Read Me
ELMPS 2012 v2.1
PUBLIC DATA RELEASE

Files:

Questionnaires
Questionnaires in English and Arabic
1—Household
2—Individual
3—Migration & Enterprises

Codebook
File: Codebook ELMPS 2012 data v2.1.xls
For created variables only. Gives variable name, type, label, ranges and states, universe, years covered, and notes.

Data Files
Cross-section (xs) data for 2012
Files: elmps 2012 xs v2.1 all.dta (all variables)
      ---OR---
      elmps 2012 xs v2.1 pt 1.dta (created variables and q0198-q11226_2 (questionnaires 1 & 2))
      elmps 2012 xs v2.1 pt 2.dta (created variables and q12101-q14510 (questionnaire 3))

NOTE: the "elmps 2012 xs v2.1 all.dta" data file contains more than 2,000 variables. If you are using STATA IC, this file has too many variables to open. Use "elmps 2012 xs v2.1 pt 1.dta" and "elmps 2012 xs v2.1 pt 2.dta" instead. If you can open a file with more than 2,000 variables, you can just use "elmps 2012 xs v2.1 all.dta."

hhid: unique household identifier
indid: unique individual identifier
expan_12: expansion factor (must be used as an analytical weight or a probability weight).

Note: Data is not self-weighted. For details on the creation of sample weights, see: Assaad, R., and Krafft, C. (2013). The Egypt Labor Market Panel Survey: Introducing the 2012 Round. *Economic Research Forum Working Paper No. 758*

Repeated cross section for 1988, 1998, 2006, 2012
File: elmps 2012 comptbl 88_12 v2.1.dta
Important Variables:
round: Identifies round of survey (1988, 1998, 2006, 2012)
hhid: unique household identifier, begins with two-digit round number (88, 98, 06, 12)
indid: unique individual identifier, begins with two-digit round number (88, 98, 06, 12)
expan: expansion factor (must be used as an analytic or probability weight)

NOTE: This file contains only the created, compatible variables across 1988-2012. However, not all variables are available in all years.

Panel data
Files: elmps 2012 panel 98_12 v2.1.dta

NOTE: The data are constructed as panel data (wide format) directly from the repeated cross section. Variables repeat as 'variable'_# where # is round, i.e. pn_98, pn_06, pn_12 for the person numbers in 1998, 2006, and 2012 for an individual. If an individual is not present in a round, the variables for that round will be missing. In addition, an indicator variable attr_98_06_12 categorizes individuals based on when they were observed (i.e 98_06_12, 98_06 only, 06_12 only, etc.).

The elmps 2012 98_12 panel dataset includes unique individuals from all rounds, regardless of whether or not they were successfully followed in all rounds. For example, people who were added in 2012 are included, but have missing data for their 98 and 06 variables. If you only wish to work with individuals followed across two or more rounds, use the attr_98_06_12 variable to delete the observations you do not need.

NOTE: The datasets include the expansion factors for each wave as a cross section (e.g. expan_12, expan_06, expan_98) and panel weights (panel_wt_98_06, panel_wt_98_06_12, panel_wt_06_12). Make sure to use the appropriate weight for the specific panel you are working with.

Important Variables:
hhid_#: unique household identifier, begins with two-digit round number (88, 98, 06, 12)
indid_#: unique individual identifier, begins with two-digit round number (88, 98, 06, 12)
expan_#: expansion factor for cross-sections ONLY (must be used as an analytic weight)
panel_wt_98_06: panel weight for analyzing panel data with individuals observed in 1998 and 2006
panel_wt_98_06_12: panel weight for analyzing panel data with individuals observed in 1998, 2006, and 2012
panel_wt_06_12: panel weight for analyzing panel data with individuals observed in 2006 and 2012
attr_98_06_12: categorizes individuals based on whether they were in 1998, 2006, and 2012

***Do Files*:**
The do files used to create the 2012 data from the raw data and the previous round data are available. All of the do files run in sequence based on "Master do file ELMPS v29" If you are interested in particular variable's creation, you can also search for it among the do files, but be aware that some variables are created or modified across multiple files.