User Notes for PMA-Ethiopia Panel Cohort 2 (1-year Follow-up) Survey Dataset, Version 2.0

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.

PMA Ethiopia

Performance Monitoring for Action Ethiopia (PMA Ethiopia) builds on the previous success of PMA2020/Ethiopia and PMA Maternal and Newborn Health study in the Southern Nations, Nationalities and Peoples Region (SNNP). PMA Ethiopia is a five-year project implemented in collaboration with Addis Ababa University, Johns Hopkins University, and the Federal Ministry of Health. It measures key reproductive, maternal and newborn health (RMNH) indicators.

Cross-sectional data, including a health facility-based survey, are collected annually in all regions. Longitudinal data (following pregnant women through one year postpartum) are collected in two cohorts of women (2019-2021 and 2021-2023) in four large, predominantly agrarian regions: Tigray, Oromiya, Amhara, and Southern Nations, Nationalities, and Peoples' Region, and one urban region, Addis Ababa. Afar is included in the first cohort (2019-2021) of the longitudinal survey.

Sampling

PMAET Panel Cohort 2: 1-Year Follow-up survey followed women who were eligible to remain in the Panel Cohort 2 study and did not refuse to follow-up at their previous interview. Women who consented to follow-up at the baseline and were still living at the time of the last interview; did not report a still birth, miscarriage, or abortion at the previous interview; did not refuse the previous interviews; and did not refuse future follow-up at the previous interview were included in this survey.

A total of 1,858 mother or caregivers completed the 1-year follow-up survey. Panel women could become ineligible for 1-year follow-up if they refused follow-up at some point, died, or had no live births. The data collection for the 1-year follow-up survey occurred between September 2022 and August 2023.

For more information on the PMA survey methodology and sampling, please refer to PMA Survey Methodology at https://www.pmadata.org/data/survey-methodology.

Materials included in with this data zipped file

The data zipped file includes:

1. PMAET_Panel_Cohort2_1yrFU_v#_Date (Dataset in 3 file formats)
2. PMAET_Panel_Cohort2_1yrFU_FemaleQuestionnaire_v#_Date (pdf)
3. PMAET_Panel_Cohort2_1yrFU_Usernotes_v#_Date (pdf)
4. Example_Merge_BL_6wkFU_6moFU_1yrFU_v#_Date (Sample Stata dofile)

Codebook

The latest version of the PMA Ethiopia Household and Female Survey master codebook can be downloaded from the https://www.pmadata.org/data-codebooks.

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**). The *"Do not Know"* value for dates is *Jan 1, 2030*. For each date question where the woman did not know the month but knew the year, the value is Jan 1 of that year and the variable with the same name as the date variable but ending in **_m_dnk** will have a value of 1.

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. Multi-select options are generally, though not always, transformed into binary variables for analysis.

Variable Response Options

Select one: For the select one numeric variables, consistent values for option choices were given across PMA countries. Some select one variables, however, such as **region**, have answer options that vary across countries.

Select multiple: Similarly, most select multiple variables have the same response options across all PMA countries.

See the PMAET HQFQ Master Codebook for complete details on variables and answer choices for each survey.

Specific Variables

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 survey years.

RE_ID: Identification number of the resident enumerator (RE), or interviewer. RE names are masked with the PMA-Ethiopia random numbers in the household dataset. The same random number is applied to the same REs across different surveys of the PMA-Ethiopia grant.

PMA2020_RE_ID: The resident enumerators (REs), or interviewers, who were involved in the previous PMA2020 surveys (2014-2018), also had PMA2020 IDs. This ID is consistent for all survey rounds of the PMA2020 grant.

participant_ID: A unique ID of each respondent participating in panel surveys. The same ID is given to the same female respondent across different surveys. This variable should be used to merge or identify women throughout the study.

OYmetainstanceID: OYmetainstanceID is the unique ID generated by ODK for each female form submitted to the central server. For PMA-Ethiopia, the variable OYmetainstanceID is unique for each female in the one-year follow-up survey.

OYFUweight: Weight for the six-week follow-up survey data

current_methodnum: The numbering scheme for contraceptive methods is consistent across all PMA countries. For example, female sterilization is equal to 1 in every PMA 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.

Preparation Data for Panel Analysis

If you wish to do longitudinal analysis, you can merge data from the baseline, 6-week follow-up and 6-month follow-up to the 1-year dataset using **participant_ID**. The Information and the example stata dofile on how to combine the baseline, 6-week follow-up and 6-month follow-up data can also be found in the 6-month follow-up dataset zipped file. Example Stata dofile for merging the previous survey data to the 1-year follow-up data is included in this dataset zipped file for your reference.

There will be 340 women from baseline who do not merge with the 1-year data. These women became ineligible between screening and baseline, refused follow-up at some point, died, or had no live births.

Analyzing Children's Data Across Surveys

The data for the individual children born to the women are in wide format, since a woman could have had twins. So, all of the variables specific to each child born are duplicated and have a suffix of 1 or 2 depending on if they reference the first or second child born.

During cleaning, data managers ensured that child 1 from the 6-week data was the same infant as child 1 in the 6-month and 1-year data, so there is no need to individually merge children. Instead, to analyze data at the child level, after merging the 1-year dataset with the 6-month, 6-week and baseline dataset by participant_id, you will need to reshape the data to account for the women who had twins.

```
drop if participant_id==""
reshape long VARIABLES OF INTEREST, i(participant_id) j(index)
bysort participant_id: drop if _n==2 & pregnancy_type!=2
```

GPS Variables

GPS coordinates are not released in this dataset.

Notes for 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, and Service Delivery Point 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/ 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.

In April 2024, a new version (version 2.0) of this dataset was released with the following updates:

- **totay todaySIF** variables were added to the dataset.
- **OY_result** variable was renamed as OY_result_cc for standardization purpose
- List of choice names: **more_children_list** and **providers_list** were updated.

- **methods_list_num** choice list was corrected.

Dataset Citations

Suggested citation: Addis Ababa University School of Public Health and the William H. Gates Sr. Institute for Population and Reproductive Health at The Johns Hopkins Bloomberg School of Public Health. Performance Monitoring for Action Ethiopia (PMA-ET) Panel: Cohort 2 - One-Year Follow-up Survey (Version 2.0), PMAET-Panel-C2-1yrFU. 2023. Ethiopia and Baltimore, Maryland, USA. https://doi.org/10.34976/kzcb-6b45

To report errors or inconsistencies:

Please email datamanagment@pmadata.org.