Exercise 3

Applied Longitudinal Data Analysis

Deadline: Upload your assignment by Monday (February 26, 5 p.m.) Upload **one** file only (.pdf). Include the R-code into the Appendix. Please cite data and readings (and preferably also R-packages). Label and number figures and tables.

Exercise 3.1

DATA01:   Load the individual "female file" of the DHS for a country of your choice.
DATA02:   Delete missing information on your key variables of interest.
DATA_URBAN: Urban population
DATA_RURAL: Rural population
Construct the variables that you require for the analysis. Provide the sample statistics by urbanity and (briefly) describe the table.

Exercise 3.2

Do you expect a "boy preference" for your country of choice? Do you expect that there is stronger relationship in rural than in urban areas? Formulate testable hypotheses. Buttress the hypotheses. The readings on moodle may be helpful in this context.

Exercise 3.3

Estimate Cox models and control for the covariates as listed below. Describe and interpret all hazard ratios (HR). Do the results support your hypothesis?

Model 1:   GENKIDS01
Model 2:   GENKIDS01+COHORT+EDU+AGEKID1+URBAN

Exercise 3.4

Estimate separate models for urban and rural settings. Describe and interpret the results in a concise manner. You do not have to describe each single HR.

Model 3(URBAN): GENKIDS01+COHORT+EDU+AGEKID1
Model 3(RURAL): GENKIDS01+COHORT+EDU+AGEKID1

Exercise 3.5

Omitted variable bias is a serious shortcoming of regression analysis. This also applies to event history analysis. Do you think that the results for GENKIDS01 may be biased? Explain!

Raw Data

Name	Realization	Variable label	Class
v212	10	Age at first childbirth	haven
	…		
	49		
v010	1965	Year of birth	haven
	…		
	2003		
	NA Refusal		
v007	2015	Year of interview	haven
	2016		
v025	1: urban	Rural or urban	haven
	2: rural		
v106	0: No education	Level of education	haven
	1: Primary education		
	2: Secondary education		
	3: Higher education		

Generated Variables

Name	Realization	Variable label	Class
TIME	0	Duration until third birth/censoring	Numeric
	..		
	14		
EVENT	0	Variable that equals 1 for cases with event (third child) and 0 otherwise	Numeric
	1		
URBAN	Urban	Urbanity	Factor
	Rural		
GENKIDS01	Boys	Gender mix of previous children	Factor
	Girls		
	Mix		
EDU	Please use a meaningful classification	Level of education. Please use a meaningful classification	Factor
COHORT	Please use a meaningful classification	Birth cohort.	Factor
AGEKID1	Please use a meaningful classification	Age at first childbirth.	Factor