



**USAID**  
FROM THE AMERICAN PEOPLE

## **Feed the Future FEEDBACK**

### **Malawi 2015 Zones of Influence (ZOI) Interim Survey Public Release Datasets and Documentation**

#### ***README Data Description Document***

**February 2017**

This deliverable package contains CSV, SAS, SPSS, and STATA datasets; a codebook; and a Privacy Assurance Statement (PAS) for the Malawi 2015 ZOI Interim Survey. Please note that these datasets also contain harmonized data for two ZOI districts (Balaka and Machinga) from the coordinated 2015 Food for Peace (FFP) baseline survey in Malawi. (This FFP survey, titled *Baseline Study of FY 2014 Food for Peace (FFP) Development Food Assistance Projects in Malawi*, was conducted July-September 2015 by ICF International.)

The sample for the Feed the Future Malawi ZOI interim assessment comprises rural areas of seven districts distributed across Malawi's Central and Southern regions: Michinji, Lilongwe, Dedza, Ntcheu, Mangochi, Balaka, and Machinga. Data for Balaka and Machinga were collected in the coordinated FFP/ICF baseline survey, harmonized, and merged with ZOI interim survey data.<sup>1</sup>

Fifteen datasets are included with this deliverable: Household, Household Roster, Women, Children, three WEAI files, and eight Expenditure files. Note that all modules were collected in the Malawi 2015 ZOI interim survey (i.e., all standard Feed the Future indicators for the Malawi interim assessment were calculated from primary ZOI interim survey data [including the harmonized records from the FFP baseline survey data]).

These Malawi ZOI interim survey datasets contain observations from households with a completed interview. The variable order is generally the same as the Malawi interim questionnaire (also included in this deliverable package). New/derived variables are labeled and typically located at the end of each file. (However, there is a separate "WEAI\_Recode" dataset containing the derived variables from the WEAI analysis.)

#### **Privacy Notice**

These datasets are freely available to the public without restriction. However, all personal identification information that was present in the original data was removed. In addition, variables that could potentially be used indirectly (in combination with other variables or with other data sources) to identify any particular household or respondent were recoded to prevent that disclosure.

Location variables, unless otherwise noted, were randomized such that the location codes no longer reflect the original administrative codes or codes used in the sample frame. The anonymous codes are

---

<sup>1</sup> A crosstab in the household file of the variables "a05" (District) and "survey" (Interim Feed the Future or Baseline FFP) illustrates for data users this coordinated data collection, and presents the ZOI interim assessment records collected in the seven districts of the ZOI by each of the two subcontractors (Westat [Feed the Future] and ICF International [FFP]).

provided such that analysts may identify which records were collected from the same areas. The anonymous codes cannot be used to merge or compare with any other data source.

Households with large numbers of members can potentially be identified if large households are unique or rare in the survey districts. To effectively avoid revealing the identity of any large household that may be the only one or among a very few of its size in its district, data for a number of household members were removed from several households. The number of households' data affected and the number of household members' data removed through this procedure will not be disclosed. The members' data that were removed were not eligible for interview elsewhere in the questionnaire.

Because of these modifications, it may not be possible for analysts to exactly reproduce all of the results in published country reports and indicators.

## A. Datasets

In each of the **CSV, SAS, SPSS, and STATA** subfolders, there are 15 datasets for the Malawi 2015 ZOI Interim Survey. In addition to the data files, please note that for the SAS version of the datasets, there is a SAS formats catalog file (containing the value labels for the SAS datasets): **formats.sas7bcat**.

1. **MWI\_HOUSEHOLD\_PR** – This dataset (n=1,021, vars=90) is a household-level file, and thus contains records for each sampled household with a completed interview in the Malawi 2015 interim assessment (including the harmonized records from the FFP survey). The unique identifier for this file is pbs\_id, and the dataset includes variables from Modules A, D, and F. Please note that all key household-level derived variables are also included in the household dataset, including, for example, gendered household type (“ghht”) and the household hunger indicator (“hhunger”). These derived household-level variables can be merged into other datasets as needed using the unique household identifier pbs\_id.
2. **MWI\_HHMEMBERS\_PR** – This dataset (n=4,779, vars=36) is the household member roster for the households included in the Malawi 2015 interim assessment. It has one record per household member from Module C. The unique identifiers for this file are pbs\_id + idcode; these variables uniquely identify each household member in each sampled household.
3. **MWI\_WOMEN\_FILE\_PR** – This dataset (n=942, vars=75) is an individual-level dataset with all women age 15-49 with a completed interview in Module H. The unique identifiers for this file are pbs\_id + idcode.
4. **MWI\_CHILDREN\_FILE\_PR** – This dataset (n=843, vars=120) contains records for all children under 5 years of age in the sampled households. It includes data from Module I for children's anthropometry and infant and young child feeding practices. The anthropometry Z-scores were calculated in SAS during the data management process, using the World Health Organization (WHO) SAS “igrowup” package. The unique identifiers for this file are pbs\_id + idcode.
5. **MWI\_WEAI\_PR** – This dataset (n=948, vars=166) is the first of two datasets needed to calculate the WEAI-related measures. It contains Module G data from the primary adult (18+) female decisionmaker within each household (for the sub-sample of households with a primary adult female decisionmaker). The unique identifiers for this file are pbs\_id + idcode. (Note that pbs\_id alone can uniquely identify records in this WEAI file because, for the interim survey, there was only one person per household [the primary adult female decisionmaker] interviewed for the WEAI module.)

6. **MWI\_WEAI\_TIMEUSE\_PR** – This dataset (n=17,064 vars=117) is the second of the two WEAI datasets used to calculate the WEAI measures. It includes all of the 24-hour time allocation data from Module G6, the time use questionnaire, and thus each respondent in Module G has multiple records—one for each of the 18 time use activities (17,064 records ÷ 18 activities = 948 WEAI respondents). The unique identifiers are pbs\_id + idcode + activity.
7. **MWI\_WEAI\_RECODE\_PR** – This dataset (n=948, vars=168) contains the analytical variables derived in the WEAI analysis. The unique identifiers in this individual-level file are pbs\_id + idcode.
8. **MWI\_EXP\_PR** – This dataset (n=1,021, vars=73) contains variables from Module E, the Household Consumption Expenditures module used to calculate the poverty and expenditure indicators. It includes household-level derived variables (including the expenditure and poverty indicator variables), as well as variables from sub-Module E6: Housing Expenditures. The unique identifier in this household-level file is pbs\_id.
9. **MWI\_EXP\_E1\_PR** – This dataset (n=115,373, vars=26) contains data from sub-Module E1: Food Consumption Over Past 7 Days. Each household with food consumption data has multiple records (for the 113 food items in sub-Module E1). (115,373 records ÷ 113 food items = 1,021 Module E households with sub-Module E1 data.) The unique identifiers in this long file are pbs\_id + e101.
10. **MWI\_EXP\_E2\_PR** – This dataset (n=10,210, vars=16) contains data from sub-Module E2: Non-Food Expenditures Over Past 7 Days. Each household with data for non-food expenditures over the past week has multiple records (for the 10 non-food items in sub-Module E2). (10,210 records ÷ 10 non-food items = 1,021 Module E households with sub-Module E2 data.) The unique identifiers in this long file are pbs\_id + e201.
11. **MWI\_EXP\_E3\_PR** – This dataset (n=26,546, vars=16) contains data from sub-Module E3: Non-Food Expenditures Over Past One Month. Each household with data for non-food expenditures over the past month has multiple records (for the 26 non-food items in sub-Module E3). (26,546 records ÷ 26 non-food items = 1,021 Module E households with sub-Module E3 data.) The unique identifiers in this long file are pbs\_id + e301.
12. **MWI\_EXP\_E4\_PR** – This dataset (n=39,819, vars=16) contains data from sub-Module E4: Non-Food Expenditures Over Past Three Months. Each household with data for non-food expenditures over the past three months has multiple records (for the 39 non-food items in sub-Module E4). (39,819 records ÷ 39 non-food items = 1,021 Module E households with sub-Module E4 data.) The unique identifiers in this long file are pbs\_id + e401.
13. **MWI\_EXP\_E5\_PR** – This dataset (n=32,672, vars=16) contains data from sub-Module E5: Non-Food Expenditures Over Past 12 Months. Each household with data for non-food expenditures over the past 12 months has multiple records (for the 32 non-food items in sub-Module E5). (32,672 records ÷ 32 non-food items = 1,021 Module E households with sub-Module E5 data.) The unique identifiers in this long file are pbs\_id + e501.
14. **MWI\_EXP\_E5CONS\_PR** – This dataset (n=2,250, vars=20) contain data from sub-Module E5 regarding non-food items that may or may not have been purchased (e.g., construction items such as wood/bamboo poles and thatching grass). Each household with data for these 2-3 construction items over the past 12 months has multiple records. (Note that there is an “other” category of construction items for records from the harmonized FFP Baseline Survey data.) The unique identifiers in this long file are pbs\_id + e504.
15. **MWI\_EXP\_E7\_PR** – This dataset (n=31,651, vars=20) contains data from sub-Module E7: Durable Goods Expenditures. Each household with data for durable goods expenditures has multiple records (for the 31 durable goods in sub-Module E7). (31,651 records ÷ 31

non-food items = 1,021 Module E households with sub-Module E7 data.) The unique identifiers in this long file are pbs\_id + e701.

## **B. Codebook**

The Malawi 2015 ZOI Interim Survey codebook includes two worksheets. The “Variable\_List” worksheet lists the dataset name, variable name, variable label, variable description, variable type, and (where applicable) value label names. The “Notes” column in this worksheet provides helpful information regarding specific variables (e.g., e701, e703). A second worksheet, “Formats\_Value\_Labels,” provides value label names, values, and value labels for the applicable categorical variables. Both original and derived (analysis) variables are included in this codebook for the Malawi 2015 ZOI Interim Survey.

## **C. Privacy Assurance Statement (PAS)**

A Privacy Assurance Statement for the Malawi interim assessment is also included in this data deliverable package. The PAS document presents a general overview of the requirements and process to ensure respondents’ privacy and confidentiality prior to public release of survey data. This document also includes the informed consent statements from the USAID-approved questionnaires (for both the Feed the Future survey as well as the Food for Peace survey). Informed consent was obtained by trained interviewers from each eligible respondent to the survey.