

Components of the Income Aggregate: “Household Living Standards Survey (VHLSS), Vietnam 2002”¹

Prepared for the Rural Income Generating Activities (RIGA) Project²

of the Agricultural Development Economics Division,

Food and Agriculture Organization

December, 2009

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 Vietnam Household Living Standards Survey (VHLSS) was carried out over a period of twelve months in 2001 and 2002.

The VHLSS was designed by two samples consisting of households and communes/wards selected in 61 central provinces: one large sample of 45,000 households which mostly was concentrated on the income of households to assess living standard for national, regional and provincial levels and one smaller sample of 30,000 households with information about both income and expenditures to evaluate intensive living standard at central and provincial levels. For the purpose of the creation of the RIGA income aggregate and subsequent analyses, the smaller sample of 30,000 households is used.

A total of 30,000 were surveyed for the purpose of creating the smaller more detailed sample, however 29,530 households actually have complete surveys in the raw data. There are 22,621 rural households and 6,909 urban households in the dataset. In the original data, URBAN02 is the variable distinguishing urban from rural households.³ For the calculation of the income aggregate, URBAN02 is renamed to URBAN in order to use the same variable name across different surveys.

The households surveyed for 2002 were not selected proportional to the population (over-sampling of certain areas was deliberately undertaken), therefore population weights are

¹ The information provided in this document relies substantially upon the General Introduction Document, provided with the 2002 VHLSS 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/lsm>.

³ The definition of rural is not documented in the survey.

necessary to apply when generating any summary statistics from the income aggregate and other raw or constructed variables.

The average household size in Vietnam in 2002 was 5.1 persons.⁴ All money amounts are in the local currency, Dong. In 2001, the official exchange rate was 15.403 Dong = 1.0 USD.⁵

The income aggregates are calculated at the household level and all aggregates are annualized.

Household identifier

The household identifier was not identically laid out in all the household-level modules. In particular; the two datasets needed to create the household roster, the *hhexp02.dta* and the *tongcong.dta*, established the household identifier in a different way. Hence, to create the household roster, it was necessary to modify this and establish a unique way to identify households. The following table illustrates the discrepancies between the two files:

hhexp02.dta	tongcong.dta
tin	xa
huyen	
xa	
diaban	hoso
hoso	

To solve the problem, we renamed xa “xat” and hoso “hosot” in *hhexp02.dta*. We then generated the new xa and hoso variables:

- xa = combining (tin huyen xa)
- hoso = combining (diaban hoso)

After performing this correction in the data, there was still a further inconsistency. For households with “hosold” between 299 and 400, the hoso identifier had been transformed further: we had to subtract 200.

A third inconsistency was found: **xa=109415 with hosold between 100 and 121** suffered a specific transformation that was corrected in the do file.

All these transformations were saved in *hhexpe02new.dta*, in order to leave *hhexpe02.dta* without modifications along with the rest of the raw data.

After this reconstruction of the two variables that create the unique household identifier, the various household-level modules of the VHLSS survey can be linked using **XA** and **HOSO**. These variables grouped together create a variable HH, which is used as a unique household identifier for consistency across countries within the study. The household and community data can be merged by using the variable **XA**.

⁴ RIGA calculation from the VHLSS dataset.

⁵ The exchange rate used is the end of period exchange rate for 2002 obtained from the IMF International Financial Statistics database.

The following Stata syntax was used for making the necessary transformations for the construction of the household identifier described in this section:

```
***
use "S:\RIGA\Vietnam02\RawData\Household\hhxpe02.dta", clear
sort xa hoso

isid tinh huyen xa hoso

*SOLVING PROBLEM N.1: in all the other raw datasets, the hh identifier is
"xa hoso", where:
  *xa is a combination of tinh huyen xa
  *hoso is a combination of diaban and hoso

rename xa xaold
rename hoso hosold

*Generate xa that is equal to the xa variable found in all the other raw
datasets
gen xat=string(tinh)+string(huyen, "%02.0f")+string(xaold, "%02.0f")
destring xat, generate(xa)

*Generate hoso that is equal to the hoso variable found in all the other
raw datasets
destring diaban, generate(diaban1)
gen long hoso=(diaban1*1000+hosold)

*SOLVING PROBLEM N.2: when 299<hosold<400, the hoso identifier has been
transformed subtracting 200
replace hoso=(diaban1*1000+hosold)-200 if hosold>299&hosold<400

*SOLVING PROBLEM N.3: xa=109415 suffered a specific transformation
replace hoso=((diaban1+10)*1000+hosold)-100 if
xa==1090415&(hosold>100&hosold<121)

isid xa hoso
count
save ../Tempdata\hhxpe02new.dta, replace
***
```

Comments

- 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**. **All variables included in the aggregate income variable are net of costs unless otherwise noted.**
- 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.
- In the Crop Production section, the reference period is the previous 12 months. Two total crop income variables are created: *cropincome1* and *cropincome2*. *Cropincome1* includes estimates of own crop consumption based on the agricultural production module of the household questionnaire. *Cropincome2* includes estimates of own crop consumption based on the food expenditure section of the questionnaire in which household consumption was recorded and updated periodically over a given time frame.
- In the Food Expenses and Home Production section, the survey disaggregated between Holiday (Tet) and Regular (Non-Holiday) expenditures and consumption.

- The reference period for the consumption of own produce (from the food expenditure section of the questionnaire, corresponding to *Cropincome2*) is the last 12 months. However, for holiday expenses and consumption, these were assumed for the calculation of the income aggregate, to only cover two weeks of the year.
- For the Livestock, Other Income, Rentagric and Transfers sections, the reference period is the previous 12 months.
- Other Income includes income from non-labor, nonfarm sources.
- Rentagric calculates gross income from the rented and sharecropped agricultural land. The expenditures taken into account are the expenditures a household makes on for land it cultivates and owns (i.e. land for which the household must pay land rights to a public authority). Payments made for land that the household does not own are not owned is not considered in this estimate.
- For Transfer income the estimate calculated is of gross receipts. Households are not asked about transfer outflows therefore a net estimate cannot be calculated. Remittances income includes gifts which results in higher participation rates when expected.
- In the Wage employment section, the reference period is the specified by the respondent in the questionnaire for the primary and secondary jobs of the past 7 days and the primary and secondary jobs held in the last 12 months. All household members 6 years and older were interviewed.
- The survey establishes its own industrial and occupational classifications, which, for the income aggregate are modified to follow the United Nations International Standard Classification of Occupations (ISCO-88) codes. Given these standards, the employment 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.
- Earnings from wage employment include all in-cash and in-kind benefits in cash received from the employer and are net of contributions to social security and health insurance.
- The Self Employment (Selfemp) section accounted for income from nonfarm enterprises owned by the household. For expenditures, the reference period was chose by the respondent and for revenues, the reference period is the last 12 months. Income from nonfarm enterprises is disaggregated by industry (following the ISCO-88 Codes) in order to convey the diversity of household activities.
- 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, 25 households are dropped from this survey.
- 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. This table can also be viewed, with the individual components disaggregated, in Microsoft Excel from the file *Income.xls*.

Table 1

Program Name (Stata “do” file)	Input Data Files	Output Data Files	Output Variables
Sample	tongcong.dta hhexpe02.dta	Sample.dta	
Prices	muc5b21_23.dta muc6a2.dta	price_prod_tinh.dta price_prod_urban02.dta price_prod_unit.dta price_purch_xa.dta price_purch_tinh.dta price_purch_urban02.dta price_purch_unit.dta	price_prod_xa price_prod_tinh price_prod_urban02 price_prod_unit price_purch_xa price_purch_tinh price_purch_urban02 price_purch_unit
Food	muc6a1.dta muc6a2.dta price_prod_xa.dta price_prod_tinh.dta price_prod_urban02.dta price_prod_unit.dta price_purch_xa.dta price_purch_tinh.dta price_purch_urban02.dta price_purch_unit.dta Sample.dta	Food_Tet.dta Food.dta	foodtetpurchimp foodtetownimp foodtetown_cropimp foodtetown_lvstkip fooddown_cropimp fooddown_lvstkip foodpurchimp foodownimp
Aquainc	muc5b62.dta Sample.dta	Aquacinc.dta	aquaexp aquainc aquanetimp
Cropincome1	muc5b25.dta	cropexp.dta	cropexpimp
Cropincome2	muc5b21_23.dta muc5b2ho.dta	cropinc.dta cropbyprinc.dta	cropsoldimp cropownimp cropbyprincimp
Cropincome3	muc5b4.dta muc5b5.dta cropinc.dta cropbyprinc.dta farmservicesnet.dta sylviculturenet.dta Food.dta Food_Tet.dta	farmservicesnet.dta sylviculturenet.dta cropincome1imp cropincome2imp	farmservicesnetimp sylviculturenetimp cropincome1imp cropincome2imp
Employment	muc5a.dta muc3.dta SAMPLE.DTA	<i>Employmainjob.dta</i> <i>employsecnjob.dta</i> Employment.dta	wge, wgem, wgimp (by industry and skill level)

Livestock1	muc5b32.dta muc5b31.dta muc5b5.dta Food.dta Food_Tet.dta SAMPLE.DTA	<i>livstexp.dta</i> <i>livstinc.dta</i> hunting.dta Livestock.dta	livestockexpimp livestockincimp livestockbyprincimp livestockownconsimp huntingnetimp livstockimp
Otherincome	muc5d.dta SAMPLE.DTA	Otherincome.dta	otherincimp nonfarmrntimp
Rentagric	muc5b1.dta SAMPLE.DTA	Rentagric.dta	farmrntimp
Selfemp	muc5c2.dta muc5c1.dta SAMPLE.DTA	selfemp1.dta selfemp2.dta selfemp3.dta selfemp4.dta Selfemp.dta	selfempimp selfempimp selfempimp selfempimp selfimp1 selfimp2 selfimp3 selfimp4 selfimp5 selfimp6 selfimp7 selfimp8 selfimp9 selfimp10
Transfers	muc5d.dta SAMPLE.DTA	Transfers.dta	transfersgrossimp privtransferimp pensionsimp socialtransimp pubtransferimp
Aggregateincome	Sample.dta Aquainc.dta Cropincome.dta Employment.dta Livestock.dta Otherincome.dta Rentagric.dta Selfemp.dta Transfers.dta	Income.dta	agr_wge nonagr_wge crop1 livestock selfemp transfers other

Table 2

Vietnam 2002		22,330 Rural HH Observations		Rural, Weighted, Dongs				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	2,684	11.75%	5,719.59	672.05	6.18%	6.59%	374	44
nonagr_wge	Wage Employment- Nonfarm	6,130	28.73%	8,491.93	2,440.03	16.16%	23.93%	556	160
crop1	Crop Production	18,644	83.18%	2,919.28	2,428.36	43.23%	23.81%	191	159
livestock	Livestock Production	17,537	78.21%	1,021.24	798.68	5.75%	7.83%	67	52
selfemp	Non-ag Self Employment	7,761	35.33%	7,003.63	2,474.32	16.29%	24.26%	458	162
transfer	Total Transfers	18,316	83.50%	1,387.44	1,158.57	10.49%	11.36%	91	76
other	Other Income Sources	5,530	24.87%	906.03	225.35	1.91%	2.21%	59	15
totincome1	Total Household Income	22,324	99.98%	10,199.41	10,197.35	100.00%	100.00%	668	667

Percent Rural (Weighted)	73.19%
Dong/USD (2002)	15.280

Notes:

1. Source data: 2002 Living Standards Survey (VLSS).
2. Exchange rate used is obtained 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 = crop harvested minus crop sold. For crop2, owncons is based on the consumption Holiday Expenses and Regular (Food) Expenses sections of the 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 Transfers, which are gross receipts)

Table 3

Vietnam 2002		22,330 Rural HH Observations		Rural, Weighted, Dongs				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	2,684	11.75%	5,719.59	672.05	6.32%	7.05%	374	44
nonagr_wge	Wage Employment- Nonfarm	6,130	28.73%	8,491.93	2,440.03	17.25%	25.60%	556	160
crop2	Crop Production	19,645	87.83%	2,007.74	1,763.43	40.46%	18.50%	458	115
livestock	Livestock Production	17,537	78.21%	1,021.24	798.68	5.37%	8.38%	373	52
selfemp	Non-ag Self Employment	7,761	35.33%	7,003.63	2,474.32	17.31%	25.96%	353	162
transfer	Total Transfers	18,316	83.50%	1,387.44	1,158.57	11.23%	12.15%	177	76
other	Other Income Sources	5,530	24.87%	906.03	225.35	2.07%	2.36%	125	15
totincome2	Total Household Income	22,326	99.99%	9,533.79	9,532.42	100.00%	100.00%	624	624

Percent Rural (Weighted)	73.16%
Dong/USD (2002)	15.280

Notes:

1. Source data: 2002 Living Standards Survey (VLSS).
2. Exchange rate used is obtained 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 = crop harvested minus crop sold. For crop2, owncons is based on the consumption Holiday Expenses and Regular (Food) Expenses sections of the 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 Transfers, which are gross receipts)