Introduction to Stata

Lucas van der Velde

October 2018

We would like to estimate the determinants of working full time from a random sample of Polish workers. The raw data was obtained from the International Social Survey Programme (ISSP), which gathers data from several countries across the globe.

In particular we would like to estimate the model

$$P(y = \text{employed}) = f(age,\ gender, education\,level,\ household\,characteristics) \tag{1}$$

1. Load the database into Stata.
2. Create a dummy variable (emp) that equals 1 if the individual is working and 0 otherwise.
3. Create the dependent variables: a dummy for gender, a dummy for each education level (primary, secondary, tertiary) and one for working partner.
4. Keep only observations of people aged between 25 and 55 (*prime age* workers)
5. Provide descriptive statistics (age, gender, education level) of people employed and unemployed (*Hint:* Remember the variable type (continuous, categorical) before deciding which statistic to show.)
6. Estimate two models where the dependent variable is emp: a linear probability model (linear regression) and a probit model. Interpret the results.
7. Save the output to Excel