Readme for "Interpreting trends in intergenerational mobility"
Swedish data

Authors: Nybom & Stuhler
Date: 2023-12-18

1. Data sources and access

The paper uses proprietary data from various administrative registers in Sweden. The data have been made available to us by Sweden's national statistical agency. Unfortunately, we are not in a position to make these data available to third parties as we are bound by the agency's conditions of access and by national legislation.

We here provide and explain the programs used to generate our samples and results. We will also cooperate in any efforts to replicate our results, including assistance in gaining permission to access the data from the agency or welcoming visitors who are interested in reviewing the data at the Swedish Institute of Social Research at Stockholm University.

2. Set up of directories and master file (master.do)

It is necessary to open the master file and change the globals “path” and “path2” to the correct directory with the raw data and the local project directory, respectively. The master do-file then automatically creates three subfolders “Do”, “Dat”, and “Log”. In your local project folder, you need to copy in all the do-files and auxiliary data sets (“Help files”). The subfolder “Do” should include the codes. The master do-file also creates a subfolder “Dat” in which the “Help files” should be placed and which also hosts temporary data sets used by the do-files. The subfolder “Log” will include log-files and all results.

3. Raw data files and help files used

As explained in more detail in section A.11 of the Online appendix, our source data set is based on a 35 percent random sample of the Swedish population born between 1932 and 1967 (the “source sample”). In addition, the source data includes links to all children and parents of this source sample. We then have demographic information, education and incomes for all these individuals (i.e. the source samples and their relatives.

The codes use the following raw micro data files:

datafil10.dta — Census data (“Folk- och bostadsräkningen”)
forsamlingfob60_m.dta — Parish and municipality of residence in 1960
forsamlingfob65_m.dta — Parish and municipality of residence in 1965
datafil01.dta — Basic demographics for the source sample and all their relatives included in any of the different data sets (year of birth, sex, etc)
datafil02.dta — Links to all (birth) children of those in the source sample
datafil03.dta — Links to the (birth) parents of those in the source sample
datafil_11_utb[YEAR].dta — The highest educational attainment of all in the source sample and their relatives measured in year [YEAR].

ar[YEAR]k.dta — Annual incomes of all in the source sample and their relatives measured in year [YEAR].

The codes also use the following auxiliary “Help files”:

slutgiltiga_reformkommuner_fob60.dta — Derived reform status by cohort and municipality (see Holmlund, 2007).
sun2000_converter.dta — A crosswalk between older and newer coding systems for highest educational attainment provided by Statistics Sweden.

4. Data cleaning

The data cleaning and preparation file “dataset.do” loads and cleans source data and creates the variables used in the analysis. A main data file “main_file.dta" is then saved as well as a file “main_file_income.dta” which stores additional information on annual incomes across ages and years.

5. Analyses

The files “descriptives.do”, “estimation.do”, and “sensitivity_analyses.do” produce all results presented or discussed in the paper. All are run automatically from the master.do file.

The file “descriptives.do” produces sample statistics and two graphs with descriptive trends (Table A.2; Figure 3; Figure A.5). The file “estimation.do” produces the main results (Table 2; Figure A.2; Figure A.6), while the file “sensitivity_analyses.do” produces robustness tests (Table A.3; Figure A.7).