ELMPS 2023 Read Me
Version 2.0
October 18, 2024
FOR PUBLIC RELEASE

Data files

All data files are available as STATA .dta files.

2023 cross section: ELMPS 2023 xs v2.0.dta

- indid: unique individual identifier
- hhid: unique household identifier
- For those who cannot load the many variables in this data set, you can separately load pt 1, pt 2 and pt 3 versions (pt 1 includes created variables; pt 2 includes individual questionnaire variables; pt 3 includes household questionnaire variables)

1998-2023 panel: ELMPS 2023 panel v2.0.dta

- Findid: unique individual identifier from when individual was first observed
- The data are constructed as panel data (wide format) directly from the repeated cross section.
- This file contains only the created variables across waves.
- Variables repeat as 'variable'_# where # is wave, i.e. pn_06, pn_23 for the person numbers in 2006 and 2023. If an individual is not present in a wave, the variables for that wave will be missing.

1988/1998/2006/2012/2018/2023 repeated cross section: ELMPS 2023 rep xs v2.0.dta

- round: Identifies wave of survey (1988, 1998, 2006, 2012, 2018, 2023)
- indid: unique individual identifier
- hhid: unique household identifier
- This file contains only the created, compatible variables across time.
- The data are long structure, observation is an individual & year combination

Do files

All current data runs out of "Master Do File ELMPS 2023 v69.do" This do file calls sub-do files (all in do files folder)

- **Note:** For Sections: housing mobility, education, fathers' and mothers' tables, job mobility, marriage and fertility, and for those who responded (panel individuals) in any of the previous sections that there are no changes in their status between 2018 & 2023, their missing answers have been replaced with their original data (original variables) in 2018.

Weights

- expan_hh: this is the cross-sectional *household* expansion weight for the full sample and can be used with the rep xs and xs files. This is the main weight you should use in cross-sectional analyses on the *household* level (using variables from household

questionnaire, household level analyses). This variable is available in the panel, but should not be used for panel analyses; see panel weight below.

- expan_indiv: this is the cross-sectional *individual* expansion weight for the full sample and can be used with the rep xs and xs files. This is the main weight you should use in cross-sectional analyses on the *individual* level (using variables from individual questionnaire or individual level analyses). This variable is available in the panel, but should not be used for panel analyses; see panel weight below.
- expan_ref_hh: this is the weight, specific to 2018, that is a cross-sectional expansion weight based on only the refresher sample. This is a weight you can use in cross-sectional analyses on the *household* level (using variables from household questionnaire, household level analyses).
- expan_ref_indiv: this is the weight, specific to 2018, that is a cross-sectional expansion weight based on only the refresher sample. This is a weight you can use in cross-sectional analyses on the *individual* level (using variables from individual questionnaire or individual level analyses).
- panel_wt_*: this is the panel weight, designed for analyses spanning different combinations of years (_*). It is not an expansion weight, but normalized to 1.

Enforcement of skip patterns

Skip patterns were only enforced uni-directionally; the programming ensured that individuals answered every question they were supposed to. However, if there was a misunderstanding/data entry such that they went backwards and revised to follow a different skip path, stray/entered data remained even if no longer relevant. Analyze according to the universe who should have answered per skip patterns.

Questionnaires

The questionnaire is divided into two parts:

1. Household questionnaire
2. Individual questionnaire

English and Arabic questionnaires for each of these parts are available.

Codebooks

Created variables are codebooked in: Codebook ELMPS 2023 data v2.0

For created variables only. Gives variable name, type, label, ranges and states, universe, years covered, and notes.

There are additional Arabic codebooks for:

- Country codes
- Geographic codes (gov.xls and stat_u.xls for mapping locations to urban/rural)
- Primary school codes
- Preparatory school codes
- Secondary school codes
- University codes
- Higher education specialization codes (specialization.xlsx)