

# Components of Income Aggregate: “Living Standards Survey, Nigeria 2003-2004”

*Prepared for the Rural Income Generating Activities (RIGA) Project<sup>1</sup>*

*of the Agricultural Development Economics Division,*

*Food and Agriculture Organization*

*February, 2010*

---

*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 Nigeria Living Standards Survey (NLSS) survey was carried out for twelve months from September 2003 through to August 2004, households being interviewed over seven visits.

The total sample was 19,158 households. The average household size in Nigeria is 4.74 persons,<sup>2</sup> money amounts are in Nigerian Naira. In August, 2003, the official exchange rate<sup>3</sup> was 132.89 Naira = US\$1.0. **The income aggregates are calculated at the household level and all aggregates are annualized.**

In the original datasets, the various modules of the NLSS data households can be linked by the unique household identifier, CASE\_ID. This variable is renamed to HH in the final aggregate for consistency across surveys.

In the original datasets, SECTOR is the variable that identifies urban households from rural households. There are 14,512 rural household and 4,646 urban households in the dataset. In the do files, SECTOR is renamed to URBAN in order to use the same variable name across different surveys.

Regarding income from different sources, revenues and costs were disaggregated when such information was available. The disaggregated sources for each income component are

---

<sup>1</sup> 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/lsm>.

<sup>2</sup> Author’s calculations using source data.

<sup>3</sup> Exchange rate used comes from the World Bank World Development Indicators database.

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**

## Comments

- **Please Note: In 2009 an update to the original data for the NLSS 2003-4 was obtained from the Nigeria Ministry of Statistics. As a result, these newer datasets were used to update the RIGA income aggregate for Nigeria from the earlier set of data, originally received in 2006.**
- 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. Outlier checks are done at three standard deviations from the median.
- In the Crop Production section, the reference period is the previous 12 months. Two total crop income variables are created. Cropincome1 includes estimates of own crop consumption based on the crop section of the household questionnaire. Cropincome2 includes estimates of own crop consumption based on the food expenditure section of the questionnaire.
- In the Livestock, Other Income, and Transfers sections the reference period is the previous twelve months.
- In the Wage section, the reference period is the duration of employment as specified in the questionnaire for the respondent's main, second, third and fourth jobs held in the last 12 months. Employment earnings are disaggregated by industry following the occupation activity code (the industry codes are not reported in the survey) and descriptions provided by the respondent. The industry groups, which correspond with the United Nations International Standard Industrial Classification for All Economic Activities (ISIC) include: (1) Agriculture 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 and (9) Services and (10) Unknown. The employment categories are further disaggregated into skilled, unskilled, and other based on the ISIC Occupation Classification code. Employment income can be reported by the respondent in-cash or in-kind as net or gross of taxes and it includes all bonuses, commissions, tips and benefits received.
- The Self Employment (Selfemp) section accounted for income from non-farm enterprises owned by the household. For expenditures, the reference period was the last two weeks, three months and twelve months. Given problems with reporting, we use twelve months expenditures, and replace them with annualized 3 months expenditures, or annualized 2 week expenditures if annual expenditures are not reported. For revenues, the reference period is the last two weeks, with the exception of the past 12 months for revenues from the rental of enterprise assets. Annualization is done based on the reported number of months, the enterprise was in operation.
- 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. 686 enterprises are missing their industrial classification, but report expenditures and revenues, consequently they were classified as Other.

- 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, 70 households are dropped from this survey.
- Only information on present household members is considered in the final income aggregate. Present is defined as a having been absent from the households for 6 or fewer months in the previous 12 months for non-head household members. Household heads are included regardless of their location status since they can still serve as primary income sources. Thus, we drop observations if S1Q18 > 6 & S1Q3 != 1.
- Participation and income share variables for all income components are included in the final income aggregate.
- **Data problems:**
  - The data for section 9E (other agricultural by-products) was miscoded in the raw data and showed a suspicious pattern of over-reporting for the three states 22, 23 and 24, therefore, a recoding was implemented for hunting, palm wine and fruit and berries reporting (variables s9eq1, s9eq3, s9eq4). As a result, we were able to obtain 100, 122 and 76 clean observations for the three problematic states. The following recoding was implemented to correct for this problem:
    - replace s9eq1 = . if s9eq1 > 111600 (161 observations affected)
    - replace s9eq3 = . if s9eq3 > 99010 (382 observation)
    - replace s9eq3 = . if s9eq3 == 10042| s9eq3 == 6022| s9eq3 == 6032| s9eq3 == 4032| s9eq3 == 4042 (6 observations)
    - replace s9eq3 = . if s9eq3 == 24022| s9eq3 == 24052| s9eq3 == 12032| s9eq3 == 12012| s9eq3 == 12062 (7 observations)
    - replace s9eq3 = . if s9eq3 == 4011| s9eq3 == 3011| s9eq3 == 3012| s9eq3 == 1011| s9eq3 == 1012| s9eq3 == 1031 (21 observations)
    - replace s9eq4 = . if s9eq4 > 20000 (10 observations)
  - The reported labor expenditures for section 9G (crop by-product processing) were miscoded in the raw data and showed a suspicious pattern of over-reporting for the three states 22, 23 and 24, therefore, a recoding was implemented for the labor expenditures reporting (variable s9gq7). As a result, we were able to obtain 93, 224 and 23 clean observations for the three problematic states. The following recoding was implemented to correct for this problem:
    - replace s9gq7 = . if s9gq7 > 240000 & s9gq7 != .0 (2265 observations affected)
  - The data set for section 9A3 (Agricultural equipment rent expenditures) was not available in the revised version of the raw data (received in 2009) and consequently the original raw data file `sec9a3_agri_equipment.dta` was used. The variable s9aq36 (yes/no question) was miscoded including values 0 and 5. Zeros were recoded to 2 (2194 observations) In addition, s9aq37 was miscoded in the raw data and showed a suspicious pattern of over-reporting for the three states 22, 23 and 24, therefore, a recoding was implemented for the rental income reporting (variable s9aq37). As a result, we were able to obtain 93, 224 and 23 clean observations for the three problematic states. The following recoding was implemented to correct for this problem:
    - replace s9aq37 = . if s9aq37 > 350000 & s9aq37 != .0 (6 observations affected)
  - The data for section 9F (crop expenditures) was miscoded in the raw data and showed a suspicious pattern of over-reporting for the three states 22, 23 and 24, therefore, a

- recoding was implemented for the variable s9fq2. As a result, we were able to obtain 88, 35 and 12 clean observations for the three problematic states. The following recoding was implemented to correct for this problem:
- replace s9cq2 == . if s9cq2 > 572620 (56 observations affected)
- The data for section 9C2 was miscoded in the raw data and showed a suspicious pattern of over-reporting for the three states 22, 23 and 24, therefore, a recoding was implemented for the s9cq22 variable. As a result, we were able to obtain 26, 9 and 0 clean observations for the three problematic states. The following recoding was implemented to correct for this problem:
- replace s9fq2 = . if s9fq2 > 1114500 & s9fq2 != . (1100 observations affected)
- The data for the section on self-employment expenditures was miscoded in the raw data and showed a suspicious pattern of over-reporting for the three states 22, 23 and 24, therefore, a recoding was implemented for the s11bq7, s11bq8, and s11bq9 variables. As a result, we were able to obtain 420, 805 and 867 clean observations for the three problematic states. The following recoding was implemented to correct for this problem:
- replace s11bq7 = . if s11bq7 > 920000 & s11bq7 != . (28 observations)
  - replace s11bq8 = . if s11bq8 > 620720 & s11bq8 != . (8 observations)
  - replace s11bq9 = . if s11bq9 > 300000 & s11bq9 != . (1357 observations)
  - replace s11bq9 = . if s11bq9 == 122221 (1 observation)
- 646 rural households<sup>4</sup> had incomplete income information in the survey such that these households do not have a constructed income aggregate. The missing households are primarily on the upper end of the expenditure distribution and are geographically concentrated in the south of the country (70.51%). The average household size is 3.35 persons with an average education of 6.92 years of schooling among the working-age individuals and 5.12 years of schooling among the household heads. Nearly 25% of the heads are female, 46% are single, and 24.7% are Muslim. Further, the households are primarily landless (97.24%); the 16 land-owning households hold on average 4.24 hectares.

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 dta files	output dta files	output variables
Sample.do	\$RAWHH\Sec 1 roster.dta	\$CON\Sample.dta	
Interper.do	\$RAWHHOLD\intro1_visit.dta \$RAWHHOLD\intro2_visit.dta \$RAWHHOLD\intro3_visit.dta	\$CON\interper.dta	
Price.do	\$RAWHH\Sec 9H own consumption.dta \$RAWHH\Sec 9G agric processing.dta \$RAWHH\Sec 9C1 agric harvest cash.dta \$RAWHH\Sec 9C2 agric harvest subsistence.dta	\$TEMP\price_prod_*.dta \$TEMP\price_purch_*.dta	price_prod_* price_purch_*

<sup>4</sup> A corresponding 738 urban households also had missing income information. Although these urban households may have similar characteristics, the profile of the missing households pertains only to rural households.

Foodown.do	\$TEMP\interper.dta \$RAWHH\Sec 9H own consumption.dta \$TEMP\price_*.dta	\$TEMP\cropown.dta \$TEMP\livstown.dta \$TEMP\fishown.dta \$CON\foodown.dta	cropownimp livstownimp fishownimp foodown
Agrent.do	\$RAWHH\Sec 9a1 agric land.dta \$RAWHH\Sec 9B agric plots.dta \$RAWHHOLD\sec9a3_agri_equipment.dta	\$TEMP\farmrntexp.dta \$TEMP\agrent.dta \$INC\farmrnt.dta"	farmrntexpimp agrentimp shrcropincimp farmrntimp
Agricbyprod.do	\$RAWHH\Sec 9G agric processing.dta \$RAWHH\Sec 9E agric other income.dta	\$TEMP\agother.dta \$TEMP\agricbyprod.dta	cropbypr2imp livstbypr2imp cropbyprimp
Cropinc1	\$RAWHH\Sec 9F agric costs & expenses.dta	\$TEMP\cropexpenses.dta	cropexpimp
Cropinc2	\$RAWHH\sec 9C1 agric harvest cash.dta \$RAWHH\sec 9C2 agric harvest subsistence.dta \$TEMP\cropinc.dta \$TEMP\cropexpenses.dta \$TEMP\agrent.dta \$TEMP\farmrntinc.dta \$TEMP\agricbyprod.dta \$TEMP\cropown.dta \$TEMP\agother.dta	<b>\$INC\cropincome.dta</b>	cropsoldimp cropinputimp sharecropexpimp cropown1imp laboreximp  <b>cropincome1imp</b> <b>cropincome2imp</b>
Employment.do	\$RAWHH\Sec 4 employment.dta	\$TEMP\employ1.dta \$TEMP\employ2.dta \$TEMP\employ3.dta \$TEMP\employ4.dta <b>\$INC\Employment.dta</b>	wgeimp <b>wgeimp1_1</b> <b>wgeimp2_1</b> <b>wgeimp3_1</b> <b>wgeimp4_1</b> <b>wgeimp5_1</b> <b>wgeimp6_1</b> <b>wgeimp7_1</b> <b>wgeimp8_1</b> <b>wgeimp9_1</b> <b>wgeimp10_1</b> <b>wgeimp1_2</b> <b>wgeimp2_2</b> <b>wgeimp3_2</b> <b>wgeimp4_2</b> <b>wgeimp5_2</b> <b>wgeimp6_2</b> <b>wgeimp7_2</b> <b>wgeimp8_2</b> <b>wgeimp9_2</b> <b>wgeimp10_2</b> <b>wgeimp1_3</b> <b>wgeimp2_3</b> <b>wgeimp3_3</b> <b>wgeimp4_3</b> <b>wgeimp5_3</b> <b>wgeimp6_3</b> <b>wgeimp7_3</b>

			<b>wgeimp8_3 wgeimp9_3 wgeimp10_3</b>
Fishing.do	\$RAWHH\Sec 9F agric costs & expenses.dta \$RAWHH\Sec 9a2 agric livestock.dta" \$TEMP\fishown.dta	\$TEMP\fishexpenses.dta \$TEMP\fishincomes.dta <b>\$INC\fishinc.dta</b>	fishnptimp fishsoldimp fishrentimp <b>fishincimp</b>
Livestock.do	\$RAWHH\Sec 9F agric costs & expenses.dta \$RAWHH\Sec 9a2 agric livestock.dta \$TEMP\livstown.dta	\$TEMP\livstexpenses.dta \$TEMP\livstinc.dta \$INC\livestock.dta	livstexpimp livstrevimp livstrentimp <b>livstincimp</b>
Otherincome.do	\$RAWHH\Sec 13A3&A4 misc income.dta	\$INC\otherincome.dta	<b>otherincimp</b>
Selfemp.do	\$RAWHH\Sec 11A enterprise basic.dta \$RAWHH\Sec 11B first enterprise expd.dta \$RAWHH\Sec 11B second enterprise expd.dta \$RAWHH\Sec 11B third enterprise expd.dta \$RAWHH\Sec 11E net income and inventory.dta \$RAWHH\Sec 11D enterprise revenues.dta	\$TEMP\selfemplexp.dta \$TEMP\selfempl.dta \$TEMP\nonfarmxter.dta <b>\$INC\selfemp.dta</b>	hhrevyrimp hhexpimp <b>selfimp1 selfimp2 selfimp3 selfimp4 selfimp5 selfimp6 selfimp7 selfimp8 selfimp9 selfimp10</b>
Transfers.do	\$RAWHH\Sec 13A1 transfersOUT.dta \$RAWHH\Sec 13A2 transfersIN.dta \$RAWHH\Sec 13A3&A4 misc income.dta	\$TEMP\transferprivout.dta \$TEMP\transferprivinc.dta \$TEMP\pensions.dta <b>\$INC\transfers.dta</b>	transfersentimp transferincimp pensionsimp socialtransimp privtrans pubtrans <b>transferstotimp transfersgrossimp</b>
<b>AggregateIncome.do</b>	<b>\$INC\Farmrnt.dta \$INC\Fishinc.dta \$INC\Cropincome.dta \$INC\Employment.dta \$INC\Livestock.dta \$INC\Otherinc.dta \$INC\Selfemp.dta \$INC\Transfers.dta</b>	<b>\$INC\Income.dta</b>	<b>Agr_wge Nonagr_wge Crop1 Crop2 Livestock Selfemp Transfers Other</b>

**Table 2**

<i>Nigeria 2004</i>		13,821 Rural HH Observations	Rural, Weighted, Naira				Rural, Weighted, USD			
<i>Variable</i>		<i># Participants</i>	<i>Participation Rate</i>	<i>Returns to Participation- Participant HHs</i>	<i>Returns to Participation- All HHs</i>	<i>Share of Total Income- All HHs (Mean of Shares)</i>	<i>Share of Total Income- All HHs (Share of Means)</i>	<i>Returns to Participation- Participant HHs</i>	<i>Returns to Participation- All HHs</i>	
<b>agr_wge</b>	Wage Employment- Agriculture	156	1.2%	125,857	1,462	0.6%	1.25%	947	11	
<b>nonagr_wge</b>	Wage Employment- Nonfarm	1,261	9.7%	247,197	23,857	7.4%	20.47%	1,860	180	
<b>crop1</b>	Crop Production	11,713	84.5%	79,409	67,115	73.9%	57.59%	598	505	
<b>livestock</b>	Livestock Production	5,712	39.9%	11,152	4,454	6.0%	3.82%	84	34	
<b>selfemp</b>	Non-ag Self Employment	2,158	17.3%	103,968	17,986	9.9%	15.43%	782	135	
<b>transfer</b>	Total Transfers	769	6.1%	18,162	1,109	1.5%	0.95%	137	8	
<b>other</b>	Other Income Sources	546	3.8%	14,669	556	0.6%	0.48%	110	4	
<b>totincome1</b>	Total Household Income-crop1	13,356	97.0%	120,177	116,538	100.00%	100%	904	877	

<b>Percent Rural (Weighted)</b>	<b>54.80%</b>
<b>Naira/USD (August 2003)</b>	<b>132.89</b>

1. Source data: 2003-2004 Living Standards Survey
2. Exchange rate used comes from the World Bank World Development Indicators database.
3. The variable "crop1" is distinguished from "crop2" in the way home consumption of own production of crops (owncons) is calculated. In crop1, owncons is the amount reported in the crop section of the questionnaire. For crop2, owncons is calculated from the "Consumption of Own Produce" section of the household questionnaire. Total household income "totincome1" and "totincome2" are therefore calculated with the corresponding crop income variable.
4. All values reported are annual and net of costs (with the exception of income from transfers and land rent, which are gross receipts).

**Table 3**

<i>Nigeria 2004</i>		13,821 Rural HH Observations		Rural, Weighted, Naira			Rural, Weighted, USD		
<i>Variable</i>		<i># Participants</i>	<i>Participation Rate</i>	<i>Returns to Participation- Participant HHs</i>	<i>Returns to Participation- All HHs</i>	<i>Share of Total Income- All HHs (Mean of Shares)</i>	<i>Share of Total Income- All HHs (Share of Means)</i>	<i>Returns to Participation- Participant HHs</i>	<i>Returns to Participation- All HHs</i>
<b>agr_wge</b>	Wage Employment- Agriculture	156	1.2%	125,857	1,462	0.6%	1.24%	947	11
<b>nonagr_wge</b>	Wage Employment- Nonfarm	1,261	9.7%	247,197	23,857	7.2%	20.30%	1,860	180
<b>crop2</b>	Crop Production	12383	88.8%	76,654	68,102	75.9%	57.95%	577	512
<b>livestock</b>	Livestock Production	5712	39.9%	11,152	4,454	4.6%	3.79%	84	34
<b>selfemp</b>	Non-ag Self Employment	2158	17.3%	103,968	17,986	9.7%	15.30%	782	135
<b>transfer</b>	Total Transfers	769	6.1%	18,162	1,109	1.5%	0.94%	137	8
<b>other</b>	Other Income Sources	546	3.8%	14,669	556	0.5%	0.47%	110	4
<b>totincome2</b>	Total Household Income-crop2	13794	99.8%	117,800	117,525	100.00%	100%	886	884

<b>Percent Rural (Weighted)</b>	<b>54.80%</b>
<b>Naira/USD (August 2003)</b>	<b>132.89</b>

1. Source data: 2003-2004 Living Standards Survey
2. Exchange rate used comes from the World Bank World Development Indicators database.
3. The variable "crop1" is distinguished from "crop2" in the way home consumption of own production of crops (owncons) is calculated. In crop1, owncons is the amount reported in the crop section of the questionnaire. For crop2, owncons is calculated from the "Consumption of Own Produce" section of the household questionnaire. Total household income "totincome1" and "totincome2" are therefore calculated with the corresponding crop income variable.
4. All values reported are annual and net of costs (with the exception of income from transfers and land rent, which are gross receipts).