**Notes on the Togo Formal Firms Survey 2015 (TFFS 2015)**

- In the interest of protecting the identity of the respondents:
    - Some questions have been omitted or recoded in the published dataset.
    - Financial variables have been rounded to the nearest 5,000 FCFA (approximately $10 USD)
- The following answer codes were possible throughout the survey:
    - -6 = Does not apply
    - -7 = Missing (coded by data entry agents)
    - -8 = Refusal
    - -9 = Does not know
- Given the weighted stratified sampling design, it is necessary to use weights in the analysis. The weighted analysis is representative of tax-paying firms in the largest urban areas in Togo.
- While there is a general weight variable that takes into account firm response to the survey (named weight in the dataset), we strongly recommend using weights specific to each variable or set of variables used in the analysis due to the relatively small sample, incomplete questionnaires, and item non-response. This can easily be done using the genweight.ado program provided.
    - The syntax of this program is genweight varlist, gen(newvar)
    - The program creates a new weight variable that is specific to the variable or variables that are included in varlist.
- To ensure smaller standard errors, we strongly recommend using the svyset and svy commands in STATA for the analysis. We recommend the following:
    - svyset [w=weight], strata(strata) fpc(fpc) singleunit(centered)
    - As aforementioned, the weight variable used should be the weight variable created for the variables that you are analyzing.
    - For example, to regress last year's sales on location and sector, one could use the following commands after converting the negative answer codes to missing:
      *genweight p11q1a p2q1 sector1, gen(regweight)*
      *svyset [w=regweight], strata(strata) fpc(fpc) singleunit(centered)*
      *svy: reg p11q1a p2q1 sector1*
- The dataset includes information from all sections of the Togo Formal Firm Survey 2015. Variable names are constructed using p to identify the section of the questionnaire from which the variable was taken and q to identify the question number in that section.
- The available documentation includes:
    - The de-identified and cleaned survey data in STATA format.
    - The questionnaire used to administer the survey
    - The surveyor manual used to train the enumerators. This includes detailed descriptions of each question.
    - The implementation report, which describes the sampling, fieldwork organization, and challenges.
    - The genweight.ado file, which enables the quick calculation of weights specific to a specific question or set of variables.
- Use of the dataset must be acknowledged using a citation which would include:

- The Identification of the Primary Investigator
- The title of the survey (including country, acronym and year of implementation)
- The survey reference number
- The source and date of download
- Example:

  Hillary C. Johnson, World Bank, Theresa Osborne, World Bank. Togo Formal Firm Survey (TFFS) 2015.  Ref. TGO_2015_TFFS_v01_M. Dataset downloaded from *[URL]* on *[date]*.