README

Rising Inequality? A Stocktake of The Evidence. Australian Productivity Commission Research Paper: Chapter 5: Economic Mobility

(Modified Replication – ECON7030)

Gorkem Koroglu

SID: 520312809

2024 Semester 2

Data Introduction (Metadata)

The code name is "A6_code_Final Reserach Paper_ECON7030". The code for the final research paper consists of two main parts, (1) data preparation and processing and (2) data analysis.

Under data preparation and processing, there is a data check for setting and sorting the dataset into panel data format. Next step is to generate, clean and process categorical variables in the order of hhstate (state), hhsos (section of the state), ancob (country of birth), vaccreq (vaccination requirement), cvemp (collectiveness of the job), edhigh1 (highest education received), icagef2 (age), hgsex (sex), and hhfty (family formation. There is also processing continuous variables of interest, income and wealth.

Under data analysis, there is processing of "wave" data by only leaving 19 (2019) and 21 (2021) waves. The households are tracked by "xwaveid". The ids are counted and kept if they repeat two times in both waves in the remaining dataset. Then, the income deciles are set according to the organization in Australian Productivity Commission Report 2018 page 32. As the income deciles are set, the difference in each repeating "xwaveid" income decile position is taken, which generates the "difference" variable. The "difference" variable histogram is visualized and the difference frequency chart abd the histogram is exported.

Under A6 regression analysis section, the households who show upward mobility is collected under the dummy variable "up_mobility". The probit regression model is run and extracted. There is also a correlation check for the explanatory and control variables at the bottom. Lastly, all the dataset is saved.

Data Sources

The data source is HILDA Release 21, which can be accessed from this shared OneDrive folder: "Dataset/hilda_econ7030_cleaned.dta". The used data dictionary for the dataset is the following Melbourne Institute of Applied Economic and Social Research, HILDA Survey, HILDA Online Data Dictionary: https://hildaodd.app.unimelb.edu.au/default.aspx.

Instructions

All the variables are generated and stored under the name "hilda_econ7030_cleaned", which can be seen under the data source section. When doing the code check for the output, it is recommended to check the lines 269 and onwards (after the "Results Up Until A5 (Presentation)).