

Project Assignment: Business Analytics and Econometrics 2021

Open the file Assign???.dta containing individual cross-section data about a large number of Germans. The variables are defined in the data file. The aim is to analyze the emotional stability of the individuals.

1. Describe the data with appropriate statistical measures. Give the scale type of all variables. Give the complete correlation matrix for all variables.

2. Check the data for outliers or errors. If necessary, eliminate or correct them cautiously and redo the descriptive analysis.

Hint: Do NOT eliminate or correct the observations with LHHINC = 0 (because transfer incomes are not included in that variable).

3. Consider the variable DOCVIS. Shortly explain why a linear relation between doctor visits (as indicator for illness) and emotional stability may be inadequate. Add a regressor to your data set to account for this potential non-linearity.

4. Provide the OLS estimates of the linear regression model with emotional stability as endogenous variable and all available regressors. Include the beta coefficients.

5. Interpret the parameters (with exact values and units) of LIFESAT and CHILDO_1. Interpret the age parameters (with exact values and units) for age = 20 and age = 60.

6. Derive mathematically whether the relation between age and emotional stability is monotonous for the individuals in your data set. Interpret your result.

7. Are we able to interpret the age parameters as causal effect of age on emotional stability? Shortly explain why (not).

8. Give the 3 variables with the highest economic significance and the 3 variables with the highest statistical significance. Shortly interpret your results for LIFESAT.

9. Generate the interaction term between MALE and NOT_EMPLOYED and re-estimate your linear regression model from 4. including the interaction term. Interpret the parameter of that interaction term.

10. Re-estimate your linear regression model from 4. excluding LHHINC. Compare this result with the results from 4. Shortly explain potential differences.

11. Print and sign the following declaration:

I honestly declare that the empirical work of this project was performed by myself. Only my own empirical results have been used to answer the questions on the project sheet. My cooperation with other students or outside parties was restricted to the discussion of methodological approaches, but did not include the implementation of the used approaches or the interpretation of my specific empirical results.

12. Document your analysis in a pdf file of at most 10 pages (individually, not in groups!) and send this file until Sun, 07.11.2021, 23:59 h to my Email address: Jensen@stat-econ.uni-kiel.de.

Good luck!