Kagera Health and Development Survey

Price Index

1. Introduction

KHDS-2 corrects prices in two ways:

1. Each household living in Kagera is linked to a price questionnaire, which was administered in all 4 waves from 1991/94, as well as in 2004. From this a Laspeyres index is calculated using the wave 1 of KHDS-1 as a base.
2. For households interviewed in 2004 who had moved out of Kagera we did not collect price information. Here we use the inflation rate of Kagera (which is close to the national inflation rate) to correct for price changes in time and correct for spatial differences between different regions by applying a Fisher index based on the Household Budget Survey (HBS) of the National Bureau of Statistics (NBS).

The corrections depend both on differences in location and time of the interview. Table 1 gives an overview of the source of information of different 2004 locations of the household.

Table 1: Sources of Information per 2004 location of the household

location in 2004	information basis	
	for Laspeyres Index	for Fisher Index*
In or near original cluster	price questionnaire administered in that cluster	= 1
Elsewhere in Kagera	price questionnaire administered in closest KHDS cluster	=1
Another region of Tanzania	Kagera's inflation rate between the surveys	from NBS, assuming the clusters are urban
Uganda	Kagera's inflation rate between the surveys	exchange rate

* Kagera is the base and equal to one

2. Households living in or near the original cluster

Each KHDS cluster has 5 price questionnaires, one in each of the four waves from 1991 to 1994 and one in 2004. The base price for the Laspeyres is the average price in the region per expenditure item i in wave 1 (denote B_i). Say the price of good i in cluster c is P_{ic} then the price index in cluster c is calculated as

$$index_c = \sum_i W_i \frac{P_{ic}}{B_i}$$

where W_i is the expenditure weight of good i.

Expenditure weights are calculated using the 1991/94 expenditure weights (see KHDS-1 User's Manual for details).

Missing prices for waves 1-4 were imputed using regression estimates (see KHDS-1 User's Manual for details).[1] For KHDS-2 missing prices were imputed differently. Per good i its average price increase across the Kagera region was calculated, denoted by I_i. If a price in KHDS-2 was missing, then it is imputed as its KHDS-1 price multiplied by I_i.

The file "price2004_cl.do" calculates the Laspeyres index for all KHDS clusters. It takes as given the KHDS-1 expenditure weights and the KHDS-1 missing prices imputations, but it recalculates the KHDS-1 indices in all 4 waves. This is important as the result calculated in the current do-file differ slightly from those calculated in the KHDS-1 do-files. This is because, contrary to KHDS-1, we treat all urban clusters as (No. 44-47) as one price observation. More on this below.

KHDS-2 has up to three price observations per cluster. Any 2004 prices that deviate more than 3 standard deviations from the regional average for that good (pooled data with all observations for that good in all clusters) are put to missing. Then the average across the non-missing price observations is taken per good per cluster. Table 2 shows the consumption basket, the expenditure weights per item, the number of non-missing observations and the average growth between the two surveys, used to impute missing values.

Table 3 gives, per cluster, the Laspeyres index and the number of non-missing, non-imputed price observations on which it is based. Prices increased by a little more than a factor of four in between the two survey rounds. This estimate is based on a total of 951 price observations in 2004.

There are 4 urban clusters (No. 44-47). These clusters are only a maximum of 2 km apart along the busiest tarmac road in the region. There is no reason to make different price indices for them. We take the average across the price observations between all these clusters and treat them as one cluster. Because this same procedure was not followed in KHDS-1, the 1991/94 price index calculated in KHDS-2 differs slightly between the two surveys. Note that because this procedure changes the base price it changes the indices in all clusters, compared to the index described in the KHDS-1 User's Manual. Table 4 compares the new index (calculated as described here) and the old one (calculated as described in the KHDS1 User's Manual) and shows that differences are minimal.

[1] KHDS-1 reports two indices, one where only missing values are estimated through regressions analysis and one where all values are replaced by the estimates. We opted for the second option when recalculating the KHDS-1 indices.

Table 2: consumption basket used to calculate the Laspeyres index

Good type	expenditure weight (W_i)	No. of non-missing observations	average price growth 1991/94 to 2004*
Raw Cassava	0.025	27	4.55
Dry Cassava	0.026	13	3.84
Cassava Flour	0.043	28	3.31
Sorghum	0.017	32	3.58
Wheat Bread	0.093	18	2.86
Rice Grade 3	0.038	46	3.05
Sugar Grade 3	0.042	48	2.72
Sweet Potato	0.019	32	3.71
Irish Potato	0.023	19	3.58
Kidney Beans	0.031	33	3.22
Salt Grade 3	0.009	48	2.27
Groundnuts	0.015	40	3.06
Tomato	0.018	45	2.83
Cooking Banana	0.088	26	3.39
Sweet Banana	0.018	28	4.29
Orange	0.010	16	4.33
Cooking Oil	0.011	47	2.43
Local Brew	0.045	34	3.23
Onions	0.011	45	2.71
Chicken Eggs	0.016	48	3.75
Chicken	0.018	18	4.54
Beef	0.034	34	4.16
Goat Meat	0.021	24	3.55
Fish	0.053	36	4.24
Fresh Milk	0.032	18	2.83
Milk Powder	0.032	10	8.77
Kerosene	0.025	45	4.87
Firewood	0.063	20	3.57
Battery	0.055	48	1.39
Linen	0.015	15	2.13
Charcoal	0.056	10	2.26
Total	1.000	951	

* the base is the average across the four waves of KHDS-1 administered between 1991/94

Table 3: The 2004 Laspeyres Index per Cluster and the No. of Price Observations on Which it is Based

cluster	Laspeyres index	no. of items out of the consumption basket for which prices were not imputed*
1	4.39	18
2	4.2	22
3	3.29	15
4	4.32	20
5	4.4	16
6	4.38	14
7	4.24	26
8	3.24	18
9	3.72	8
10	3.36	12
11	4.12	14
12	3.52	25
13	3.52	18
14	3.48	19
15	4.29	14
16	3.47	23
17	3.51	28
18	3.88	14
19	4.19	21
20	3.48	16
21	3.56	18
22	4.13	16
23	3.46	20
24	3.31	23
25	3.24	20
26	2.87	17
27	3.29	8
28	4.51	23
29	4.37	22
30	3.76	19
31	4.12	16
32	3.85	19
33	4.69	29
34	5.02	27
35	4.97	24
36	4.25	19
37	4.43	20
38	4.39	28
39	4.01	24
40	4.76	15
41	4.44	23
42	5.15	25
43	4.56	17
44	4.45	30
45	4.45	30
46	4.45	30
47	4.45	30
48	4.01	28
49	4.2	23
50	4.88	26

51	3.47	11
Mean	4.05	20

Table 4: Comparison of the wave 1-4 Laspeyres Index as calculated in KHDS-1 & KHDS-2

		Wave			
cluster	index	1	2	3	4
1	new	1.06	1.13	1.29	1.41
	old*	1.05	1.12	1.28	1.4
2	new	1.05	1.15	1.28	1.41
	old	1.04	1.14	1.27	1.39
3	new	0.83	0.88	0.95	1.08
	old	0.82	0.87	0.95	1.07
4	new	1.04	1.16	1.3	1.47
	old	1.03	1.15	1.29	1.46
5	new	1.02	1.13	1.25	1.41
	old	1.02	1.12	1.24	1.4
6	new	1.15	1.25	1.34	1.49
	old	1.14	1.24	1.33	1.48
7	new	1.21	1.3	1.45	1.67
	old	1.19	1.29	1.44	1.65
8	new	0.83	0.87	0.96	1.08
	old	0.82	0.86	0.95	1.07
9	new	0.91	1	1.12	1.26
	old	0.9	1	1.11	1.25
10	new	0.81	0.89	1	1.16
	old	0.81	0.89	0.99	1.15
11	new	1.1	1.18	1.33	1.5
	old	1.09	1.17	1.32	1.49
12	new	0.82	0.88	0.98	1.11
	old	0.82	0.87	0.97	1.1
13	new	0.85	0.93	1.03	1.14
	old	0.85	0.92	1.02	1.13
14	new	0.82	0.88	0.98	1.11
	old	0.81	0.87	0.97	1.1
15	new	1.07	1.18	1.33	1.5
	old	1.06	1.17	1.32	1.49
16	new	0.81	0.86	0.94	1.06
	old	0.8	0.85	0.94	1.05

17	new	0.81	0.87	0.97	1.07
	old	0.81	0.87	0.96	1.06
18	new	0.91	1	1.1	1.25
	old	0.9	0.99	1.09	1.24
19	new	1.07	1.17	1.27	1.44
	old	1.07	1.16	1.26	1.43
20	new	0.88	0.97	1.07	1.21
	old	0.88	0.96	1.06	1.2
21	new	0.84	0.9	0.98	1.1
	old	0.84	0.9	0.98	1.09
22	new	0.9	0.97	1.1	1.27
	old	0.9	0.96	1.1	1.26
23	new	0.81	0.88	0.97	1.1
	old	0.81	0.87	0.96	1.09
24	new	0.91	0.94	1.04	1.18
	old	0.9	0.94	1.03	1.17
25	new	0.85	0.9	0.99	1.13
	old	0.84	0.9	0.98	1.13
26	new	0.72	0.78	0.85	1
	old	0.71	0.77	0.85	1
27	new	0.83	0.91	0.99	1.16
	old	0.82	0.9	0.98	1.16
28	new	1.16	1.29	1.46	1.62
	old	1.15	1.28	1.44	1.61
29	new	1.17	1.28	1.43	1.62
	old	1.16	1.27	1.42	1.6
30	new	0.85	0.9	1	1.15
	old	0.84	0.9	0.99	1.15
31	new	1.05	1.14	1.26	1.45
	old	1.04	1.13	1.25	1.44
32	new	0.9	0.98	1.09	1.23
	old	0.9	0.97	1.08	1.22
33	new	1.22	1.28	1.42	1.62
	old	1.21	1.27	1.41	1.6
34	new	1.22	1.32	1.45	1.63
	old	1.21	1.31	1.44	1.62
35	new	1.22	1.31	1.47	1.71
	old	1.21	1.3	1.46	1.69

36	new	1.21	1.34	1.5	1.72
	old	1.2	1.33	1.49	1.71
37	new	1.17	1.31	1.44	1.64
	old	1.16	1.3	1.43	1.62
38	new	1.13	1.24	1.37	1.55
	old	1.12	1.22	1.36	1.54
39	new	1.12	1.22	1.37	1.53
	old	1.11	1.21	1.36	1.52
40	new	1.16	1.25	1.41	1.6
	old	1.14	1.24	1.39	1.58
41	new	1.15	1.27	1.43	1.62
	old	1.14	1.26	1.41	1.6
42	new	1.2	1.34	1.52	1.69
	old	1.19	1.32	1.5	1.67
43	new	1.17	1.3	1.47	1.68
	old	1.16	1.29	1.45	1.66
44	new	1.19	1.32	1.45	1.63
	old	1.14	1.24	1.38	1.52
45	new	1.19	1.32	1.45	1.63
	old	1.14	1.24	1.38	1.52
46	new	1.19	1.32	1.45	1.63
	old	1.17	1.3	1.45	1.63
47	new	1.19	1.32	1.45	1.63
	old	1.23	1.37	1.48	1.69
48	new	0.9	0.97	1.09	1.23
	old	0.89	0.96	1.08	1.22
49	new	1.04	1.15	1.29	1.49
	old	1.03	1.14	1.28	1.48
50	new	1.02	1.11	1.25	1.43
	old	1.02	1.1	1.24	1.41
51	new	0.85	0.92	1.03	1.19
	old	0.84	0.91	1.02	1.18

* 'Old' refers to the price index reported in the KHDS-1 User's Manual. This price index treats clusters 44-47 as separate observations. 'New' refers to new calculation after treating these clusters as one price observation. KHDS-1 had two indices, the 'fitted' where all prices were estimates and the 'actual' where only missing prices were estimates. Numbers reported in this table are the 'fitted' KHDS-1 index.

3. Comparison with secondary data

The obtained results were confronted with the consumer price index over the same period in Tanzania reported in the World Development Indicators (World Bank). Table 3 shows that, on average, KHDS estimates prices to have increased in Kagera with a factor of 4.05 since 1991/94. Table 5 below shows that this is very close to the estimate for Tanzania in the same period.[2]

Table 5: Comparison with the national CPI

year	inflation consumer prices, average (annual percentage)	1992 base	1993 base
1992	22	100	
1993	25	125	100
1994	33	166	133
1995	28	213	170
1996	21	257	206
1997	16	299	239
1998	13	338	270
1999	8	365	292
2000	6	386	309
2001	5	406	325
2002	5	426	341
2003	4	443	354

4. Households not living in or nearby the original cluster

For households living in *Kagera*, but not in the original cluster, the nearest KHDS cluster with price data was identified and recorded in a variable named "si11code" in the "hh.dta" file.

For households living in *Tanzania* in a region other than the Kagera Region inflation over time was imputed to equal to the Kagera average (i.e. 4.05, see Table 3). Spatial deflation was obtained by applying a regional Fisher price index obtained from the Household Budget Survey (NBS). The NBS reports per region both urban and rural Fisher price indices. KHDS did not collect, directly, information on whether the household is living in an urban or rural area. Most households who have moved out of the region are likely to live in urban areas and the urban index is applied to all of them. At the time of writing work is being planned to manually check the location data, compare them to the administrative divisions in Tanzania and determine whether the location is urban or rural.

[2] To make full use of the available data and to avoid noise to the extent possible, the base for the KHDS inflation figure of 4.05 is the *average* across the 4 KHDS-1 waves between 1991 and 1994. The bulk of these price data are from 1992 and 1993. Therefore we would expect the KHDS rate to lie in between the CPI figure with 1992 and 1993 as a base in Table 5. Note that this does not yet take account of the inflation in 2004. All price questionnaires were administered in the first half of 2004 and the 2004 inflation rate was not applied. The NBS reports inflation to be 4.2% in 2004.

To be sure: this work has not been completed yet and the current index assumes households are urban.

For households living in *Uganda* the Kagera inflation rate was applied for corrections in time. Corrections in space only include a correction for the exchange rate, as values in these questionnaires are reported in Ugandan Shillings.

Table 6: Fisher corrections for households living outside of Kagera

Region	NBS Fisher price	KHDS fisher price*
DODOMA	0.89	0.96
ARUSHA	0.79	0.85
KILIMANJARO	0.87	0.94
TANGA	0.95	1.02
MOROGORO	0.98	1.05
PWANI	0.88	0.95
DAR ES SALAAM	0.79	0.85
LINDI	0.73	0.78
MTWARA	0.75	0.81
RUVUMA	0.93	1.00
IRINGA	0.92	0.99
MBEYA	1.00	1.08
SINGIDA	0.97	1.04
TABORA	0.97	1.04
RUKWA	1.14	1.23
KIGOMA	0.98	1.05
SHINYANGA	0.94	1.01
KAGERA	0.93	1.00
MWANZA	0.94	1.01
MARA	0.91	0.98
UGANDA**		0.65

* = makes Kagera the reference region (=1), by dividing all NBS Fisher indices by the Kagera value (.93)
** = exchange rate as questionnaires indicate Ugandan Shillings

5. File Structure and Computing Price Corrections

"price2004_hh.do" creates a file with, for each household its 2004 Laspeyres price index and Fisher price index. The outcome of this file is saved as "price2004_hh.dta". The wave 1-4 deflators of 1991/94 are in a separate file, which is discussed next.

"price2004_cl.do" creates a file with, per cluster, 5 price corrections: one for each of the four rounds of KHDS-1 and one for KHDS-2. KHDS-2 is coded as wave=5. The outcome of this file is saved as price2004.dta. Cluster can readily be used as an identifier

to merge in price data for waves 1-4 of KHDS-1. However, in KHDS-2 *only for households living in or near their original cluster* (i.e. the variable“si2c” in “hh.dta” is equal to one or two) can the first two digits of hhid2 be used to identify the price cluster. It is suggested that this file only be used to obtain price indices for waves 1-4 of KHDS-1, but that the price2004_hh.dta file be used to identify the relevant price index for 2004.

To transform a 2004 monetary into its wave 1 (from KHDS-1) equivalent it should be divided by the Laspeyres index and multiplied by the Fisher index (=1 for all households living in Kagera). For 1991/94 it should only be divided by its Laspeyres index (all 1991/94 households live in Kagera and so their Fisher Index would be equal to one).