

SECOND Supplemental Survey to the 1999/2000 Post Harvest Survey – Zambia
Food Security Research Project – Michigan State University
Data Documentation – Updated June 2010

The Food Security Research Project from Michigan State University conducted the second supplemental survey to the 1999/2000 Post Harvest Survey in June/July 2004. The Central Statistical Office (CSO) provided the personnel for the survey. The sample was all households that completed the First supplemental survey conducted in 2001. From this sample of 6922 households, 5419 households completed the survey. However, of those 5419 households, 77 did not match the household from the previous sample. 5344 households appear to match the previous survey (first supplemental survey).

Response status

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Completed	5419	78.3	78.3	78.3
Refusal	14	.2	.2	78.5
Moved out of SEA	707	10.2	10.2	88.7
Hh dissolved	390	5.6	5.6	94.3
Non-contact	362	5.2	5.2	99.6
HH skipped & not interviewed	30	.4	.4	100.0
Total	6922	100.0	100.0	

Panel household identification:

This dataset is part of a panel study. In the file, “id04.sav” a variable has been added, called “paneltype”. 5 different types of households have been identified:

paneltype Panel classification of household

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 matches all 3 years	4288	79.1	79.1	79.1
1 2000 only	0	0	0	79.1
2 2000 & 2004 only	1070	19.7	19.7	98.9
3 2004 & 2008 only	52	1.0	1.0	99.8
5 none of years match	9	.2	.2	100.0
Total	5419	100.0	100.0	

A value of 0 identifies the household as having been interviewed all three years of the panel study. A value of 2 indicates the household was interviewed only in 2000 and 2004. A value of 3 indicates that the household interviewed in 2001 did not match the 2004 household; however this household does match the household interviewed in 2008. A value of 5 indicates the household was interviewed but is not the same household as was interviewed in 2001.

Recommended file structure:

C:\.....\Zambia\

hh04\

anal - data files and syntax created from analysis

data - working data files and syntax used to convert from archive format to SPSS

docs - documentation for this survey

lookup – lookup files are stored in this directory

newvars –

demog – adult equivalent values and size of household

income - standardized computed variables for crop income, off-farm income, livestock income, productive assets income, landholdings, retained caloric values, fertilizer costs

tmp - the directory where temporary files can be stored - to be erased when finished

Data files descriptions (\.....\Zambia\hh04\data):

Household level variables - prov, **dist**, **csa**, **sea**, **hh**

Conventions:

If a value in a variable was missing and the question was suppose to have been asked, a user missing value of -8 with a value label of 'not collected' was entered. This value (-8) was declared as "user missing".

Extremes were run on all continuous variables. If a value was considered an extreme, that case was verified by examining the questionnaire. If the value in the data file was the same as the value on the questionnaire, the questionnaire was further examined to determine if the enumerator had recorded the value incorrectly. Where it was possible, data were corrected based on information from other variables.

A "comment" variable was added to each file. The people responsible for cleaning the files were to write a comment in this variable to explain any unusual values for the case. Not all of the people cleaning the data followed this instruction.

Data file description:

Survey Section	File Name	Key Variables	Variable names start with ...	Number of cases
Front Page	allsurid04.sav – all surveyed households	dist csa sea hh	mnemonics	6,922
	id04.sav – only households that completed the survey	dist csa sea hh	mnemonics	5,419 – see “paneltype”
	id04.sav notes: Weight variables: 2 variables are included in this file. The first, wgt_pan, should be used for panel analysis. The second variable, wgt_pop, should be used if the data are to be analyzed separately and an estimate made for the total population.			
Household level questions	hh04.sav – data from many different sections. Sections 1, 2, 3 (land use), 4 (cassava stocks and sales), 5 (maize grain processing), 6 (fertilizer acquisition), 7 (agricultural loans), 8 (ag commodity prices), 11 (perceptions/attitudes)	dist csa sea hh	hh...	5,419
	<u>hh04.sav notes</u>			
Section 1 – Demography	adults04.sav – adult family member	dist csa sea hh mem	da..	19,116
	<u>Adults04.sav notes</u> : Only members who were in the previous supplemental survey will be in this file. Note that the head of the household could be in the additional adults file if there is no head for the household in this file.			
	adults_addit04.sav – additional adults 12 years and older not listed in 2001	dist csa sea hh mem	ad...	6,831
	Adults_addit04.sav notes: Member numbers start at 31. Children who have become 12 years and older who were members of the household in 2001 will be listed here as well as additional adults who have joined the household.			
	child04.sav – children younger than 12	dist csa sea hh mem	dc...	11,828
	Child04.sav notes: Member numbers start at 51			
Previous Deaths	prevDeaths04.sav – deaths within hh since April 2001	dist csa sea hh pdmem	pd...	905
Section 2 – Off-farm income remittances	salwage04.sav – salaried employment or informal wage labour activities	dist csa sea hh wact w01	w...	1,448
	SalWage04.sav notes:			
	bus_activ04.sav – formal or informal business activities	dist csa sea hh bact b01	b...	2,065

Survey Section	File Name	Key Variables	Variable names start with ...	Number of cases
	<u>bus_activ04.sav</u> Notes: This file has duplicate cases. These are valid duplicate cases – the data in the rest of the variables are different – different incomes-expenditures were received for the same activity in the same location.			
	remit04.sav – remittances given and received	dist csa sea hh rmem	r...	2,944
	Remit04.sav notes:			
	gg_foodaid04.sav – gifts, grants and/or food aid given and received	dist csa sea hh gnum gg	g...	6,380
	<u>gg_foodaid04.sav</u> notes: The two tables in the survey instrument were combined into one data file. The variable “gg” identifies the case with respect to whether the commodity was sent by the household (1) or received (2).			
Section 3 – Farm Land and Use	field04.sav – details about each field. Fallow land area added. A variable was added to indicate if the fallow land was imputed	dist csa sea hh field	f...	17,070
	<u>field04.sav</u> notes: The area for fallow fields was not recorded. Students were assigned to review the surveys that indicated they had fallow fields and record the hectares indicated on the sketch. If no area was recorded a -8 – not collected was recorded in the file. Some missing areas were imputed. The procedure used was: If a household reported that they had fallow land but no recording on sketch: Case 1: If land cultivated in 04 is less than land cultivated in 2000 plus fallow - the difference between the two was the imputed fallow land in 2004. Essential, I assumed no change in land size. Case 2: If land cultivated in 04 was greater than land cultivated in 2000 plus fallow in 2000 - then I just took it that the household did not have any fallow land.			
	crop04.sav – details about each crop planted within each field – seed information, planting time and harvested	dist csa sea hh field crop	c...	13,442
	<u>crop04.sav</u> notes: These questions were asked with respect to whether the field was monocropped (one table) or intercropped (another table). After the data were entered, the two files were merged into one crop file. The field04.sav contains a variable to indicate if the field was monocropped or intercropped.			
Section 4 – Crop Stocks and Sales	sold_stocks04.sav – Sales of crops, quantity in storage	dist csa sea hh crop	s...	11,894
	Sold_stocks04.sav notes:			
	cassava04.sav – Cassava production – area planted, variety planted, quantity harvested	dist csa sea hh field	cs...	5,103
	<u>Cassava04.sav</u> notes: sales information is in the hh04.sav file			
	veg_fruit04.sav – vegetables, fruits and green maize sold	dist csa sea hh vfcode	vf...	1,974

Survey Section	File Name	Key Variables	Variable names start with ...	Number of cases
	<u>veg_fruit04.sav notes:</u>			
Section 5 – Food purchases and processing	purch04.sav – food purchases over two periods (May to Dec 2003) (Jan 2004-April 2004)	dist csa sea hh purch	p...	4,175
	<u>purch04.sav notes:</u>			
	food_pay04.sav – food obtained as payment for labor over two periods (May to Dec 2003) (Jan 2004-April 2004)	dist csa sea hh lpurch	lp...	1,116
	prod_maize04.sav – products obtained from maize grain	dist csa sea hh prod	pr...	19,332
Section 6 – Fertilizer acquisition	fertacq04.sav – acquisition of fertilizer through transaction channels	dist csa sea hh fertch	fr...	14,120
Section 8 – Advice Provision	service04.sav – advice received from organizations	dist csa sea hh srcode	sr...	8,728
Section 9 – Assets	asset04.sav – production assets / implements	dist csa sea hh asset	ast...	102,942
Section 10 – Livestock, Poultry and Fish Farming	livestock04.sav – Cattle, pigs, goats, sheep, donkeys, poultry – owned and sold	dist csa sea hh lstock	ls...	12,216
	milk_egg_fish04.sav – production over last 12 months and price	dist csa sea hh mef	lv...	1.292
	<u>milk_egg_fish04.sav notes:</u>			

Lookup tables

C:\...\Zambia\hh04\lookup

Type of data	File name	File to be used with	Key variables	Number of cases	Comments
Crop quantity conversion to kgs	Cropconv04.sav	crop04.sav	crop, unit	259	Converts all harvested/sold crop units to kgs.
Land conversion to hectares	ha_conver.sav	field.sav	f03	4	Converts area to hectares
Price per kg for a	crop_prices_04.sav	crop04.sav	prov, crop	190	Median price per kg for the crop at the provincial level – kwacha.

Type of data	File name	File to be used with	Key variables	Number of cases	Comments
specific crop					Generated using 'crop_prices.SPS'
Livestock median price	livestock_price_04.sav	livestock04.sav	prov, lstock	135	Unit price for livestock – Kwacha. Generated using 'livestock_prices.SPS'
Livestock product prices	livprod_price04.sav	milk_egg_fish04.sav	prov, mef	27	Unit price for livestock products. Generated using 'livestock_products_prices.SPS'
Conversion of purchases to kgs	purconv.sav	purch04.sav	purch, unit	139	Converts all purchased units to kgs
Purchases median price	purch_prices_04.sav	purch04.sav	prov, purch	93	Median price of purchased product. Generated using 'purch_price.SPS'
Seed quantity conversion to kgs	seedconv.sav	crop04.sav	crop, cm07	122	Converts various units to kgs
Food aid quantities standardized to kgs	foodaidconv.sav	gg_foodaid04.sav	gnum, g03	218	Converts various units to kgs
Median kg price for products	foodaid_prices_04.sav	gg_foodaid04.sav	prov, gnum	161	Median price per kg for various products. Generated using 'foodaid_price.SPS'
Standardize cassava chips units to kgs	cassava_syntax_conv.SPS	hh04.sav			Contains syntax to convert units of dried cassava to kgs to calculate total kgs stored of cassava.

New Computed Variables

C:\...\Zambia\hh04\newvars

Type of data	File name	Key variables	Number of cases	Variables	Syntax File
Subdirectory “demog”					
Adult equivalents and size of household	adult_equivalents_2004.sav	dist, csa, sea, hh	5,419	Aehh04 – Full time adult equivalents hhsz04 – household size age15_59 – number of prime age members educhead – level of education of the head edu_max – maximum level of education in the household malehead – Male headed household femalehead – Female headed household	adult_equivalents_2004.sps – see note at end of documentation regarding method used to compute adult equivalents
Subdirectory “income”					

Crop income computation	crop_income_sales_ss02.sav	dist, csa, sea, hh	5,419	gvharv_p: gross value of harvest for crops found in all years for panel, gvharv: gross values of all crops (same as gvharv_p for this year) gvsale_p: gross value of sales for crops found in all years for panel gvsale: gross value of sales cerealK, tubersK, com_crpK, otherK – categories of crops – value of production cerealS tuberS, comcrpS, otherS – categories of crops value of sales vegsales – no production information, only sales. Also included are variables of harvest and sales by each crop	crop_income_sales_ss02.sps
Off farm income	off_farm_income_ss02.sav	dist, csa, sea, hh	5,419	Net_off_farm: total net off farm income business costs subtracted (tot_work+ tot_bus+ remit_in) Gross_off_farm: total gross off farm income where business expenses are not subtracted (tot_work+ gv_busincome + remit_in) Detailed variables are: tot_bus(net business income), gv_busincome, gv_buscost, remit_in, tot_work, non_ag, ag_wage, agric_sf (small farm ag wages) agric_lg (commercial farm ag wages)	off_farm_income_ss02.SPS
Live animal valuation	livestock_income_ss02.sav	dist, csa, sea, hh	5,419	livinc04 – total income for livestock and livestock products (milk and eggs) Fishinc – income from fishfarming	livestock_income_ss02.SPS
Productive asset valuation	productive_assets_ss02.sav	dist, csa, sea, hh	5,419	past0104: draft animals and equipment value (Kwacha) for assets matching 2001 asst0104: assets (without animals) matching 2001. Prodass: all animal and equipment assets for 2004 Assetall: assets (without animals) for 2004	Productive_assets_ss02.SPS
Land holdings	land_variables_ss02.sav	dist csa sea hh	5,419	Cropland: area under crops Landholdpanel: area under crops plus fallow land imputed Landholdpanel_fa_noimp: area under crops plus actual fallow land – no imputation Includes variables for area planted for each crop. Area was not asked for garden, virgin land, or	Land_variables_ss02.sps Area under fallow was not collected in the data file. Students examined the questionnaires and added area specific under fallow. Where no area was specified, the area was

				orchards.	imputed, see documentation above.
Fertilizer costs	hh_fertilizer_costs_kgs_ss02.sav	Dist csa sea hh	5,419	Tot_fert_kg: total kgs of fertilizer used for all crops Tot_fert_cost: total cost of fertilizer used for all crops Tot_fert_kg_p: total kgs of fertilizer used for crops asked about for panel analysis Tot_fert_cost_p: total cost of fertilizer used for crops asked about for panel analysis Fertilizer was applied at the field level so it was not possible to assign quantity of fertilizer used for each of the crops. Transport costs were also calculated.	hh_Fertilizer costs_ss02.sps Median price for basal and top dressing was calculated from the household level questions and used to value fertilizer. See definition of how costs were apportioned below based on source of acquisition of fertilizer.*
Calories retained	hh_level_crop_detail_calories_ss02.sav	dist csa sea hh	5,419	Individual variables for each crop computing calories retained	hh_level_crop_detail_calories_ss02.sps
Income (gross and net)	income_gross_net_observed_ss02.sav	dist csa sea hh	5,419	Gross_income_panel = gvharv_p+ vegsales+ tot_livs+ livprodprod+ gross_off_farm Net_income_panel = gvharv_p+ vegsales+ tot_livs+ livprodprod+ net_off_farm – tot_fert_cost_p Gross_income_allhhs and net_income_allhhs used gvharv and tot_fert_cost in the calculations.	Income_gross_net_observed_ss02.sps

***Definition of how fertilizer costs were apportioned (Bill Burke, 2010):**

We would like to find fertilizer costs (FC) in 2004 and 2008, where:

$$FC = q_c^u p_c + q_F^u p_F + q_g^u p_g + q_r^u p_r \quad (1)$$

where q_c^u = quantity of commercial fertilizer used

q_F^u = quantity of FSP fertilizer used

q_g^u = quantity of free (gift) fertilizer used

q_r^u = quantity of carry over (residual) fertilizer used

p_s = price of fertilizer from source s

The problem is that we don't observe quantities *used* at the source level, and prices are not equal across source. We can deal with prices because p_c is observed, we know $p_F = p_c/2$, $p_g = 0$, and we can assume $p_r = p_c$.¹ Quantity used, on the other hand, is not observed by source, but we do observe total quantity used ($Q^u = q_c^u + q_r^u + q_F^u + q_g^u$) and we observe quantity *acquired* from commercial purchases, FSP, and gifts, or q_c^a , q_F^a , and q_g^a respectively. The total amount of fertilizer acquired in a given year is obviously ($Q^a = q_c^a + q_F^a + q_g^a$), so we can calculate the carry over quantity used:

$$q_r^u = Q^u - Q^a \quad \text{if } Q^u > Q^a, \text{ and } 0 \text{ otherwise.} \quad (2)$$

Then, if we assume a constant ratio of quantity acquired to quantity used, we can come up with values for q_c^u , q_F^u , and q_g^u as follows (for $s = c, F, g$):

$$\tilde{q}_s^u = \frac{q_s^a}{Q^a} (Q^u - q_r^u) \quad (3)$$

Note that if $Q^u \geq Q^a$, then $Q^u - q_r^u = Q^a$, and $\tilde{q}_s^u = q_s^a$. In other words, if used is greater than acquired, we act as if all acquisitions were used, and will price accordingly. If $Q^u < Q^a$, then we will have $\tilde{q}_s^u < q_s^a$, and we assume the following "constant ratio":

$$\frac{q_s^u}{q_c^u + q_F^u + q_g^u} = \frac{q_s^a}{q_c^a + q_F^a + q_g^a} \quad (4)$$

In any case, the equality will hold that ($Q^u = \tilde{q}_c^u + \tilde{q}_F^u + \tilde{q}_g^u + q_r^u$), and we can calculate total costs as:

$$\tilde{FC} = \tilde{q}_c^u p_c + \tilde{q}_F^u p_F + \tilde{q}_g^u p_g + q_r^u p_r \quad (5)$$

¹ Alternatively, it may be more consistent to assume some other value for the price of carry over fertilizer based on the households share of acquired fertilizer by source, but for now let's not.

Documents:

SS04_DataFileNotes.pdf. Food Security Research Project, 2005. Second Supplemental Survey Data Documentation Notes. Michigan State University, October, updated June 2010.

SS04_Orig_Questionnaire.pdf . Republic of Zambia, 2004 Second Supplemental Survey to the 1999/2000 Post-Harvest Survey (for small and medium scale holdings) **Original Survey Instrument.** Republic of Zambia, May.

SS04_QUESTchanges_for_enumerators.pdf. Republic of Zambia, 2004 Checklist of most important changes to the questionnaire and other key issues - Second Supplemental Survey to the 1999/2000 Post-Harvest Survey (for small and medium scale holdings). Republic of Zambia, May.

SS04_Suppsurvey Enumerator manual.pdf. Republic of Zambia, 2004 Second Supplemental Survey to the 1999/2000 Post Harvest Survey (for small and medium scale holdings) **Interviewers' Instruction Manual.** Food Security Research Project, Michigan State University and Republic of Zambia, May.

SS04_Synthetic_Questionnaire.pdf. Republic of Zambia, 2004 Second Supplemental Survey to the 1999/2000 Post-Harvest Survey (for small and medium scale holdings) **Synthetic Survey Instrument.** Republic of Zambia, May.

Questions regarding data validity:

If you find data that you wish to have verified with the original survey instruments, please follow these instructions:

Send a copy of the syntax file that will produce a listing of data in question to Margaret Beaver at beaverm@msu.edu. She will either make corrections to the data and then issue new copies of the data to everyone or send a syntax file to everyone to run to make the corrections. It is important that you send a syntax file that will list the questionable data rather than just describing the problem.