User notes for PMA2013/CD Round 1 (Kinshasa) Household and Female data, version 1

Disclaimer: PMA2020 cannot provide in-depth support for data analysis or data related questions, however, to assist the end-user, explanation of some variables is provided below.

Generic

SIF variables: Data and time variables are provided in both string format and as Stata Internal Format (SIF) values. The variable name of any variable that has been changed into SIF is appended with SIF (e.g. system_date and system_dateSIF). For all questions requiring a date entry, if the respondent answered either "Do Not Know" or refused to answer the question, the date was recorded as January 1, 2020.

Select multiple variables: Some questions allow for the selection of multiple answers. The values for these variables are the concatenation of answer choices (e.g. if a household respondent said that they use two sources of water, choices a and c, the value of the observation would read "a c"). Multi-select options are generally, though not always, transformed into binary variables for analysis.

Country specific variables: All variables in PMA2020 have consistent values for option choices across countries (e.g. marital_status==1 is equivalent to currently married in all countries) with the exception of the following

1. Livestock questions: The specific livestock options (cow, rabbit etc) vary across countries
2. school: education categories for female schooling vary across countries
3. fp_provider: provider of current or most recent method of family planning vary across countries
4. roof/wall/floor: Household materials vary across country
5. assets: The household assets used to construct wealth scores vary across countries as do the binary variables that are created from the multi-select asset question
6. wealthquintile/wealthtertile: In some countries, wealth quintiles are provided, in other wealth tertiles. The continuous variable score is included to allow for reconstruction of various wealth categories.

Specific variables

EA: The primary sampling unit

metainstanceID: metainstanceID is the unique ID generated by ODK for each form submitted to the central server. For PMA2020, the variable metainstanceID is unique for each household but will be repeated within the household. **memberID** will provide a unique ID for each person within the household.

FQmetainstanceID: FQmetainstanceID is the unique ID generated by ODK for each female form submitted to the central server. For PMA2020, the variable FQmetainstanceID is unique for each female surveyed.

current_recent_methodnum, current_methodnum, recent_methodnum: The numbering scheme for contraceptive methods is consistent across all PMA2020 countries. For example, female sterilization is equal to 1 in every PMA2020 country, whether or not there are any reported uses of female sterilization in the dataset. In some countries, therefore, the numbering will be non-consecutive if some method choices are not selected.

cp, mcp, tcp: Variables that identify current users of any contraceptive method (cp), a modern contraceptive method (mcp), and a traditional contraceptive method (tcp) are included in publicly available datasets so that PMA2020 estimates involving current contraceptive use and method mix can be replicated. Values for these variables are 0 (no) or 1 (yes). PMA2020 codes cp, mcp, and tcp based on the variable current_methodnum with the following caveats:

1. Women who report not being a current user of contraception (current_user=0), but who report using EC in the past 12 months (recent_methodnum=8. emergency) are coded as cp=1 and mcp=1. During analysis, current method is classified as EC in the method mix; however, current method is not changed in the data that is publicly available.
2. Women who report using LAM as a current method (current_methodnum=14. LAM) must satisfy the three conditions listed below to be coded as mcp=1. If any of these conditions are not met, these women are coded as tcp=1. During analysis, current method is classified as LAM or traditional method; however, current method is not changed in the data that is publicly available.
    a. Less than six months post-partum
    b. Amenorrheic
    c. Indicating that they are using LAM with the intention of preventing pregnancy
3. Women who report female sterilization as their first contraceptive method (first_methodnum=1. female sterilization), but who do not report currently using female sterilization are coded as cp=1 and mcp=1. During analysis, current method is classified as female sterilization in the method mix; however, current method is not changed in the data that is publicly available.

Country specific notes

Question wording and option choices varied between Round 1, Round 2, and subsequent rounds of data collection. For variables that are included in each round (such as roof, wall, and floor) but where either wording choice or option choice varied across Round 1 and Round 2, variables end with CDR1 and CDR2. Rounds 3 and Rounds 4 have consistent wording and answer choices and are thus taken as the standard choices within the country. Variables in these datasets are not named with CDR3 or CDR4.

In Round 1, non-response at the female level could not be calculated, therefore only a household level weight was created. All women in the household are selected and non-response is generally low at the enumeration area level, so HHweight can be used in lieu of FQweight.

GPS Variables

No GPS coordinates for either household or service delivery points will be released for any reason.

Sampling

PMA2013/Kinshasa-R1: the first round of data collection in Congo Democratic, used a representative urban sample in Kinshasa. A sample of 60 enumeration areas (EA) was drawn (PPS). For each EA, 30 households were selected. Households were systematically sampled using random selection. Households with eligible females of reproductive age (15-49) were contacted and consented for interviews. The final sample included 1. 813 completed households, and 2, 132 completed de facto females. Data collection was conducted between October 2013 and January, 2014.

Analytic sample

PMA2020 analyses include only observations from completed household interviews. The female sample includes only completed female interviews from completed households. The majority of indicators include only de facto women (women who slept in the household the night before). All observations, however, are included in the dataset to allow end users to calculate response rates.

Dataset version updates

Any updates made to datasets after their initial release will be documented here.

In January 2017, all previously released datasets were modified as below:

1. The value of **age_at_first_use_children** is 0 for women who have ever used family planning and who have never given birth. Previously, such women had a missing value for age_at_first_use_children.
2. The values for **water_sources_main_drinking** and **water_sources_main_other** equal the value of water_sources_all if a household has one water source. Previously, such households may have had a missing value for these variables.

3. The value for **sanitation_main** equals the value of sanitation_all if a household has one sanitation facility. Previously, such households may have had a missing value for this variable.

All datasets released after January 2017 will have these changes included.

In May 2017, **version 2** of the DRC (Kinshasa) Round 1 dataset was released. The updated version corrects the following:

- The variable **current_recent_methodnum** to recode users of cycle beads to the rhythm method. In DRC (Kinshasa) Round 1, due to a programming error, users of cycle beads and the rhythm method were combined during data collection. PMA2020 has coded these women as rhythm method users.
- The variables **HHweight** and **FQweight** were normalized so that the mean of each weight is 1. Previously, these weights were not normalized.

To report errors or inconsistencies:

Please email datamanagement@pma2020.org