DHS Wealth Index Construction

1. There have been four stages of development of the DHS Wealth Index
    a. In the first stage, the WI made use of existing variables in the DHS survey, which were gathered for reasons other than determining economic status. A single principal components analysis (PCA, Factor Analysis) was used.
    b. In the second stage, specific questions were added to the household questionnaire for the purpose of having a more precise index and one that perhaps was less weighted toward urban areas. A single PCA was used as well. Generally information on number of farm animals and agricultural land area was not used.
    c. In a third stage, separate indexes for urban and rural households were produced and then combined into a national wealth index, by regressions onto a common index. Number of farm animals by type and agricultural land area were used in both the urban and rural index calculations with the expectation that the relationships would be different in each area. Three PCAs were produced, one for a set of indicators common to both urban and rural areas, which did not contain number of farm animals or agricultural land area. Two other PCAs were produced, one for urban areas and one for rural areas, both of which contained number of farm animals by type and agricultural land area.
    d. In a fourth stage, the number of farm animals were binned by type, due to the observation that the relationship between the number of farm animals by type was far from linearly related to the national, the urban, and the rural wealth indexes. Land area was also investigated for binning but did not have a particularly non-linear relationship.
    e. In later surveys that did not have questions on the number of animals by type and land area, only a single national index was produced.
2. The DHS wealth indexes for some surveys were created from the standard recode file while others were created from the raw data file.
3. There are two steps programs used to construct the wealth index: CSPRO (and previously ISSA), to export the data and SPSS to calculate the index. A typical export is given in csexport.pdf, which is applicable to all surveys with a little modification. The variables DOMESTIC, HOUSE and LAND are created in CSPRO due to it hierarchical data structure. Other programs could also be used to export the data with or without these three variables. The wealth index could also have been created in other statistical programs.
4. Each survey's information comes from the SPSS syntax used to produce the wealth index. The syntax files (with .sps extensions) have been converted to Adobe PDF files (.pdf extensions). Some surveys have two .sps files, one from the export of data and dictionary from ISSA or CSPRO to create the SPSS data file (extension .sav) and the other with the syntax for indicator variable construction and wealth index calculation. In other surveys, there is only one syntax file covering both procedures.
5. There are also Excel files which have information taken from the SPSS output files:

a. Unweighted means and standard deviations for each of the indicator variables used in the PCAs.
b. The component score coefficients for each of the indicator variables.
c. Two columns for producing a wealth score, one if the household has the item and the other if it does not. These are produced by subtracting the mean of the item from 1 and 0, respectively, and dividing by the standard deviation (to get a standardized) score and multiplying by the component score coefficient for the item. Continuous variables, such as number of members per sleeping room and land area do not have entries into this column and need to be calculated directly from the value for each household.
d. In surveys where there are urban and rural indexes, three spreadsheets are included for the PCA results, one for the common PCA, one for the urban PCA, and one for the rural PCA.
e. A final spreadsheet gives the cutpoints (with other statistics) for the wealth index quintiles. The quintiles are based on the weighted dejure household population (not households). A table of the mean value of each indicator by national quintile and for total is included, as is a histogram of the wealth index scores for many of the surveys. For surveys with urban and rural areas, the regression coefficient outputs and formulas for producing the combined national wealth index are also given.
f. See the publications "The DHS Wealth Index", Comparative Report 6 (http://www.dhsprogram.com/publications/publication-CR6-Comparative-Reports.cfm), "The DHS Wealth Index: Approaches for Rural and Urban Areas" Working Paper 60 (http://www.dhsprogram.com/publications/publication-WP60-Working-Papers.cfm), and "Steps to constructing the new DHS Wealth Index", in the Wealth Index for Web folder.

6. The wealth index files are grouped together by country with the name of the country and date of the survey. MIS and AIS surveys also are indicated in the name of the files; otherwise the files are for DHS surveys.
7. To construct a wealth index comparable to the national index, say for a local area survey, or for users of services, follow these steps:
    a. Ask the questions in the DHS household questionnaire that are used for the wealth index (usually section 2) with the same categories for each question. Also needed is the number of dejure household members. Less important, there are three questions derived from the woman's and men's questionnaires that are asked of each individual questionnaire respondent that are used to determine whether there is a live-in domestic servant (service occupation of an individual respondent or her husband and who is not related to the head of the household), DOMESTIC, whether the respondent owns a dwelling either individually or jointly, HOUSE, and whether the respondent works on agricultural land own individually or jointly, LAND. If any yes for any individual respondent then the household is given a yes value (1). Not all surveys have these variables and they do not change the value of the wealth index much if not asked.
    b. Create the wealth index indicator values from categorical questions into the dichotomous variables given in the spreadsheet following the SPSS syntax. Be sure that

yes-no questions are coded 1 for yes and 0 otherwise (also give missing values a 0). The missing categories for categorical variables are not transformed into wealth index indicator variables. Also create the members per sleeping room variable, MEMSLEEP, and the land area variable, LANDAREA, and the binning variables for the number of farm animals by type.

c. Standardize the wealth indicator variables for each case by subtracting the mean given in the spreadsheet from the case's value and dividing the difference by the standard deviation from the spreadsheet. Do not use the mean and standard deviation from the new data.
d. Multiply the standardized value of each indicator by its component score coefficient and sum over all the resulting products to get the wealth score for that case. (Note that as a short cut, the last two columns of the last spreadsheet does steps c. and d. except for continuous variables.)
e. If the DHS (AIS, MIS) survey has common, urban, and rural indexes, produce the three wealth scores as in c. and d. Note that urban areas will have common and urban wealth scores but no rural wealth score and vice versa. Then apply the urban and rural regression coefficients to get the combined wealth score.
f. Recode the national/combined wealth score into national quintiles using the cutpoint values in the spreadsheet. Note that the quintiles will in the new data represent national quintiles at the time of the DHS survey and will not be 20% of the cases in the new data.