

Methods Used for Poverty Analysis in Rwanda Poverty Update Note

Andy McKay and Geoffrey Greenwell
Consultants to Oxford Policy Management
a.mckay@sussex.ac.uk and gpg@greenwell.org

First draft, 16 February 2007

1. Introduction

The Rwanda Poverty Update Report published in December 2006 presented results on trends in poverty in Rwanda between 2000/01 and 2005/06, the years of the two major recent household surveys, the *Enquêtes Permanentes sur les Conditions de Vie des ménages*. The Update Report focused on three main dimensions: poverty and inequality in monetary terms; social indicators; and economic activities/time use; in each case focusing on changes between the periods covered by the two surveys. All three of these dimensions are important aspects of poverty in their own right; this note though explains the methods used for the calculation of the basic poverty and inequality measures. This is particularly relevant for the analysis in chapter 2 of the Poverty Update Report; but it is also relevant for the definition of the quintile groups used in the analysis of other variables.

While there were some revisions made to the questionnaire between the EICV1 and EICV2 surveys, the data needed for the estimation of poverty and inequality was deliberately maintained the same. Therefore, the approach used in the analysis of poverty and inequality in the Poverty Update Report is entirely consistent with that used in the earlier Poverty Profile report published by MINECOFIN based on the EICV1 results, with only some very minor revisions to ensure complete consistency with the EICV2 survey data. The EICV1 results in the Poverty Update Report are therefore almost exactly the same as those published before.

2. Overview: the basic approach to the estimation of poverty and inequality

The analysis of monetary poverty from the EICV1 and EICV2 household surveys is based on household consumption expenditure data. This follows standard international practice. While data on household incomes is also collected, it is generally the case that it is harder to measure income accurately, especially for agriculture and other self employment activities. In addition households may have many different sources of income. The level of household consumption – including purchases but also consumption from other sources including own production and payments received in kind – therefore is usually considered a better indicator of the resources available to a household. For Rwanda household consumption can be estimated using the data collected by the EICV surveys.

But in seeking to compare consumption between households, and use it assess poverty and inequality, it is important to take account of two additional points.

- First, households face different prices depending on the time of year when they are reporting their consumption (for example whether it is before or after the harvest periods); and also depending on where they live in the country (for example, households in the City of Kigali generally face higher prices for food than households in many rural areas). These differences in prices need to be taken into account, and this is done so here by calculating a price deflator.
- Second, households differ in the number of members they have and in the age of these members, so that larger households or those with a higher proportion of prime age adults are likely to have higher consumption needs. This is done here by means of an adult equivalent scale, which allows household size to be measured in terms of “adult equivalents”, recognising that the consumption needs of younger children for instance will be less than those of prime age adults.

The rest of this note explains in turn the procedures used for the measurement of household consumption; the construction of the price deflator; the adult equivalence scale used; the computation of the adjusted consumption standard of living measure; and the estimation of the poverty line.

3. Estimating household consumption expenditure

The EICV questionnaire collects detailed information on household expenditures, as well as on consumption obtained from non-purchased sources – for example consumption of food crops grown by the household. The contents of the consumption aggregate used for the analysis of poverty and inequality are summarised in Table 1, which also gives the sources of the data from the questionnaire. For most households the most important components are purchases of food items, consumption of food items that the household produced itself (*autoconsommation*) and purchased non-food goods and services. In addition the consumption aggregate includes spending on education, everyday health expenses (routine consultations), expenses on housing and utilities (water, electricity), as well as other smaller items such as the value of wages received in kind and of transfers made by the household to other households.

The construction of the consumption aggregate follows standard international practice on what items to include and which to exclude (Deaton and Zaidi, 2002). Care was taken to exclude non-consumption expenditures of the household (e.g. purchases of business of farm inputs); to exclude purchases of large durable goods the consumption of which is spread over many years; and also to exclude large, one off exceptional items of expenditure such as weddings and funerals, or hospital stays. In the case of durable goods instead an estimation is made of estimated consumption flows based on the current value of any durable goods the household owns and commodity-specific depreciation rates. The exceptional items of consumption are excluded precisely because of their one-off nature: a household may have incurred a very high expenditure on a wedding in the period covered by the survey, but if this was included in the consumption aggregate it would exaggerate the household’s normal level of consumption.

Table 1: Contents of household consumption aggregate used for poverty analysis

Component	Description of contents, and items covered	Source of data in questionnaire
Expenditure on education	Household expenses on costs of schooling for all members currently enrolled in any level of education	Section 2, questions s2aq10a – s2aq10h
Regular health expenses	Expenses on consultations for those that were ill in the two weeks preceding the interview	Section 3, question s3aq16, s3aq19
Lodging: rent and utilities	Imputed rental value of owner-occupied dwellings (respondent provided valuation); rent in cash and in kind; expenses on water and electricity	Section 5, questions s5bq1, s5bq2, s5bq4a, s5cq2a, s5cq3, s5cq7a
Employer provided benefits in kind	Payment received by employees in kind (...); and in the form of subsidised transport	Section 6, question s6eq10, s6eq14
Non-food expenses	Consumption of own produced tobacco Infrequent non-food items: purchases in past year, but excluding purchases of durable goods and items already reported elsewhere	Section 9D, questions s9dq4—s9dq13, and s9dq15 Section 9A1, question s9a1q3, but excluding the following items: 806, 809, 810, 811, 812, 813, 814, 815, 821, 822, 823, 824, 830, 1001—1008 inclusive
	More frequently purchased non-food items: based on purchases in last month	Section 9A2, question s9a2q3
	Frequently purchased non-food items	Section 9A3, based on questions s9a3q4—s9a3q13
Food purchases	Purchases of all food items	Section 9B, questions s9bq4—s9bq13
Consumption of own produced food	Consumption of own produced food items	Section 9D, questions s9dq4—s9dq13, and s9dq15, excluding the following (non-food) commodity codes:
Transfers paid out	Transfer made to other individuals or households, in cash or in kind	Section 10A, question s10aq8—s10aq10
Other consumption expenses	Estimated consumption flows derived from durable goods (based on current value and estimated depreciation rates)	Section 11A, questions s11aq5a—s11aq5c, and commodity specific estimates of depreciation rates
	Other expenditures: ...	Section 10C, questions s10aq2, commodity codes 17 and 18 only

In the EICV surveys consumption values are collected from households with differing recall periods depending on the frequency of purchase or consumption. Thus for infrequently purchased items such as households are asked to report their purchases over the last year, or last month in the case of more regular non-food purchases such as transport. But for the most frequent items, including all food items, households are asked over several visits to report their purchases or *autoconsommation* over the two or three days (in rural and urban areas respectively) since the last visit of the enumerator. This is considered to provide more accurate recall for the two or four week (in rural and urban areas respectively) period covered. However, all consumption needs to be computed with reference to a consistent time period. Here all consumption is expressed on an annual basis, grossing up the consumption reported with reference to shorter time periods to give annual values.¹

In the process of calculating consumption aggregates, attention must be given to the potential problem of outliers, the presence of extremely large values which mean that the value of household consumption is overestimated. One common reason for this in surveys is data entry errors, where for example extra zeros are added or decimal points are missed. In fact this is very unlikely in this case due to the thorough process of double data entry; and because the careful storage of the questionnaires from the two EICV surveys made it very straightforward to check apparently extreme observations against the original questionnaires.

Nevertheless the possibility of outliers still remains. Hence in constructing the consumption aggregate, mean values and standard deviations at the regional level were computed for each individual consumption commodity (on a per equivalent adult basis in the case of food purchases, *autoconsommation* and the main non-food purchases), and this was used to identify extreme observations as those lying more than three and a half standard deviations away from (above in practice) the mean value. Extreme values thereby identified were replaced with the mean value (multiplied by the number of equivalent adults in the household where relevant). once the data was annualised a search was conducted for extreme values, as observations lying more than three and a half standard deviations away from the mean value.² The regional disaggregation used for this process distinguished the following locations: the City of Kigali, Other Urban areas, rural Eastern province, rural Northern province, rural Southern province and rural Western province. Doing these at a regional level allows for the fact that average consumption levels of a commodity may differ significantly by location.

In practice very few observations were identified as outliers, substantially less than 1% [*add details*]. Moreover because the re-estimation is done at a highly disaggregated level, the total consumption of the household is unlikely to be

¹ This does mean that an individual household's annual consumption will tend to be over- or under-estimated if the household happened to be surveyed in a peak or slack season respectively for consumption; but the way in which the sample is designed (covering all geographic areas simultaneously) means that the average value of consumption for households in a particular location should be unbiased.

² The choice of the three and a half standard deviation criterion reflects the fact that if the statistical distribution of the variable can be approximated by a normal distribution, the probability that an observation lies so far away from the mean value is very small.

substantially affected by this procedure. While the value of the consumption of the commodity which was re-estimated will be significantly reduced by this procedure, it this is only one of a very large number of consumption commodities (the vast majority of which will not be re-estimated for the same household). Thus there is no reason to think that estimates of poverty or inequality will be significantly affected by this procedure.

Table 2: National average values of household consumption expenditure, EICV1 and EICV2

Expenditure category	Average value in current prices (thousands of FRw/household/year)		Share of total consumption	
	EICV1	EICV2	EICV1	EICV2
1. Food purchases	119.0	218.3	27.2%	25.8%
2. Consumption of own produced food	110.3	208.9	25.2%	24.7%
3. Undeclared food consumption (1+2)	229.3	427.2	52.4%	50.5%
4. Non food purchases	113.1	262.0	25.9%	31.0%
5. Rent: actual and imputed	43.8	46.5	10.0%	5.5%
6. Education expenditures	12.3	33.5	2.8%	4.0%
7. Routine health care expenditures	5.0	4.9	1.1%	0.6%
8. Utilities (water and electricity)	4.2	7.2	1.0%	0.9%
9. Use value of durable goods	13.1	27.2	3.0%	3.2%
10. Transfers made by household	8.8	15.6	2.0%	1.8%
11. Wage payment in kind (consumption counterpart)	6.3	18.3	1.4%	2.2%
12. Other expenditure	1.4	2.9	0.3%	0.3%
13. Undeclared non food expenditures (sum 4 to 12 inclusive)	207.9	418.1	47.5%	49.5%
14. Undeclared total consumption (3+13)	437.2	845.3	100.0%	100.0%

Note: these averages are calculated using household level weights, and so should be representative of all households in Rwanda.

Table 2 reports the average household values for the consumption data computed from the two EICV surveys. All values reported here are in current prices, in other words they have not been adjusted for inflation over the period covered by the surveys. As will be seen below this was substantial. In nominal terms most expenditures increased substantially over this period. But in general it is more appropriate to focus on the consumption shares in this table (the last two columns). In

general these are seen to be quite stable between the two periods, which may be interpreted as an indicator of good data quality.

The share of consumption accounted for by food falls slightly over this period. While this might be viewed as an indicator of improved average living standards over the period; but before this conclusion can be drawn with confidence it is important to take account of any changes in the relative price of food and non-food consumption items.

4. Adjusting for differences in prices faced by households

As already noted, in using the household consumption aggregates computed above it is important to express all consumption values in a common set of prices, taking account of differences in commodity prices over space and time. There are three dimensions to this:

- (i) there are significant differences in price levels between different areas of the country, reflecting levels of transport costs as well as other factors including local production patterns, and where imported commodities arrive into the country;
- (ii) there are significant differences in prices of food commodities in particular across different seasons of the year (as well as inflation over the year), and the valuations provided by respondents are likely to reflect prevailing prices at the time of interview;
- (iii) in comparing poverty between EICV1 and EICV2 it is essential to take account of inflation over the period between the surveys.

Adjustments have been made for this by means of a Laspeyres price index, considering food and non-food commodities separately. This index is computed as follows:

$$P_{r,t} = \sum_{i=1}^m w_{i,0,0} \left(\frac{p_{i,r,t}}{p_{i,0,0}} \right)$$

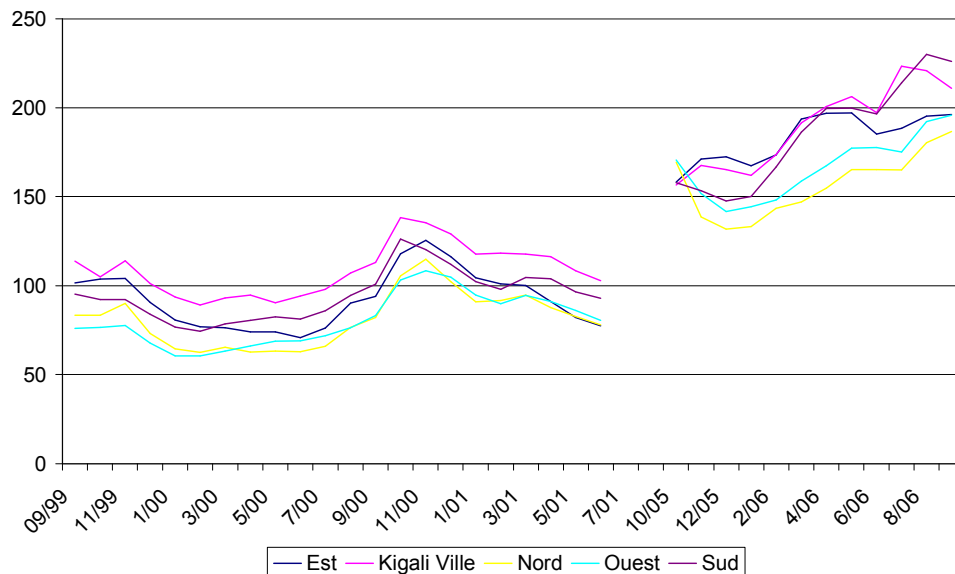
where i ($i=1, \dots, m$) indicates the commodity, $P_{r,t}$ is the price index for location r in time period t ; $p_{i,r,t}$ is the price of commodity i in location r in time period t ; $p_{i,0,0}$ is the price of commodity i in the reference location ($r=0$) and time period ($t=0$); and $w_{i,0,0}$ is the budget share of commodity i in the reference location ($r=0$) and time period ($t=0$). These budget shares are computed to include values of consumption of own production as well as purchases; this is appropriate because it is used to deflate a welfare measure which includes *autoconsommation* as well as purchases.

In this case the reference consumption basket is the national average basket in January 2001 for the poorest 60% of the population in terms of household consumption expenditure per adult equivalent (see section 5 below). The commodity shares of this basket ($w_{i,0,0}$ above) were computed from the EICV1 survey for the first poverty profile, and the same basket is used again. The basket was used as the basis for calculating the price index for each month covered by the two surveys, for each of five regions -- corresponding to the five new provinces (city of Kigali, and the East, North, south and West provinces).

The source of the price data used to compute the index differed between the food and nonfood index. In the case of the food index, the price data are obtained from the twice monthly data collected by the MINAGRI Mercuriale programme of price data

collection (previously PASAR: *Programme d'Appui à la Sécurité Alimentaire au Rwanda*). This collects standardised data on prices of main food products in 36 markets across Rwanda, 3 in each of the former provinces on a comparable basis over the EICV1 and EICV2 time periods; this translates into about 6-8 markets in each of the new provinces. This data was used to compute an average price for each commodity in each month in each of the five regions ($p_{i,r,t}$ above), and, by taking the population weighted average for each commodity in January 2001, to compute $p_{i,0,0}$. sufficiently complete price series were available for 26 commodities corresponding to ...% of the food consumption basket and including all staple commodities. The data were checked for outliers, but only one was identified reflecting a keying error; and there were no missing observations at the five province level.

Figure 1: Variation in Food Prices, by location and time

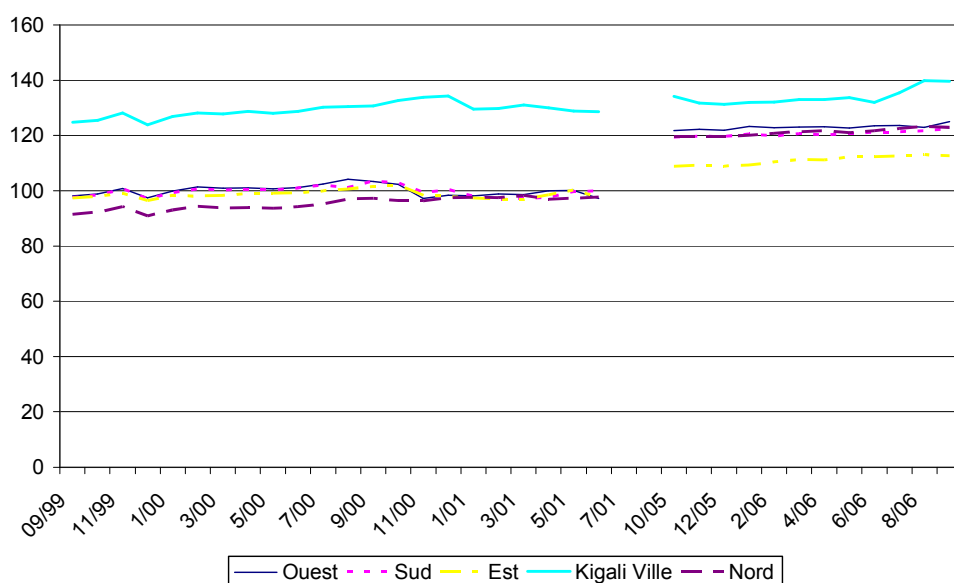


The computed values for the food index are reported in Figure 1. The values of the index show (i) quite large spatial differentials by region, with food prices being lowest in the north (important producing region of many key staples) and west, and highest overall in Kigali (though also in the South in the latter part of the EICV2 period); (ii) substantial intra-year variability of price, reflecting seasonal variations (though the patterns are less evident in the East and South in the EICV2 period, probably reflecting localised droughts in these areas in the 2006A agricultural season); (iii) significant inflation in food prices between two surveys, and over the period of the EICV2 survey (the average food index for the EICV2 period is 80-100% higher than for the EICV1 period).³ It is clearly very important to take account of these important price differences in comparing consumption levels between households and over time.

³ This is a bit greater than the CPI suggests for the same period, it is because the concept of inflation here is different, and more relevant for the analysis of poverty. Because of the purpose of the CPI, a very high weight is placed on Kigali (where a lot of purchasing takes place), and the weights reflect the consumption of the entire population (not the poorest 60% as here). This is the appropriate procedure for computing the CPI, followed in all countries. For poverty analysis though it is important to have a concept of inflation which is more relevant for the poorer groups

In the case of the nonfood index the price data are those collected for the consumer price index, which are again averaged for each month in each of the five regions, and expressed relative to the reference price. In this case some imputation of missing values was needed, when commodities were not available in specific months in specific localities, and some editing for outliers was also required when quality differences seemed to be apparent. However, a sufficiently complete set of price data was available on 41 nonfood commodities for which expenditure information was also available; these commodities accounted for 79.1% of the selected basket of purchased non-food commodities. The values of the nonfood consumption index are plotted in figure 2.

Figure 2: Non-food price index for Rwanda, by location and time

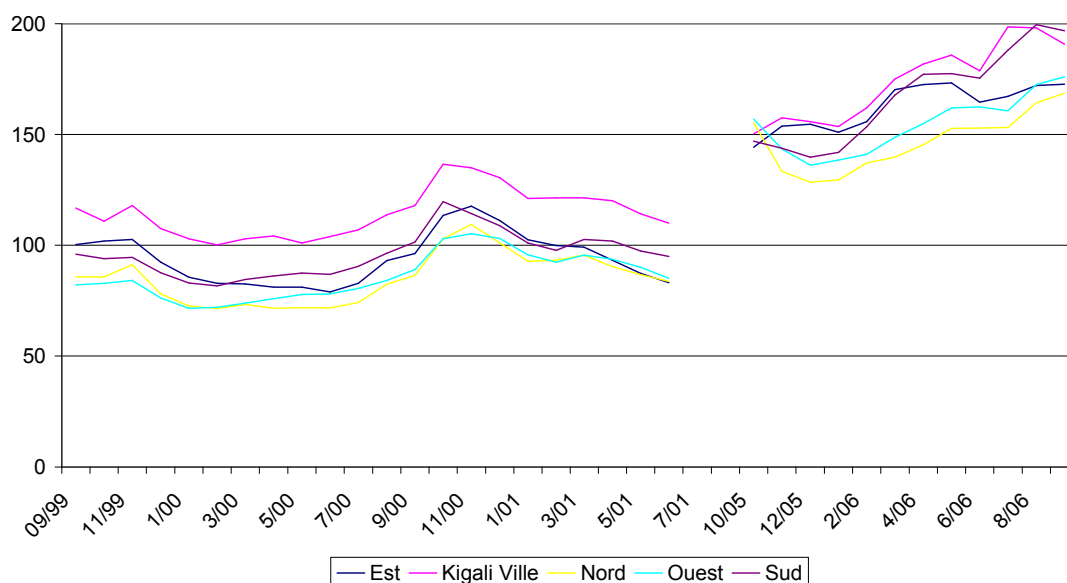


For non-food there are also important differentials by location, though these have narrowed over the period, mainly because prices have increased outside of Kigali. Nonetheless, the inflation rate for non-food commodities is much less than that for food commodities. In addition non-food commodities do not show much intra-year variation in prices, as expected.

The food and nonfood indices (each expressed relative to January 2001=100) were then combined into an overall living with weights reflecting the share of food and non-food in the reference consumption basket (71.8% and 28.2% respectively). The values of the combined index (Figure 3) therefore tend to be much more influenced by the food index. The resulting index is used to deflate the consumption measure discussed in section 3, to give consumption in the constant prices of January 2001

in the population. In Rwanda the poor seem to have faced significantly higher food price inflation.

Figure 3: Overall price index (Jan 2001=100)



5. Adjusting for differences in the size and composition of households

The differences between households in their size and composition are allowed for by computing household size measured in adult equivalents, using the scale reproduced in Table 3. This is the same scale that has been used for a long time in Rwanda, certainly dating back to the analysis of the *Enquête Nationale de Budget et de Consommation* conducted in 1983, and which was used for the EICV1 Poverty Profile study.

Table 3: Adult equivalence scale used for the construction of the consumption indicator

Age range	Gender	
	Male	Female
Less than 1 year	0.41	0.41
1 to 3 years	0.56	0.56
4 to 6 years	0.76	0.76
7 to 9 years	0.91	0.91
10 to 12 years	0.97	1.08
13 to 15 years	0.97	1.13
16 to 19 years	1.02	1.05
20 to 39 years	1.00	1.00
40 to 49 years	0.95	0.95
50 to 59 years	0.90	0.90
60 to 69 years	0.90	0.80
More than 70 years	0.70	0.70

Household effective size then is measured relative to a reference adult in the 20-39 year age range. The table reports the consumption needs of other people relative to the reference age range, and these differ by age and sex. Those aged 16-19 years are regarded as having slightly higher consumption needs than the reference, but most younger children and adults aged 40 and above have lower consumption needs. This scale was apparently chosen for the analysis of the 1983/84 survey based on discussions with the FAO at the time; it is not out of line with similar adult equivalence scales used in other countries.

To illustrate its use, consider as a made up example consider a household that comprises a 32 year old man, 27 year old woman, a 12 year old girl, a 7 year old boy and a 3 year old girl. The size of this household in terms of number of members is five. But in terms of adult equivalents using the above scale its size is 4.55 ($1.0 + 1.0 + 1.08 + 0.91 + 0.56$). This takes account of the fact that the two younger children, in particular the lowest, are estimated as having lower consumption requirements than the adults and the oldest girl.

The differences between the adult equivalent measure and simply counting household size in terms of number of members are most significant where a household comprises many very young children or older adults. The adult equivalent measure is a more accurate reflection of households' consumption needs than just counting numbers of people.

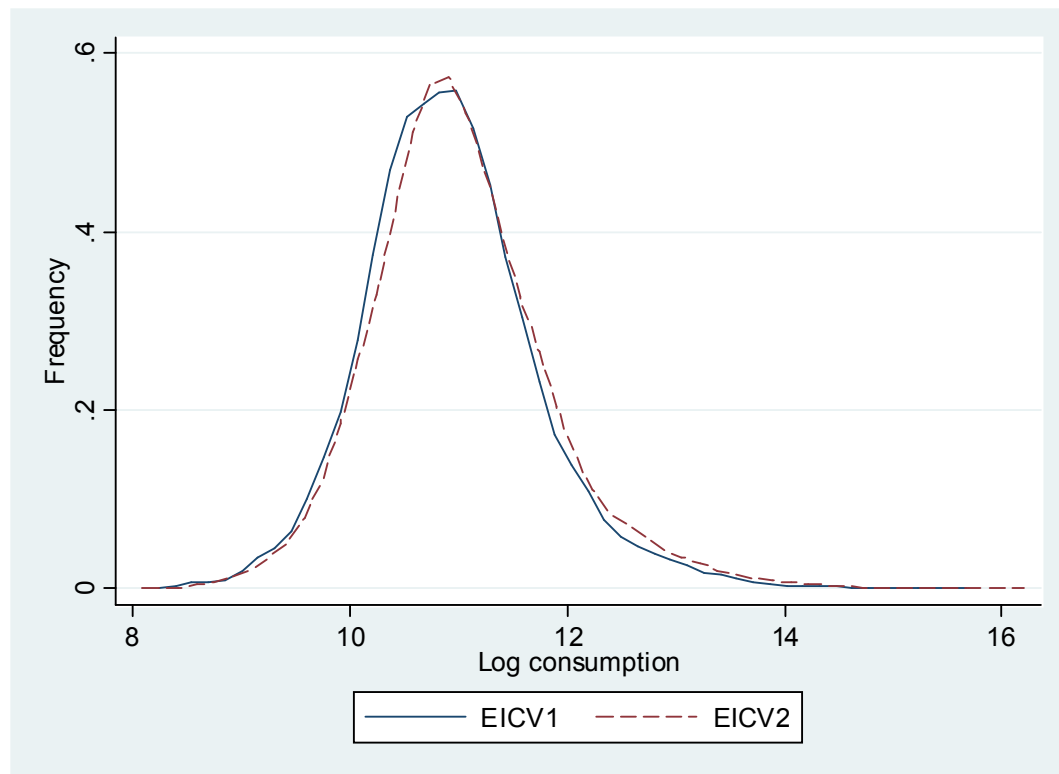
6. Construction of the consumption measure

The consumption based standard of living measure is then computed as total household consumption expenditure (computed as in section 3 above), expressed in constant prices (using the index discussed in section 4) per adult equivalent. The analysis of this variable is weighted distribution, where the weights reflect the sample weights multiplied by the household size. The sample weights are used to allow the surveyed households to be considered as representative of all households in Rwanda. The multiplication by household size is because in looking at poverty and inequality it is appropriate to look at the distribution over individuals rather than households. A poor household which comprises five members should be given a higher weight in computations of poverty than a similarly poor household but which has only two members. The distribution can then be considered as representative of the population of Rwanda.

This approach is based on a strong, but unfortunately necessary, assumption: that individuals are represented in the distribution by the consumption measure of the household they belong to. This of course fails to take account of inequality in distribution within the household. Unfortunately though (as in most surveys) the EICV surveys do not collect information on the intra-household distribution of consumption.

The distribution of the logarithm of the consumption variable is presented in Figure 4 for the two surveys. This shows that the distribution seems to have shifted to the right over this period: in general consumption levels in real terms have increased over this period.

Figure 4: Kernel density plots of the logarithm of real household consumption per adult, EICV1 and EICV2.



Note: this graph has been computed using population weights.

7. Poverty Line

The same absolute poverty line was used that was computed for the EICV1 poverty profile. In January 2001 prices the poverty line was set at FRw 64,000 per adult per year, and an extreme poverty line (below which households could not even afford the basic food consumption basket, even without spending anything on non-food items was FRw 45,000 per adult per year. In January 2006 prices these poverty lines translate into FRw 90,000 and FRw 63,500 per adult per year respectively. This translates into approximately FRw 250 and FRw 175 per adult per day respectively.

The full details of the computation of the poverty line are reported in Annex 2 to the EICV1 Poverty Profile. In looking at changes in absolute poverty between the EICV1 and EICV2 surveys it is important to use the same poverty line in real terms for both. At this point it was felt that the poverty line estimated for the EICV1 poverty profile had sufficient acceptance that it could be used again for the current analysis. There may though be a significant case for constructing a new poverty line in another five years, reflecting changes in consumption basket over this period.

As well as the analysis of absolute poverty, it is also important to consider changes in relative poverty as well. To some extent this is captured in the Poverty Update Report by also including an analysis in of inequality; but this is an issue which should be considered much more in subsequent analysis of this data set.

Reference

Deaton, A and S. Zaidi (2002), *Guidelines for Constructing Consumption Aggregates for Welfare Analysis*, World Bank: Washington D.C.