# 9th ECPR Summer School in Methods and Techniques, 24 July to 9 August 2014
# University of Ljubljana, Slovenia
# Course Description Form [1-week course, 15 hours]

**Course title**

## C10. Applied Multilevel Modelling

**Instructor details**

First name, last name: **Kim Mannemar Sønderskov**
Department/Unit: Department of Political Science
Institution: Aarhus University
Full postal address for ECPR correspondence: Bartholins Allé 7, 8000 Aarhus C, Denmark
Phone: +45 8716 5682
Fax: +45 8613 9839
E-mail: ks@ps.au.dk

**Short Bio**

Kim Mannemar Sønderskov is an associate professor at the Department of Political Science, Aarhus University, Denmark, from which he also earned his PhD. His fields of interests include political behaviour and attitudes, neighbourhood effects, and economies of scale in public organizations. His works have appeared in European Sociological Review, Political Studies, Rationality and Society and Public Choice, among others. He has written a textbook on statistical analysis using Stata (in Danish, forthcoming in English), and has taught multiple courses on applied statistics at the BA, MA and PhD level.

**Prerequisite knowledge**

*Note from the Academic Convenors to prospective participants: by registering to this course, you certify that you possess the prerequisite knowledge that is requested to be able to follow this course. The instructor will not teach again these prerequisite items. If you doubt whether you possess that knowledge to a sufficient extent, we suggest you contact the instructor before you proceed to your registration.*

Students must have a solid grasp of linear and dichotomous (logit/probit) regression analysis and the assumptions behinds these techniques. They must also be familiar with Stata (simple data management and analysis) and have a basic understanding of maximum likelihood estimation.

The following books are useful refreshers on these issues.
Achen, C.H. (1982). Interpreting and using regression. Sage Publishers.
Berry, W.D. (1993). Understanding regression assumptions. Sage Publishers.
Long, J.S., & J. Freese, (2006). Regression models for categorical dependent variables using Stata. Stata Press.
Kohler, U., & F. Kreuter (2012). Data analysis using Stata. Third edition. Stata Press.

**Short course outline**

The course equips the students with a thorough understanding of multilevel modelling and its application to political science research. The course covers linear and logistic hierarchical models with fixed and random intercepts as well as random slopes. It also deals with non-hierarchal models such as crossed random effects and it will touch upon techniques to model time in multilevel

modelling. The course is an applied course that focuses on how to analyse multilevel data using Stata and how to interpret the results. In the lectures and the lab sessions we will use examples and data from research in political science and discuss the applied estimation strategy and the results. By the end of the course the students will be able to analyse a variety of research questions using multilevel modelling and to interpret and report the results in accordance with widely accepted conventions.

**Long course outline**

Researchers are often faced with observations that are nested in groups. Voters in countries, municipalities or voting districts, citizens in neighbourhoods or organizations, children in classes, schools and school districts, to name just a few obvious examples of such observations and groups. The nested data structure may imply that the observations are not independent in the sense that observations belonging to the same group tend to resemble each other with respect to the dependent variable. Common exposure, contagion and/or self-selection may give rise to this similarity. Ignoring the dependency may lead to biased inference both in terms of flawed standard errors and coefficients. Multilevel modelling provide useful tools to handle the dependence and it also provides tools to exploit the data structure and obtain answers to relevant political science question that would be difficult to answer otherwise. It is thus no wonder that the use of multilevel modelling within political science and related fields is growing rapidly these years.

The course equips the students with a thorough understanding of multilevel modelling and its application to political science research. The full content is given below, but in brief, the course covers linear and logistic hierarchical models with fixed and random intercepts as well as random slopes. It also covers non-hierarchal models such as crossed random effects and it will touch upon techniques to include time in multilevel modelling. We will also dwell upon cross level interactions. The course is an applied course; we will focus on the challenges one face and the steps one goes through when writing a political science research paper using multilevel modelling. To be more specific, we will focus on how to identify dependency, how to deal with it and exploit it through multilevel modelling, how to interpret the results and how/what to report. This applied focus implies that we will not spend much time on the mathematical foundations of multilevel modelling. Please also note that multilevel modelling is a vast topic with an ever-expanding set of tools. Thus, the course cannot cover all aspects and all specific applications.

We will be using Stata as statistical software throughout the course. Stata allow the user to estimate a wide set of multilevel models and it also provides useful tools to interpret and visualize the results. Students are expected to be familiar with basic data management and analyses in Stata. Students are also expected to read the compulsory reading before classes and they should expect to spend an hour or two every day finishing the exercises that were introduced in the lab sessions. We discuss the exercises the following day and students are expected to participate in these discussions.

Day-to-day content:

Day 1: LECTURE: Why and when multilevel modelling. Within and between group variance. LAB: Preparing data for multilevel modelling in Stata. Descriptives. Tests and statistics (intraclass correlation, within and between variance, rho).

Day 2: LECTURE: Pooled regression vs. multilevel modelling. Linear multilevel modelling with random and fixed intercepts (random and fixed effects models). Explained and unexplained variance. Assumptions and diagnostics I. LAB: Estimating linear random and fixed effects models. Diagnostics (distribution of residuals, influential groups).

Day 3: LECTURE: Assumptions and diagnostics II (the exogeneity assumption). Random vs. fixed effects and a hybrid model. Random coefficients models. LAB: Estimating linear multilevel models. Specification tests (Hausman and similar tests).

Day 4: LECTURE: Cross-level interactions. Dichotomous dependent variables in multilevel modelling. LAB: Testing and visualizing cross level interactions. Estimating multilevel models with dichotomous dependent variables.

Day 5: LAB: Beyond the hierarchical two-level model: multiple levels, non-hierarchical models, and longitudinal models. LAB: Estimating three-level models, crossed random effects and longitudinal models.

**Day-to-day schedule (Monday 28 July – Friday 1 August)**
Each day = 3 contact hours, split in two 90' sessions
See above

| | **Topic(s)** | **Details [NB : incl. timing of lecture v/s lab or fieldwork etc. hours]** |
|---|---|---|
| Day 1 | Why and when multilevel modelling | Equal split between lecture and lab |
| Day 2 | Linear multilevel modelling with random and fixed intercepts | Equal split |
| Day 3 | Random vs. fixed effects and a hybrid model | Equal split |
| Day 4 | Cross-level interactions and dichotomous dependent variables | Equal split |
| Day 5 | Beyond the hierarchical two-level model | Equal split |

**Day-to-day reading list**

We will use the following textbook: Rabe-Hesketh & Skrondal (2012). Multilevel and Longitudinal Modelling Using Stata. Third edition. College Station: Stata Press.

| | **Readings (please list at least the compulsory reading for the scheduled day)** |
|---|---|
| Day 1 | Rabe-Hesketh & Skrondal (2012): [Vol. 1] pp. 1-7, 73-115.<br><br>Steenbergen, M.R., & B.S Jones (2002). Modelling multilevel data structures. American Journal of Political Science, 46(1): 218-237.<br><br>Hjerm, M. (2009). Anti-immigrant attitudes and cross-municipal variation in the proportion of immigrants. Acta Sociologica, 52(1), 47-62. |
| Day 2 | Rabe-Hesketh & Skrondal (2012) [Vol. 1] pp. 123-172<br><br>Franzese, R, & C. Kam. Modeling and interpreting interactive hypotheses in regression analysis. University of Michigan Press, 2009. pp. 115-130.<br><br>Stegmueller, D. (2013). How Many Countries for Multilevel Modeling? A Comparison of Frequentist and Bayesian Approaches. American Journal of Political Science. 57(3): 748-761.<br><br>van der Meer, T., M Te Grotenhuis, & B. Pelzer. (2010). Influential Cases in Multilevel Modeling. A Methodological Comment. American Sociological Review, 75(1), 173-178. |

| | |
|---|---|
| Day 3 | Rabe-Hesketh & Skrondal (2012) [Vol. 1] pp. (123-172), 181-216.<br><br>Pittau, M.G., R. Zelli & A. Gelman (2010). Economic disparities and life satisfaction in European regions. Social Indicators Research, 96(2), 339-361. |
| Day 4 | Rabe-Hesketh & Skrondal (2012) [Vol. 2] pp. 501-563.<br><br>Empirical examples [TBA] |
| Day 5 | Rabe-Hesketh & Skrondal (2012) [Vol. 1] pp. 385-460.<br><br>Dinesen, P.T., & K.M. Sønderskov (2012). Trust in a time of increasing diversity: On the relationship between ethnic heterogeneity and social trust in Denmark from 1979 until today. Scandinavian Political Studies, 35(4), 273-294. |

**Software and hardware requirements**

***Software programme***
Stata version 12 or newer

***Hardware requirements***
No specific requirements

**Literature**

Gelman, A. & J. Hill (2007). Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.

Singer, J.D., & J.B. Willett (2003). Applied longitudinal data analysis: Modeling change and event occurrence. Oxford University Press.

Snijders, T.A.B., & R.J Bosker. (2012). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling, Second Edition. Sage Publishers.

**Lecture room requirement**
Lecture room: Projector and WIFI
Lab sessions: Computer room with projector and WIFI

**Preferred time slots**
Prefer mornings, but it is not important