Teaching Statement: **Regression with categorical dependent variables using Stata** (Fiona Carmichael)

The session is on regression analysis with categorical dependent variables using the Stata software. It will include practical exercises using data series. Example data will be used to explore the most commonly used regression models for categorical outcomes: binary logit and probit and subject to time, ordinal logit and probit and multinomial logit. The emphasis in the practical component is on the application of appropriate techniques and interpreting results using secondary data. The course assumes that participants have prior knowledge of common commands in Stata to organise and handle data and undertake standard regression techniques. By the end of the session participants should have a good understanding of how to run their own regressions with categorical dependent variables using Stata and how to interpret their results.Throughout the focus is on using Stata to run the appropriate regression models and interpret the results. The format of the module and the practical work will make extensive use of the modelling and data in Scott Long and Jeremy Freese (2014) Regression Models for Categorical Dependent Variables Using Stata, Third Edition, College Station, TX:Stata Press.

Learning Outcomes/Objectives

On completing this session, students will be able to:

- Understand when to use binary/ordered/multinomial logistic/probit regression models
- Understand how these approaches differ from OLS, and understand the alternative model specifications and extensions to multi-category dependent variables with and without an 'order'.
- Develop data management skills and basic programming skills (writing code) using statistical software such as Stata to operationalise these models.
- Understand and interpret the outputs and findings of these models in both statistical and substantive terms (i.e. one that relates to the research question).
- Develop research questions and apply appropriate modelling techniques to address them according to the nature of the categorical outcome variables.

Key text:

Scott Long and Jeremy Freese (2014) Regression Models for Categorical Dependent Variables Using Stata, Third Edition, College Station, TX:Stata Press

Summary

This session introduces you to regression modelling when the dependent variable is categorical. Categorical variables can be binary, ordinal, or nominal (including count data). When the dependent variable is categorical, ordinary least squares linear regression is no longer an appropriate modeling strategy. Instead non-linear estimation methods are used. These include binomial logit and probit, ordered logit and probit and multinomial logit as well as models for count data such as the Poisson regression. The appropriate model to use will vary depending on whether the categorical variable is binary or has more than two categories and whether the categories are ordered or not.