Exercise 4

Social Inequality

Deadline: Please upload your assignment by Tuesday (**March 5, 5 p.m.)** Upload one file only (pdf). Include the R-code into the Appendix. Do not forget to cite the data. Further, integrate at least one reading into your discussion (see moodle). Use APA-style when you cite the reading(s). Label all figures and tables.

Data preparation

Please download the PISA-data from 2018 from moodle.[1] Select a country of your choice. Construct the new variables as listed in the Appendix.

Exercise 4.1

Formulate a hypothesis that deals with the influence of parental social status on children's (expected) social status. The readings on moodle may be helpful to generate a testable hypothesis. [about 250 words]

Exercise 4.2

Estimate the following OLS-regression models with expected social status as dependent variable (ISEI). Does parental background influence own social status? Is father or mother's social status more important for own (expected) social status?

MODEL01: PISEI+AGE+GENDER
MODEL02: PISEI+MISEI+AGE+GENDER

Exercise 4.3

Do patterns differ by gender? Estimate the following models and interpret the results.

MODEL03 (girls): PISEI+MISEI+AGE
MODEL03 (boys): PISEI+MISEI+AGE

Exercise 4.4 [Optional]

There is a large body of literature on the so-called "immigrant optimism hypothesis". This hypothesis states that immigrant children have higher educational and occupational aspirations once controlling for parental background. Does this hypothesis hold for your country of choice? Choose a country with a sufficiently large migrant population for this part of the analysis (e.g., France, Germany, US).

MODEL04: MIG+AGE+GENDER
MODEL05: MIG+AGE+GENDER+PISEI

[1] You may also download the data directly from the PISA-homepage. https://www.oecd.org/pisa/data/ Note that the data only come in SAS or SPSS format. However, both formats can be read into R using the package "haven" which we used before to read in data in STATA-format. Note that the student file is very large (489MB) which may exceed the abilities of your PC.

Variables in PISA 2018

Variable	Variable Label	
CNT	Country	ALB: Austria etc. (see next page)
ST004D01T	Gender	1=Female 2=Male
AGE	Age	15 16
IMMIG	Migration background	1=Native 2=Second generation 3=First generation
BSMJ	Expected ISEI	0 …100
BMMJ1	ISEI Mother	0 …100
BFMJ2	ISEI Father	0 …100

New Variable

Name	Variable Label	Realizations	Type
AGE	Age	15 16	Factor
GENDER	Gender	Boys Girls	Factor
MIG	Migration Background	Native Second generation First generation	Factor
ISEI	Expected ISEI	0..100	Numeric
MISEI	ISEI Mother	0..100	Numeric
FISEI	ISEI Father	0..100	Numeric

ALB	Albania
ARE	United Arab Emirates
ARG	Argentina
AUS	Australia
AUT	Austria
BEL	Belgium
BGR	Bulgaria
BIH	Bosnia and Herzegovina
BLR	Belarus
BRA	Brazil
BRN	Brunei Darussalam
CAN	Canada
CHE	Switzerland
CHL	Chile
COL	Colombia
CRI	Costa Rica
CZE	Czech Republic
DEU	Germany
DNK	Denmark
DOM	Dominican Republic
ESP	Spain
EST	Estonia
FIN	Finland
FRA	France
GBR	United Kingdom
GEO	Georgia
GRC	Greece
HKG	Hong Kong
HRV	Croatia
HUN	Hungary
IDN	Indonesia
IRL	Ireland
ISL	Iceland
ISR	Israel
ITA	Italy
JOR	Jordan
JPN	Japan
KAZ	Kazakhstan
KOR	Korea
KSV	Kosovo
LBN	Lebanon
LTU	Lithuania
LUX	Luxembourg
LVA	Latvia
MAC	Macao
MAR	Morocco
MDA	Moldova
MEX	Mexico
MKD	North Macedonia
MLT	Malta
MNE	Montenegro
MYS	Malaysia
NLD	Netherlands
NOR	Norway
NZL	New Zealand
PAN	Panama
PER	Peru
PHL	Philippines
POL	Poland
PRT	Portugal
QAT	Qatar
QAZ	Baku (Azerbaijan)
QCI	B-S-J-Z (China)
QCY	Cyprus
QMC	Moscow City (RUS)
QMR	Moscow Region (RUS)
QRT	Tatarstan (RUS)
ROU	Romania
RUS	Russian Federation
SAU	Saudi Arabia
SGP	Singapore
SRB	Serbia
SVK	Slovak Republic
SVN	Slovenia
SWE	Sweden
TAP	Chinese Taipei
THA	Thailand
TUR	Turkey
UKR	Ukraine
URY	Uruguay
USA	United States
VNM	Vietnam