
- The classification of non-farm enterprise activities into industries categories follows the same classification system as the employment section. Given these standards, the non-farm enterprise sectors include: (1) Agriculture, Livestock, Hunting and Fishing, (2) Mining, (3) Manufacturing, (4) Electricity and Utilities, (5) Construction, (6) Commerce, (7) Transportation, Storage and Communications, (8) Finance, Insurance and Real Estate, (9) Services and (10) Other Industries.
- For all sections, whenever information was available regarding the share of a business, enterprise, or any other income activity owned by the household, the income earned from that activity was weighted by the share owned by the household.
- A final outlier check is imposed at the end of the Aggregateincome.do file in which households with income shares from any given activity greater than or less than 3 (300%) are dropped from the final income aggregate. Using these criteria, 7 households are dropped from this survey.

[4] RIGA project calculations.
[5] Exchange rate used comes from the World Bank World Development Indicators database.

- Participation and income share variables for all income components are included in the final income aggregate

The programs that calculate each household's income aggregate component are summarized in Table 1. Tables 2 and 3 summarize the results from the created income aggregate.

Table 1

Do file	Input data files HHQ= household questionnaire AGQ= agricultural questionnaire	Output data files	Main variables constructed	Notes/Decisions
Sample.do	HHQ\SEC_A_T.dta AGQ\SEC_1_ALL.dta	Sample.dta		Create household listing of roster. Dropped hhid=="11010020050092" as all info was missing for this household.
Prices.do	HHQ\SEC_K1.dta AGQ\SEC_7B.dta AGQ\SEC_7A.dta AGQ\SEC_5A.dta AGQ\SEC_5B.dta	price_purch_ea.dta price_purch_region.dta price_purch_urb.dta price_purch_unit.dta price_prod_ea.dta price_prod_region.dta price_prod_urb.dta price_prod_unit.dta	price_purch_ea price_purch_region price_purch_urb price_purch_unit price_prod_ea price_prod_region price_prod_urb price_prod_unit	Created median prices for each crop at different administrative levels (enumeration area, region, locality) using the reported expenditures in the consumption module and the reported sales from the agricultural module. For HHQ\SEC_K1.dta: replace unit=1 if skcode==1003 & quantity==250 & unit==0
Foodown.do	INHH\SEC_K1.dta price_purch_XX.dta (from prices.do) price_prod_XX.dta (from prices.do)	Food.dta Foodonw_crop.dta Foodown_livestock.dta	foodgift foodown foodpurch foodcons foodown_crop foodown_livstock	Dropped non-food items (skcode>1103) Food own consumption if source==3 (home produced). Food consumption is classified as purchased if source==3, as gift if source==6 and as other non-purchased consumption if source==2. Did not have frequency of consumption so was obliged to annualize weekly food expenditures multiplying by 52. Performed two outlier checks: one by crop code, and one by region.
Agother.do	AGQ\SEC_11_ALL.dta AGQ\SEC_13A.dta AGQ\SEC_13B.dta	agotherexp.dta (expenditure on technical assistance and rental of farm implements)	agotherexp	Agotherexp is the sum of variables s13q5, s13q9 and s11q9 Outlier checks by region (for s13q5 and s13q9) and by item code for s11q9.
Rentagric.do	AGQ\SEC_3A.dta AGQ\SEC_3B.dta	Rentagric.dta	farmrnt farmrntexp sharecropexp	rental income: (s3aq4 if s3aq3==2) + (s3bq4 if s3bq3==2) rental expense: (s3aq30 if s3aq32==1) + (s3bq30 if s3bq32==1) share crop exp: (s3aq30 if s3aq32!=1) + (s3bq30 if s3bq32!=1) Outlier checks by region.
Cropincome.do	AGQ\SEC_3A.dta AGQ\SEC_3B.dta	cropexp1.dta (inputs) cropexp2.dta (labour)	seedexp fertexp	Input values based on section 3, questions 57, 58, 59, 63, 41, 46, 52, 56, and section 4 questions 20.

	AGQ\SEC_4A.dta AGQ\SEC_4B.dta AGQ\SEC_5A.dta AGQ\SEC_5B.dta AGQ\SEC_6A.dta AGQ\SEC_6B.dta AGQ\SEC_7A.dta AGQ\SEC_7B.dta AGQ\SEC_9_ALL.dta Foodown_crop.dta	cropinc.dta (revenues, other exp) cropbyprod.dta (by-products)	pestexp laborexp harvestv soldv transportexp storedv lostv owncons byprodinput byprodsold byprodexp cropincome1 cropincome2	In kind payments valued using median prices estimated in price_prod_XX.dta files. For labour hired-in on a daily basis, days of labour are valued according to median daily wages at different administrative levels. Own consumption from the agricultural module is estimated as the quantity of kgs harvested minus sales, stored and lost crop and is valued using median sales prices from the agricultural module. In SEC_9_ALL.dta, variable s9q6_1 recoded one quantity of 126563 to missing. By-product inputs according to median prices. Outlier checks for by-products by item code and region.
Fish.do	AGQ\SEC_12A.dta AGQ\SEC_12B.dta AGQ\SEC_12D.dta	fishexp1.dta (fish durables repair exp) fishsold.dta fishexp2.dta (fishing input exp) fishinc.dta	fishexp1 fishsold fishexp2	Outlier checks according to item code and region. Fish income and expenditure variables considered: s12q15, s12q9, s12q27, s12q28
Livestock.do	AGQ\SEC_10A.dta AGQ\SEC_10B.dta fishinc.dta Foodown_livestock.dta	livstinc.dta livstbyprod.dta Livestock.dta	livstborn livstsold livstexp livstlost livstbyprodsold livstinc	Quantities valued according to median unit prices estimated by region, locality and the overall sample. Outliers checked by item code (where relevant) and by region. Livestock income includes income from fishing activities.
Selfemp.do	HHQ\SEC_B_C_D_E1_F_G1_U.dta HHQ\SEC_E2.DTA	Selfemp.dta	selfimp1 selfimp2 selfimp3 selfimp4 selfimp5 selfimp6 selfimp7 selfimp8 selfimp9 selfimp10	Industries classified according to variable seq24 Expenditures calculated from variable sseq38, seq39, seq40. Revenues estimated from variables seq35_2 (gross income). Net income obtained by subtracting expenditures from gross income. In the cases where the net income was negative, strangely low, or did not match the net reported self employment income (Seq36_2), the value was replaced with the reported net self employment income (seq36_2)

				if the reported net income equalled the average reported self employment income (seq42). This affected 55 of 1605 observations. Outliers were checked by industry and by region. Annualization based on the number of months of the year worked (seq41).
Employment.do	HHQ\SEC_B_C_D_E1_F_G1_U.dta	Employment.dta	wge1_3 wge2_3 wge3_3 wge4_3 wge5_3 wge6_3 wge7_3 wge8_3 wge9_3 wge10_3	Industries classified according to variable seq13 Skill level could not be classified due to absence of occupational codes. For annualization assumed 4.3 weeks per month, 6 days per week worked. Both wage income (seq18_1) and in kind income (seq21_1) considered. Outliers checked by industry and region.
Transfers.do	HHQ\SEC_G2.dta HHQ\SEC_O1.dta	privtrans.dta socialtrans.dta Transfers.dta	privtrans socialtrans pubtrans transfergross	Private transfers include sgq11 Social transfers consider soq3 soq4 soq5 Total transfers are gross. Outliers checked by region and transfer type (for social transfers).
Aggregateincome.do	Sample.dta Rentagric.dta Agbyprod.dta Cropincome.dta Livestock.dta Employment.dta Otherincome.dta Selfemp.dta Transfers.dta	Income.dta	agr_wge nonagr_wge crop1 crop2 livestock other selfemp transfers totincome1 totincome2	For each income source, participation variables are constructed (prefixed by "p_") as well as share variables (prefixed by "sh1" or "sh2") Different aggregations of income sources are also constructed such as onfarm (crop and livestock), offfarm (agr_wge nonagr_wge, other, selfemp, transfers), non-farm (non-agrwge and selfemp) nonag (nonagr_wge, other, selfemp, trnasfers) and agricultural (agr_wge, crop and livestock). A final outlier check is incorporated that drops households that end up with income shares from the major categories (sh2agr_wge, sh2nonagr_wge, sh2crop2, etc) as greater than 300%. 7 observations dropped as a result.

Table 2

Tanzania 2009	2,056 Rural HH Observations	Rural, Weighted, Shillings						Rural, Weighted, USD	
Variable		*# Participants*	*Participation Rate*	*Returns to Participation- Participant HHs*	*Returns to Participation- All HHs*	*Share of Total Income- All HHs (Mean of Shares)*	*Share of Total Income- All HHs (Share of Means)*	*Returns to Participation- Participant HHs*	*Returns to Participation- All HHs*
agr_wge	Wage Employment- Agriculture	401	22%	137316	29,949	6%	4%	104	23
nonagr_wge	Wage Employment- Nonfarm	334	15%	654827	96,006	7%	13%	496	73
crop1	Crop Production	1918	96%	216469	207,136	45%	27%	164	157
livestock	Livestock Production	1198	61%	273365	167,222	13%	22%	207	127
selfemp	Non-ag Self Employment	686	34%	632038	215,336	14%	28%	479	163
transfer	Total Transfers	1172	57%	86109	49,357	13%	6%	65	37
other	Other Income Sources	19	1%	38146	442	0.3%	0.1%	29	0.3
totincome1	Total Household Income-crop1	2047	100%	768371	765,447	100%	100%	582	580

Percent Rural (Weighted)	**73%**
Shillings/USD (2009)	***1,320.31***

Table 3

Tanzania 2009	2,056 Rural HH Observations	Rural, Weighted, Shillings						Rural, Weighted, USD	
Variable		*# Participants*	*Participation Rate*	*Returns to Participation- Participant HHs*	*Returns to Participation- All HHs*	*Share of Total Income- All HHs (Mean of Shares)*	*Share of Total Income- All HHs (Share of Means)*	*Returns to Participation- Participant HHs*	*Returns to Participation- All HHs*
agr_wge	Wage Employment- Agriculture	401	22%	137316	29,949	5%	3%	104	23
nonagr_wge	Wage Employment- Nonfarm	334	15%	654827	96,006	6%	10%	496	73
crop2	Crop Production	1956	97%	456628	441,940	52%	44%	346	335
livestock	Livestock Production	1198	61%	273365	167,222	13%	17%	207	127
selfemp	Non-ag Self Employment	686	34%	632038	215,336	13%	22%	479	163
transfer	Total Transfers	1172	57%	86109	49,357	10%	5%	65	37
other	Other Income Sources	19	1%	38146	442	0.1%	0.0%	29	0.3
totincome2	Total Household Income-crop2	2051	100%	1002640	1,000,252	100%	100%	759	758

Percent Rural (Weighted)	**73%**
Shillings/USD (2009)	***1,320.31***