

## **BSM946 - Workshop 3: Discrete Choice Models**

### **Important Information:**

The workshops provided in this module are designed to help you in the practical application of the work covered in the lectures, and to support you through this step. The tutorials should ideally be attempted in advance of the session, allowing you to raise any questions or issues in the workshop. The questions outlined below are likely to take more time than the workshops run for, so attempting the questions before the sessions will allow you to receive the most greatest benefit from these sessions.

No answers will be provided for the workshop questions. Guidance will be given in the workshops, and the staff associated with the module will happily assist in office hours etc. The aim is for you to gain experience of the practical application of the methods and to start thinking about the key questions identified in the material. I would encourage you to make notes as you complete the workshop so you can reflect on these notes and raise any issues for clarification.

### **Discrete Choice Models**

- 1) Load "crime data.dta" dataset into STATA.
- 2) Create a variable that measures whether an individual has ever been arrested.
- 3) Fit a linear regression to see if being arrested is linked to the individuals employment history, the total time they have spent in prison, their income and their race.
- 4) What do the results suggest?
- 5) What is the likelihood of being arrested in 1986 if the individual's income is 250, they have spent no time in prison and the individual identifies as Hispanic? Does this figure make sense?
- 6) Why might the above not be a very good regression?
- 7) Run the same regression as above but with either a probit or a logit.
- 8) What do the results suggest?
- 9) What is the likelihood of being arrested in 1986 if the individual's income is 250, they have spent no time in prison and the individual identifies as Hispanic? Does this figure make sense?
- 10) Compare the goodness-of-fit for both the linear model and the probit/logit that you have fitted. Which gives a better fit?