Instrument-based estimation of full treatment effects with partial compliers

Replication package

Didier Nibbering Matthijs Oosterveen

This is the readme file for the replication package of the article "Instrument-based estimation of full treatment effects with partial compliers" by Didier Nibbering and Matthijs Oosterveen, hereafter referred to as the article. This readme file and the replication package are constructed according to the guidelines of the *Review of Economics and Statistics*.

1 Data access

The data of three of the four empirical applications in the article are publicly available, and can be downloaded via the following links:

- **The Oregon health insurance experiment**: `https://www.nber.org/research/data/oregon-health-insurance-experiment-data`
- **The LinkedIn job opportunities experiment**: `https://www.openicpsr.org/openicpsr/project/125461/version/V1/view`
- **The village-based schools experiment**: `https://www.openicpsr.org/openicpsr/project/113861/version/V1/view`

To download the data from the Oregon health insurance experiment one has to log in to NBER. Once logged in, the link automatically downloads the datasets from the various papers analysing the experiment. As clarified in the article, we use the dataset discussed by Finkelstein et al. (2016, NEJM). To download the data from the LinkedIn job opportunities experiment and the village-based schools experiment one has to log in to openICPSR. Both logins can be done with a Gmail account.

The data for the fourth application in the article is proprietary. Access was obtained by following the instructions laid out in the readme file of the original paper. This readme file together with the replication package can be downloaded via the following link:

- **The elite university education natural experiment**: `https://academic.oup.com/jeea/article/18/6/2824/5716347`

In particular, access to the proprietary data can be obtained by emailing the author of the original paper (massimo.anelli@unibocconi.it - `www.massimoanelli.com`), who will share the contact of the institutional officer, to which a formal written request should be send. We obtained access to the proprietary data with the permission to start a new research project, under the condition that the dataset would not be shared for any reason.

2 Description of the replication files

This replication package contains the following files:

- Nine do-files
    1. Dataprep_OHIE.do
    2. Dataprep_LinkedIn.do
    3. Dataprep_Village.do
    4. Dataprep_Elite.do
        * These do-files contain (copy-pasted) code from the original papers to clean the raw data and code to appropriately name and generate our instrumental, treatment, outcome, and control variables.
    5. ImbensManskiCI.do
        * This do-file contains the code that implements the Imbens and Manski (2004) approach to estimate confidence intervals.
    6. Figureassumptionsxlabels.do
        * This do-file contains the code to label the horizontal axis of the figures that report on the subsample analyses.
    7. Analyses.do
        * This do-file contains the code for all the empirical analyses.
    8. Combinefigures.do
        * This do-file contains the code to combine several figures (across applications) into the final figures of the article.
    9. Masterscript.do
        * This do-file sets all working directories and subsequently calls upon the do-files above to obtain the results in the article.
- One letter (ApprovalDataAccess_Bocconi.pdf) from Università Bocconi documenting the approval to access the proprietary dataset of the fourth empirical application "the elite university education natural experiment".

3 Variable descriptions

For a detailed description of the variables in each of the datasets of the four empirical applications in the article we refer to the readme files in the replication packages of the original papers. These replication packages can be downloaded via the links above.

4 Software version and execution

The empirical analysis was conducted with StataMP 18. The operating system to run this software was Windows 11 Home.

To replicate, for instance, the results for the Oregon health insurance experiment, one can download the data from NBER using the link above. Subsequently, in the do-file Masterscript.do one needs to specify the working directories appropriately and choose a value for the global called "EXP", which may either be equal to "OHIE", "LinkedIn", "Village", or "Elite", and should be set equal to OHIE (the default value). Then, one should run Section A of Masterscript.do to replicate the results for this application.

Downloading the data for the other two applications, or obtaining access to the data for the fourth application, allows for their replication by simply renaming the global EXP appropriately and re-running Section A of Masterscript.do.

Section B of Masterscript.do calls upon the do-file Combinefigures.do, which combines several figures (across applications) into the final figures of the article. This code relies on obtaining the results for OHIE, LinkedIn, and Elite.