Introduction to Multilevel Modeling

Workshop, 3rd term 2015-2016

5-7 April 2016
Seminar room, Villa Malafrasca / Seminar room, Villa Paola

Organised by Alexander H. Trechsel, Fabrizio Bernardi and Marta Fraile Maldonado, with Anna Kandyla and Irene Sánchez Vítores

Guest speaker: Héctor Cebolla Boado (UNED)

Please register online

Contact: Mariana.Spratley@eui.eu

Abstract
Since the late eighties, multilevel modeling is an extended technique in the fields of Medicine, Biology, Sociology, Social Psychology, Economics or Political Science. The development of computers - with new software and hardware which reduce the cost of its implementation -, the apparition of new data prepared to run multilevel analysis, and its conceptual and analytical appeal, have facilitated the extension of this technique. Multilevel analysis has become useful when dealing with data structured in several levels. In other words, multilevel modeling is appropriate for research designs where the smallest units of analysis can be grouped along hierarchically organized categories. For instance, the respondents of a survey can be analysed according to their individual traits, such as age or educational attainment, but they can also be looked at considering the region where they live. Multilevel modeling permits to take into consideration the individual observation in the environment that it has developed and include it in the model's estimates.

Schedule
The workshop takes place on **Tuesday 5 April, Wednesday 6 April and Thursday 7 April.**
On the afternoon of Thursday 7 April, a special session on BYOR is organised as part of the workshop (15:00-17:00).

Please see the full schedule below for details on times and rooms.

Tuesday 5 April	11:00 – 13:00	Seminar Room, Villa Malafrasca
	Lunch break	
	15:00 – 17:00	
Wednesday 6 April	11:00 – 13:00	Seminar Room, Villa Paola
	Lunch break	
	15:00 – 17:00	
Thursday 7 April	11:00 – 13:00	Seminar Room, Villa Paola
	Lunch break	
	15:00 – 17:00	BYOR session

Last updated 04.04.2016

Overview

This crash course introduces to the basic logic of multilevel analysis and its application to the study of schools as agents in the (re)production of educational inequality. All the examples and exercises provided are estimated using STATA (11 onwards).

The focus of this course is basically practical: research questions are to be translated into multilevel empirical strategies. The course gives priority to the graphical interpretation of the results. The discussion of your ongoing multilevel research is welcomed.

1. *Introduction to the multilevel logic and the use of hierarchical data. Substantive and technical consequences of ignoring any multilevel structure. Multilevel strategies: fixed effects. Requisites for random effects estimations.*
   The aim of this block is to explain the problems that arise from ignoring the clustering of individuals in groups (schools, countries): biased standard errors, unreliable contrasts and waste of theoretical potential. Alternative solutions are proposed to the estimation of multilevel random term regressions, including fixed effects.

2. *The random intercept model. The empty mode. The measurement of variation at different levels. The intra class correlation coefficient. Completing the model specification.*
   Understanding the basic hierarchical model with a random intercept and no predictors. Expanding the empty model to contrast theoretical arguments using individual level independent variables as fixed effects and level-2 explanations.

3. *Slopes as outcomes: random slope models. Centering predictors.*
   Allowing individual level independent variables to vary across clusters. Centering or not centering? Cross-level interactions.

4. *Evaluating our models: residuals, fit of the models and shrinkage.*
   To gain confidence in our results, we calculate and look at the distribution of residuals. Different tools for measuring the amount of variation explained at different levels. Increasing our awareness about how multilevel corrects group effects.

5. *Extensions to other research contexts.*
   This section presents the basic features of models estimating effects on binary dependent variables and three level research environment.

Materials

A. References for the Study of Multilevel Modeling

There is a growing number of manuals on multilevel modeling that you can use to study all technical details. My advice for someone with a proficient knowledge of regression analysis is Hox (2010). The others listed here are considered equally classic references

Allison, P. D. 2009. *Fixed Effects Regression Models*. Newbury Park, CA: Sage.

Bickel, Rorbert (2007): *Multilevel Analysis for Applied Research. It's Just Regression*!, New York, London: Guilford Press.

De Leeuw, J. and E. Meijer (eds.) (2008), *Handbook of Multilevel Analysis*. Berkeley: University of California Press.

Gelman A. and J. Hill, (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. CUP.

Goldstein, Harvey (2010) *Multilevel Statistical Models*. London: Arnold

* Hox, J. (2010). *Multilevel analysis: Techniques and applications*. New York: Routledge.

Kreft, I., and De Leeuw, J. (1998). *Introducing multilevel modeling*. London: Sage.

Raudensuh, S. W., and Bryk, A.S. (2002) Hierarchical Linear Models: Applications and Data Analysis Methods, Second Edition. Newbury Park, CA: Sage.

Snijders, Tom and Roel Bosker (1999), *Multilevel Analysis: An Introduction to Basic and Applied Multilevel Analysis*. Sage.

B. References on Software

General packages are now able to estimate and help with interpretation of multilevel modeling. Although the course uses STATA, other software (SPSS, SAS and R) can also be used. Of course, the traditional multilevel-specific software remains an option for multilevel modeling.

Rabe-Hesketh, S. and Anders Skrondal (2005) *Multilevel and Longitudinal Modelling Using STATA*: Stata Press.

Albright, Jeremy J. and Dani M. Marinova (2010) *Estimating Multilevel Models using SPSS, Stata, SAS, and R*. Online, document. University of Indiana.

Littell, R. C., Milliken, G. A., Stroup, W. A., Wolnger, R. D., and Schabenberger, O. (2006). *SAS for Mixed Models*, Cary, NC: SAS Institute.

Other software alternatives: *sometimes the smart option (if models do not converge, if large samples, if seeking advanced estimation options).*

HLM: is still not very flexible and requires to prepare you data using another software (http://www.ssicentral.com/hlm/).

MLWin: the best option for advanced multilevel modeling with a large number of estimation options (including bootstrapping). Its website provides fantastic tools for self-learning and teaching (http://www.cmm.bristol.ac.uk).

WinBUGS: an option if one seeks to do Bayesian estimation (http://www.mrcbsu. cam.ac.uk/bugs/winbugs/contents.shtml

C. Specific Topics: Centering & Sample Sizes

Given the importance of centering and the large debate existing around the issue of sample sizes, some specific references are recommended.

Kreft, Ita, Jan Leuuw and Aiken (1995) "The effect of different forms of centering in hierarchical linear models" *Multivariate Behavioural Research* 30(1): 1-21

Maas, C and Hox, T (2005) "Sufficient sample sizes for multilevel modeling" *Methodology*, 1: 86-92.

Raudenbush, Stephen W. (1989). "Centering predictors in multilevel analysis: choices and consequences". *Multilevel modelling newsletter*, 2 (1): 10-12.

Snijders, T.A.B., and Bosker, R.J. (1993). "Standard errors and sample sizes for two-level research". *J. Educational Statist*., 18, 237-259.