Replication Instructions for Empirical Application in "Inverse probability tilting for moment condition models with missing data"

by Bryan Graham, Cristine Pinto and Daniel Egel

The data used in the empirical application presented in Section 4 of the paper was downloaded from the Bureau of Labor Statistics National Longitudinal Surveys (NLS) Web page

`https://www.nlsinfo.org/investigator/pages/login.jsp`

in the summer of 2010. This download was in the form of a STATA dictionary file (`NLSY_BlkWhtGap.DCT`). This file is included in the archive of replication files.

To replicate the empirical results place the `NLSY_BlkWhtGap.DCT` in a "SOURCE_DATA" directory. Place the included STATA Do file `NLSY_BlkWhtGap_SampleCreate.DO` in a "DO_FILES" directory. Finally create a third "WRITE_DATA" directory. Open `NLSY_BlkWhtGap_SampleCreate.DO` in a text editor or STATA's own editor and change the three global string definitions with the same names as in quotation marks above to point to the created directories. These definitions appear near the top of the do file. Save the file.

Once this is done you may run the do file. It will generate two files. A STATA log file `NLSY79_Sample_Log.scml` and a tab delimited text file with the estimation sample described in the paper (`NLSY79_Sample.out`). This text file was converted into a MATLAB data file `NLSY79_Sample.mat` which should be placed in the "WRITE_DATA" directory. The STATA do file produces the estimation sample, generates some basic summary statistics and some of the estimation results. See the log file and the paper for details.

The IPT and AIPW estimation results were generated using code written in MATLAB. To reproduce these results you may run the `NLSY79_Application_Base_Summer2010.m` file from the MATLAB command prompt. This file calls various M files included in the replication bundle of files. These files should be placed in a directory that MATLAB has access to. You will also need to change a directory reference near the very top of `NLSY79_Application_Base_Summer2010.M` to point to the "WRITE_DATA" directory described above. The place where this needs to be done is flagged by a comment. The MATLAB code requires the Statistics and Optimization toolboxes in addition to the base MATLAB software. For a more detailed description of the MATLAB code see the accompanying `IPT_MATLAB_CODE_Readme.pdf` file.

Please feel free to modify the code for your own use. We do ask that you acknowledge our work by citing the original paper and explicitly mentioning the use of our replication code. This code is provided "as is", without warranty explicit or otherwise. Beyond these instructions we are unable to assist with its use. If you do find errors we would welcome you telling us. Please e-mail me at `bryan.graham@nyu.edu`.