Methodology Employed in the Calculation of Aggregate Expenditure

EXTERNAL IMPACT EVALUATION OF THE MILLENNIUM VILLAGES PROJECT, NORTHERN GHANA

Date: **February 2014**

Submitted by Itad
In association with:

itad
Results in development

IDS Institute of Development Studies

Participatory Development Associates Ltd.

Introduction

Expenditure was calculated using the same methodology employed by GSS for the calculation of poverty measures from GLSS3, GLSS4, and GLSS5. This methodology was largely designed by McKay and Coulombe and building on previous work by Glewwe and Twumba for the poverty analysis of GLSS1 and GLSS2. In some cases we could not use either of the two methodologies because our questionnaire differs from those employed by GSS in some respects and hence we referred to the paper by Deaton and Zaidi as the 'best practice' in the calculation of consumption figures.

The motivation for using the same methodology employed by the GSS is that our questionnaires are fairly similar to those employed by the GSS and secondly because we want to obtain poverty estimates that are as comparable as possible to those presented in the official country statistics.

Outliers

The procedure for dealing with the outliers is the following:

- Outliers are found by taking the natural logs of expenditure on each item and classifying as outliers all those observations whose values are three standard deviations above the mean.
- Outliers are replaced by the average expenditure for that item (after removing the outlier) in each of the four regions separately (Builsa MV, Builsa CV, West Mamprusi MV, and West Mamprusi CV – this follows GSS practice of replacing outliers within three macro-regions with urban/rural split).
- In the case of food items, expenditures are first divided by household size in order not to classify as outliers large food expenditures made by large households.
- This operation is conducted two times sequentially for each variable.

Missing values

Some households report purchasing items but are not able to report the amount spent. We interpreted these cases as missing values rather than zero expenditures.

In the case of non-food expenditures we replaced the missing expenditures by the mean household expenditure for that item in each of the four regions separately. We did this after replacing outliers using the methodology above in order to obtain a representative mean.

In the case of food expenditure, we adopted the following approach:

- When value of purchase is missing but quantities and units of measurements are available: we applied survey prices (calculated in the way illustrated below) to the quantities purchased.
- When value of purchase and the quantity purchased are missing but the unit of measurement is reported: we set the quantity purchased to one unit of the reported unit of measurement.

- When quantity consumed from own production is missing but the unit of measurement is reported: we set the quantity consumed to one unit of the reported unit of measurement.

Housing

Housing conditions in the study area are generally poor. Only five households report paying any rent and estimating the rental values of property for all households using these observations (as performed by GSS) is impossible. This means that the service value of houses is not included in the expenditure figures.

Household size

Following GSS practice we coded as non-household member any individual who was away from the household for six or more months over the previous 12 months (equivalent to being away for at least five months in question a014) UNLESS the individual is the head of household or a child less than 9 months old.

Durable goods

We calculated the user value of durable goods. Following Deaton and Zaidi this can be obtained as the annual cost of holding a particular asset:

$$uv = SP_{t1}(\partial + r - i)$$

Where:
SP is the current value of the asset (asset stock (S) times its price (P))
Delta is the annual depreciation rate of the item
r is the nominal annual interest rate
i is the annual inflation rate

The user value can be thought of as the annual cost required for replacing the item inclusive of the opportunity cost of not investing money in an alternative use (captured by the real interest rate).

SP is known from the data and estimated by the respondent directly. The real interest rate is an arbitrary value assumed to be 0.02. Finally the depreciation rate is calculated from the data in the following way:

$$\partial + i = 1 - \left(\frac{SP_{t1}}{SP_{t0}}\right)^{1/T}$$

This quantity is calculated for each asset in the dataset using information on present value, value at purchase, and the number of years it has been used. In order to remove outliers, a median depreciation rate specific to each asset is calculated from the sample and used to estimate the user value of the item. These are the depreciation rates calculated from the data.

Q270.NM. Asset Item	Mean	Std. Dev.	Freq.
Animal-drawn cart (wheels)	.11583459	0	239
Any other motor vehicle (car	.16077702	0	13
Bed	.19340217	0	846
Bicycle	.21610561	0	1562
Bucket	.24611455	0	1751
Camera	.16698791	0	5
Chair	.197869	0	779
Computer	.21610561	0	3
Engine / generator (includes	.23205692	0	42
Kerosene Lamp	.21610563	0	409
Mobile / cellular phone	.27005354	0	1024
Motorcycle or scooter	.24937615	0	219
Radio	.24611455	0	964
Refrigerator	.14821093	0	13
Sofa (pieces)	.18471698	0	187
Table	.18519226	0	1074
Tape Recorder / Cassette Play	.21007282	0	346
Television	.2129083	0	66
Three stone stove/fire	.1661931	0	401
Torch / Lamp	.37390429	0	1935
Wall Clock	.24937615	0	93
Wardrobe	.1761689	0	21
Watch	.31129476	0	250
Total	.24514145	.06478336	12242

High frequency expenditure

Outliers and missing values are replaced following the procedure outlined above.

Low frequency expenditure

Outliers and missing values are replaced following the procedure outlined above.

The following items were excluded from the computation because their user value was already obtained from the asset ownership lists: mobile phones, furniture, motorised vehicles, non-motorised vehicles. Agricultural tools were excluded from the durables list because they are not consumption items.

The following items were retained in the form of user values: home repairs, power equipment, PCs, appliances. Their user value was obtained by multiplying their purchase value by the average annual depreciation values calculated for household assets (0.25) and then dividing the figure by two, because on average purchases over the last year were used only for six months.

We also excluded the payment of tax and cash losses but we included transfers made to other households and payments for funerals, dowries, and marriages.

Food consumption

Outliers and missing values are replaced following the procedure outlined above. It is important to bear in mind that in the case of food items all expenditures are divided by household size before identifying the outliers.

Households report the value of each purchase and the calculation of food expenditures is straightforward. The valuation of the consumption of food produced by the household is more complicated because it needs to be imputed. The best practice in this case is employing farm-

gate prices at the household, locality or area level from the income section of the questionnaire, because these prices better reflect the quality and the value of the item consumed.

Unfortunately, only a few farm-gate prices can be effectively calculated from the income section of the questionnaire. Everything from the number of reported cases, the variety of items consumed, and the cases reported per items suggests that there is under-reporting of own-consumption and production in the agricultural production section of the questionnaire.

Reported cases of own-produced foods from the expenditure section.

Q278.NM. Consumption Products Name	Freq.	Per cent	Cum.
Alcohol	22	0.10	0.10
Avocado	6	0.03	0.13
Bambaram beans	899	4.03	4.15
Beans	1,130	5.06	9.21
Berries	505	2.26	11.47
Bread	1	0.00	11.48
Butter	131	0.59	12.06
Cassava	49	0.22	12.28
Cooking Oil	185	0.83	13.11
Cooking fat and margarine	2	0.01	13.12
Cream	1	0.00	13.12
Dawadawa Fruits	1,154	5.17	18.29
Eggs	678	3.04	21.33
Fish	237	1.06	22.39
Grapes	133	0.60	22.98
Green Maize	1,187	5.32	28.30
Groundnuts	1,270	5.69	33.99
Guavas	8	0.04	34.02
Irish Potatoes	29	0.13	34.15
Kernef-bra/bito	1,330	5.96	40.11
Kola Nut	15	0.07	40.18
Lettuce	4	0.02	40.19
Maize Grain(kernels/seeds)	1,460	6.54	46.73
Maize Meal/flour(after milling or grind	1,239	5.55	52.28
Mangoes	382	1.71	53.99
Meat	501	2.24	56.23
Milk(any)	226	1.01	57.25
Millet	1,449	6.49	63.73
Okra	988	4.42	68.16
Onions	22	0.10	68.26
Oranges	15	0.07	68.32
Other Fruits	113	0.51	68.83
Other Nuts/Seeds	20	0.09	68.92
Other flours	34	0.15	69.07
Other grains	60	0.27	69.34
Other green leafy vegetables	577	2.58	71.92
Other legumes/Pulses	9	0.04	71.96
Other tubers/roots	20	0.09	72.05
Other vegetables	309	1.38	73.44
Papaya(PawPaw)	36	0.16	73.60
Peas/Cowpeas	141	0.63	74.23
Pigeon beans	16	0.07	74.30
Plantains	2	0.01	74.31
Poultry	675	3.02	77.33
Pumpkin	102	0.46	77.79
Rice	878	3.93	81.72
Salt	7	0.03	81.75
Shea Fruits	1,272	5.70	87.45
Snacks	1	0.00	87.45
Sorghum	1,117	5.00	92.45
Soya beans	166	0.74	93.20
Sugar	7	0.03	93.23
Sumsume/Nasulga	598	2.68	95.91
Sweet Bananas	1	0.00	95.91

Sweet Potatoes	248	1.11	97.02
Tea	24	0.11	97.13
Tobacco/Cigarettes	41	0.18	97.31
Tomatoes	252	1.13	98.44
Water Melon	196	0.88	99.32
Yams	129	0.58	99.90
Yellow Melon	23	0.10	100.00
Total	22,332	100.00	

These are the reported cases of own-consumed food from the production section of the questionnaire.

Q330.NM. Crop	Freq.	Per cent	Cum.
AGUISHI	1	0.06	0.06
AGUISHIE	1	0.06	0.11
B,BEANS	1	0.06	0.17
BAMABARAM BEANS	1	0.06	0.23
BAMABRA BEANS	1	0.06	0.29
BAMBARA BEANS	146	8.37	8.66
BAMBARA NEANS	1	0.06	8.72
BAMBARABEANS	3	0.17	8.89
BAMBARAM BEANS	45	2.58	11.47
BAMBARAMBEANS	3	0.17	11.64
BAMBARBEANS	1	0.06	11.70
BEANS	284	16.28	27.98
CASSAVA	2	0.11	28.10
COWPEA	9	0.52	28.61
GROUNDNUT	106	6.08	34.69
GROUNDNUTS	214	12.27	46.96
GROUNDUTS	2	0.11	47.08
KENNEF	4	0.23	47.31
MAIZE	151	8.66	55.96
MILLE	1	0.06	56.02
MILLET	314	18.00	74.03
OKRA	9	0.52	74.54
OKRO	1	0.06	74.60
OTHER GREEN LEAFY VEGETABLES	1	0.06	74.66
RICE	161	9.23	83.89
SORGHUM	207	11.87	95.76
SOYA BEANS	39	2.24	97.99
SOYA-BEANS	1	0.06	98.05
SOYABEANS	3	0.17	98.22
SWEET POTATOES	1	0.06	98.28
VEGETABLE	2	0.11	98.39
VEGETABLES	7	0.40	98.80
WATER MELON	1	0.06	98.85
WATERMELON	9	0.52	99.37
YAM	9	0.52	99.89
YAMS	2	0.11	100.00
Total	1,744	100.00	

Farm-gate prices can be obtained only for the main cereals (maize, rice, sorghum, and millet) and beans. For reasons of simplicity and consistency we decided to calculate prices from the consumption section of the questionnaire rather than employing farm-gate prices.

For each item we calculated four different sets of prices: household-level prices (the 'unit value' obtained by dividing the value of purchases by the quantity purchased); village-level prices (the median unit value in each locality); area prices (the median unit value in each of the four areas considered); all sample price (the median unit value across the entire sample). Prices were calculated for 21 different units of measurement because we were not able to convert all measurements in kg. So the total number of prices calculated is 63(food items)*21(units of measurement)*4(levels of measurement).

We applied this set of prices sequentially to consumption of own production:

- Household unit values, when available (note that household unit values were applied only in those cases for which the unit of measurement of purchases and own consumption were the same – unit values are noisy and hence we applied a routine that replaced unit value outliers by the median observed across the whole sample, outlier being defined as a unit value larger or smaller than 1.28 standard deviations from the log mean).
- Village median of unit values when household unit values were not available.
- Area median of unit values when village unit values were not available.
- All sample median of unit values when area median values were not available.