Documentation file for do-files and datasets corresponding to paper titled:

"For Better or for Worse?: Education and the Prevalence of Domestic Violence in Turkey"

by Bilge Erten and Pinar Keskin

January 20, 2017

This document describes the datasets, STATA programs and R programs used to replicate the results reported in American Economic Journal: Applied Economics version of "For Better or for Worse?: Education and the Prevalence of Domestic Violence in Turkey" by Bilge Erten and Pinar Keskin. Descriptions are provided for the tables and the figures in the order in which they appear in the paper. Please see the paper for references.

Two restricted-use datasets are required to generate all the tables and the figures in the paper:

1. 2008 National Survey on Domestic Violence against Women in Turkey (2008 NSDVW): The results for Tables 1-7, and Figures 1-3, 4a, 5 (left-side figures), and 6 are generated using this dataset.

2. 2014 Household Labor Force Survey in Turkey (2014 HLFS): The results for Figures 4b and 5 (right-side figures) are generated using this dataset.

Both datasets are obtained from the Turkish Statistical Institute. Their data availability policy prohibits the distribution of data to non-registered users. Researchers have to complete a form (http://www.turkstat.gov.tr/UstMenu/body/bilgitalep/MVKullaniciTalepFormu_ENG.pdf), and send it to bilgi@tuik.gov.tr by e-mail to access these datasets. Upon receipt of this form, the Turkish Statistical Institute sends the datasets through an ftp server electronically. You can see http://www.turkstat.gov.tr/UstMenu.do?metod=bilgiTalebi more regarding to the access application.

Below are the steps to generate the results presented in the paper:

1. In order to create the variables used in the analysis, the following do files must be placed in a subfolder called "XXX/Domestic Violence/do files" and be run in this specific order:
   a. "1a Household data cleaning.do" creates the variables from household dataset (household questionnaire) in the 2008 NSDVW data.
   b. "1b Women data cleaning.do" creates the variables from the main dataset (women questionnaire) in the 2008 NSDVW data.
   c. "1c HLFS 2014 data cleaning.do" creates the variables from the 2014 HLFS data.
   d. "1d Household data cleaning for graphs.do" creates the variables from household dataset (household questionnaire) in the 2008 NSDVW data that are used in graphical analysis.

2. Once the data is cleaned, Figure 1 is generated using "2a McCrary_test.R" file. Figures 2-6 are generated using "2b Graphs.do" file. The additional do files ("rdprog_cols1_se2.do", "rdprog_cols1_se2_cutoff_reverse.do", "rdprog_cols1_se2_mod.do") are used to generate the RD plots. The program "rdprog_cols1_se2.do" has been created by Meyersson (2014).

3. The results in Tables 1-7 and Online Appendix Tables A1-A13 are generated by "3a Tables.do". Finally, "3b Tables_simes.do" provides the p-values adjusted for multiple-hypothesis testing using Simes adjustment.

References:

Meyersson, Erik "Islamic Rule and the Empowerment of the Poor and Pious", *Econometrica*, Vol. 82, No. 1 (January, 2014), 229–269.