_____

Problem Set on Intro to Stata

Prepared by Rosaria Vega Pansini and Maria Emma Santos

Using the Buthan dataset perform the following operations:

1. Start Stata and change the working directory to c:\buthan. Open the do file editor and start a new do file in which you set a new amount of memory and clear any previous dataset open and then you start a log file with the name introstata. Save the do file with the name ex1.do.

2. Open the Buthan dataset, sort the data by houseid and save the dataset without creating a new one.

3. Create two different datasets one containing information only for people living in the urban area and the second with information for people living in the rural area. Then append again these two datasets in order to obtain the original one.

4. How many observations are in the sample? Create a codebook for the variables househid, area, poor, sex, relationship with the household head and marital status.

5. How many man and women are in the sample? And how many women are married?

6. Create single dummy variables corresponding to the different level of education (b12q12). Assume that grades 1-7 correspond to primary school education, grades 8-12 correspond to secondary education. For b12q12>=13 consider all of them as 'college' education.

7. Derive frequency values for men and women for each level of education.

8. Estimate a simple linear regression model in which per capita real consumption depend on the household size, the sex of the respondent, the level of education, the marital status, the location of the household. Comment on the results of the estimation.

9. Estimate the same model estimated at point 8) using robust standard errors. Display also the post estimation command.

10. Now estimate a logit model I which the condition 'being poor' is dependent from the same independent variables considered in the model estimates at point 8).

11. Construct and display two matrices corresponding respectively to the beta and the variance-covariance matrix derived in the logistic estimation.

12. Rerun the do file, close the log file and save the dataset with the changes made during the exercise.