Read Me

llaszewski - Replication Project ReadMe - 2024-01-20

This documentation contains all data and code necessary to reproduce the results of the paper "The Primary Parental Investment in Children in the Contemporary USA is Education" Rosemary L. Hopcroft and David O. Martin. This code was written in R version 4.2.2.

Data Set: The data set you will need is the GSS dataset. The compressed RDS file is as follows: gss7218_R1.RDS. This data set will need to be in your working directory. Once it is, readRDS() will upload it into R.

Locate files: I will have everything needed in a /ReplProjTurnIn folder. This should be set as your working directory. The GSS data set should be added to this folder running any scripts. The order to run the scripts is as follows:

Data Naming: My mnemonic is "ReplProj", which will go at the beginning of all files so I can quickly search for any file that is related to my project. The RDSs I save will be the same name with .RDS instead. My script template is named "llaszewski-ReplProj-[key word for what was done]V1.R". This will change to V2, V3, etc… if I update a file.

Scripts Needed (In Order):

llaszewski-ReplProj-LoadingDataV1 - I loaded in the GSS Data set and saved it as an RDS.

llszewski-ReplProj-subsetyearsV1 - Created a subset that only includes observations from 2000-2010.

llaszewski-ReplProj-VariablesV1 - Removed all the unneeded variables, have 7 remaining. Also restructured any variable in the dataset that needed it. Race and Sex needed to be adjusted to only have observations of 0 or 1.

llaszewski-ReplProj-RemovePaseiV1 - Removed missing values from the Pasei10 variable. Were in the form ".i"

llaszewski-ReplProj-RemoveAge_EducV1 - Removed missing values in the variables Age in the form ".n", and Educ which were ".a"

llaszewski-ReplProj-RemoveSibs_DegreeV1 - Removed missing values from values in the variables Sibs in the form ".n" and ".d". Then Degree, ".a".

llaszewski-ReplProj-MetaDataV1 - Making all variables numeric. So descriptive statistics can be done. Added metadata for each of the variables. Also for the subset (d

llaszewski-ReplProj-DescStatsFrequenciesV1 - Made a table for degree variable frequencies. Using the prop.table() function.

llaszewski-ReplProj-DescStats_MeanSDV1 - Used a for loop to create a table of means and standard deviations for the rest of the variables. Then, rbind() to combine this table with the frequency table. Creating the full descriptive statistics table. (Table 1)

llaszewski-ReplProj-RegressionTable3V1.R - This script is where I got my numbers for Table 3. Using the lm() function, I made Model 1 and 2, for all cases.

llaszewski-ReplProj-RegressionTable4V1.R - This script is where I got my numbers for Table 4. I made Model 1 and 2, which was for cases where age > 24

Packages needed: the only package you will need to download is tidyverse. Download using Install.packages(tidyverse), and use with library(tidyverse). Only necessary for llaszewski-ReplProj-DescStats_MeanSDV1.R script.

Authors and Acknowledgements:
This project was completed by Logan Laszewski with support from Nicholas Bussberg.