

## **Read Me - Integrated LMPS v5.0**

August 2, 2023

### **The ILMPS**

The ILMPS is a data set that integrates and harmonizes data and variables from five rounds of the Egypt Labor Market Panel Survey (in years 1988, 1998, 2006, 2012, and 2018), two rounds of the Jordan Labor Market Panel Survey (2010 and 2016), the 2014 Tunisia Labor Market Panel Survey, and the 2022 Sudan Labor Market Panel Survey. It contains created, compatible variables that are harmonized (to the extent possible) across all rounds. The codebook includes the information on the variables included in the dataset. There are a few variables that are present only in one survey; the rest are available in two or more rounds.

### **I. Updates**

The v5.0 version of the ILMPS primarily adds in the SLMPS 2022 round. Additional revisions have been undertaken in order to harmonize the data and add variables relative to previous versions. The real and PPP international dollar conversions have been updated to 2022 terms.

### **II. Files**

#### ***Codebook: Codebook Integrated LMPS v5.0***

Gives variable name, type, label, ranges and states, universe, rounds covered, and notes.

#### ***Data File: Integrated LMPS v5.0.dta***

Format: long

Number of variables: 838

Number of observations: 275,015

Note: Data are not self-weighted.

#### ***Do Files***

All do files involved in harmonization run in sequence based on “**ILMPS master do file v61.do**”. See also the do files for each of the individual surveys.

If you are interested in particular variable’s creation, you can also search for it among the do files which reflect the labeling and coding of answers from the country questionnaires.

### **III. Variables**

#### ***Important variables***

round: Identifies round of survey (1988, 1998, 2006, 2010, 2012, 2014, 2016, 2018, 2022)

country: Identifies country (400=Jordan, 729=Sudan, 788=Tunisia, 818=Egypt)

unique household identifier: hhid

unique individual identifier: indid

#### ***Weights***

*Weights are the same across variables for Egypt 1988-2012 and Jordan 2010 and only differ for*

*Tunisia (all four different), Jordan 2016 (expan\_indiv is different), Egypt 2018 (expan\_indiv is different) and Sudan 2022 (expan\_indiv is different)*

expan\_hh: household questionnaire expansion factor

expan\_roster: household roster expansion factor

expan\_indiv: individual questionnaire expansion factor (both children and adults)

expan\_migr\_ent: household migration/enterprise questionnaire expansion factor

The weight used should be selected based on which questionnaire the *dependent* variable comes from. If data is being used across multiple questionnaires, we still recommend using the weight based on the questionnaire the dependent variable comes from, but that covariates be created so that observations missing data on those covariates are included with dummies for missing. For instance, if using receipt of remittances (covariate) to predict child schooling (dependent variable), create remittance receipt categorically/as a series of dummies for: 0 (did not receive), 1 (did receive), 2 (data missing).

### ***Variable naming conventions***

Country prefixes (400, 729, 788, 818) often added to variable values, such as for ids and governorates, to distinguish them in the integrated data.

Variable names are based primarily on those used in ELMPSs, and equivalent information for Jordan and Tunisia was merged in.

## **IV. Codebook order of presentation**

The order of presentation of variables in the codebook is consistent with the order of the created variables in the dataset, and both place related categories of variables together.