User Notes for PMA2020 Ethiopia Round 6 Follow-up Female Survey Dataset, Version 1.0

Disclaimer: PMA 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.

Ethiopia Round 6 Follow-up Survey Sampling

PMA2020 Ethiopia Round 6 Follow-up (ETR6FU) Female Survey sample included women who participated in PMA2020 Ethiopia Round 6 (ETR6) Female Survey in 2018 and consented to be re-contacted.

ETR6 survey used a two-stage cluster design with urban-rural, and major regions as strata and included 221 enumeration areas (EAs) from the previous round drawn by the Central Statistical Agency from its master sampling frame. For each EA, 35 households were selected. Households were systematically sampled using random selection. Eligible females of reproductive age (15-49 years) were consented and interviewed for ETR6 Female Survey.

For the ETR6FU Female Survey, eligible women were consented and asked the Network Scale-Up Method (NSUM) module and the Confidante module along with additional population, pregnancy, contraceptive use and discontinuation questions. The final sample included 4,910 *de facto* females who completed the follow-up interviews. Data collection was conducted between March and April 2020.

Materials included in this Data Zipped File

This data zipped file includes:
1. Female questionnaire pdf file
2. User notes pdf file
3. Codebook excel file
4. Dataset in 3 file formats: .csv, .xlsx and Stata .dta

General Variables

SIF variables: Date 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**).

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

Variable Response Options

Select one: Most select one numeric variables have consistent values for option choices across countries (e.g. marital_status==1 is equivalent to currently married in all countries). Exceptions include the variables **school**, **fp_provider**, **floor**, **roof**, and **walls**, which have country-specific options and numbering.

Select multiple: Similarly, most select multiple variables have the same response options across countries. Some select multiple variables, however, such as **assets**, have answer options that vary across countries.

See the Ethiopia Round 6 Follow-up Codebook for complete details on variables and answer choices.

Specific Variables

FQmetainstanceID: FQmetainstanceID is the unique ID generated by ODK for each female form submitted to the central server. This variable FQmetainstanceID is unique for each female interviewed for the Round 6 Follow-up Survey.

R6FQmetainstanceID: This is the unique female IDs of the Round 6 Survey. This variable can be used to merge the follow-up data with the Round 6 data.

EA_ID: The primary sampling unit masked with a random number for anonymity. The same random number is applied to the same EA across multiple rounds.

RE_ID: Identification number of the resident enumerator (RE), or interviewer. The same ID is applied to the same REs across different rounds of the PMA2020 surveys.

GPS Variables

GPS coordinates are not released in this dataset. For some PMA2020 countries a separate dataset of displaced GPS coordinates at the EA level is released.

Missing Data

In Stata, Missing data is expressed as "." in the cell. Generally, Stata commands perform computations of any type handle missing data by omitting the row with the missing values. However, this may vary across commands. PMA does not impute missing values. Missing data in datasets should be studied and/or treated before proceeding to analysis.

Reasons for missing data:
Normal situations:
1. Incomplete forms: If a household, female, or SDP form is not marked as completed (HHQ_result, FRS_result, SDP_result not equal to 1), the observation is likely to miss most of the information. Incomplete forms should not be included in the analysis.

2. Observations that are ineligible for subsequent forms: Only eligible respondents will receive subsequent forms. For example, males and ineligible females will not receive female questionnaires in family planning surveys, hence their observations will have all missing values in female forms.

3. Question not administered due to skip logic: PMA surveys use ODK's skip logic function. The subsequent questions are administered selectively based on the respondent's previous answers. Irrelevant or inapplicable questions are skipped. For example, a woman who is not a contraceptive user will not be asked questions about contraceptive usage subsequently.

Uncommon situations:
1. Lost forms: Due to technical constraints in some challenging data collection areas, forms can be lost in the process of data submission. Although most forms were recoverable, there are occasionally a few that cannot be found. For example, an observation from an eligible woman with completed female form information but missing household form information, or vice versa. These observations may be dropped based on analysis needs.

2. Missing due to incorrect skip logic: PMA surveys were conducted under rigorous quality control. However, in rare cases, there can be incorrect skip logic, which skipped a question that was supposed to be administered, resulting in missing values. These errors are documented in the PMA codebook, which can be downloaded from PMA website. It's not necessary to drop the entire observation since this will likely affect only a few questions.

Distinguish missing data from negative values:
1. -99: No response. The respondent was administered with the question but did not provide an answer. PMA survey requires consent from the respondent and the respondent has the right to refuse to answer any questions at any point. -99 is recorded to reflect that the respondent did not provide an answer to a certain question.

2. -88: Did not know. The respondent consented to answer a specific question but without knowing the answer.

3. -77: Not applicable. The question is administered to the respondent but not applicable to the respondent's situation.

PMA GitHub Repository

The PMA GitHub Repository https://github.com/PMA-DM/PMA_Analyses_Public has Stata .do files which could be used to generate indicators in the briefs, using the Household and Female, Service Delivery Point, and Client Exit Interview Datasets that are publicly available.

Note: Data presented in the online briefs represent preliminary results. Therefore, there may be slight differences between the .do file results and those in the brief. Please access the PMA DataLab https://datalab.pmadata.org/ or the Snapshot of Indicator (SOI) pdf file available online to cross check any discrepancies and get the final estimates.

Dataset Version Updates

Any updates made to datasets after their initial release will be documented here. PMA releases new versions on as needed basis and users who have been approved previously to get access to the datasets will be notified via email upon the release of the new versions. Users can then log onto the PMA website and re-download the datasets without having to submit a new request.

Users should always be using the latest version of the datasets available on the PMA website. However, if users need an access to the old versions of the datasets for any reasons, users can contact datamanagement@pmadata.org directly with an explanation for why the access to the old version was needed.

Dataset Citations

Suggested citation: Addis Ababa University School of Public Health; and the Bill & Melinda Gates Institute for Population and Reproductive Health at the Johns Hopkins Bloomberg School of Public Health. Performance Monitoring and Accountability 2020 (PMA2020) Ethiopia Round 6 Follow-up Female Survey (Version 1.0), PMA2020/ETR6FU-FQ. 2020. Ethiopia and Baltimore, Maryland, USA. https://doi.org/10.34976/01zk-gb17

To report errors or inconsistencies:

Please email datamanagement@pmadata.org