Panel Data Analysis: Home Assignment 1

Comparative Social Research

Kirill Gasanovich Guliev

25 Feb 2022

knitr::opts\_chunk$set(echo = TRUE, message = FALSE, warning = FALSE,  
 fig.align = "center")  
  
path <- "C:/Users/Кирилл/Desktop/CSR2023/panel\_data/panel-data-analysis/"

library(haven)  
library(tidyverse)  
library(margins)  
library(psych)

# Problem 1. Load and subset data. Wave = 3

wvs\_data3 <- read\_dta(paste0(path, "data/soc\_hw1.dta")) %>%   
 filter(wave == 3) %>%   
 mutate(index = as.character(index),  
 wave = as.character(wave),  
 sex = as.factor(sex),  
 emp = as.factor(emp))

# Problem 2. Descriptive statistics

## Report descriptive statistics of all the variables

summary(wvs\_data3)

## index wave country satis   
## Length:76187 Length:76187 Length:76187 Min. : 1.000   
## Class :character Class :character Class :character 1st Qu.: 5.000   
## Mode :character Mode :character Mode :character Median : 7.000   
## Mean : 6.338   
## 3rd Qu.: 8.000   
## Max. :10.000   
## NA's :2734   
## chldr fsatis sex age emp   
## Min. :0.000 Min. : 1.000 0 :39542 Min. :15.00 0 :31830   
## 1st Qu.:0.000 1st Qu.: 3.000 1 :36588 1st Qu.:28.00 1 :40458   
## Median :2.000 Median : 5.000 NA's: 57 Median :39.00 NA's: 3899   
## Mean :1.843 Mean : 5.234 Mean :40.99   
## 3rd Qu.:3.000 3rd Qu.: 7.000 3rd Qu.:52.00   
## Max. :8.000 Max. :10.000 Max. :95.00   
## NA's :3754 NA's :4887 NA's :177

We see that in our filtered dataset we have 76’187 observations. We have 39’542 females and 36’588 males in our sample; there also 57 missing values concerning the question on respondent’s sex. The minimum age of the respondent is 15, the maximum is 95. The mean age in the sample is 40.99; 177 respondent’s have a missing value in their age. 40’458 respondents from the sample are employed, 31’830 are not employed. Variable on employment has the greatest number of missing values across the wave - 3’899.

The mean value of life satisfaction (satis variable) is 6.338. The median is 7 - which means that 50% of the sample estimate their life satisfaction above this point, and 50% - below. The financial satisfaction seems to be the most balanced variable - a half of sample estimated it above 5/10, and another half - below.

The maximum number of children in a family is 8, minimum is 0; 3’754 values are missing. 50% of the sample have more than 2 children.