ANALYSING COMPARATIVE LONGITUDINAL SURVEY DATA USING MULTILEVEL MODELS

Dr Malcolm Fairbrother

School of Geographical Sciences · Centre for Multilevel Modelling · Cabot Institute
University of Bristol

Short bio – Malcom Fairbrother

Malcolm Fairbrother (http://seis.bris.ac.uk/~ggmhf/) is Reader in Global Policy and Politics and an associate member of the Centre for Multilevel Modelling at the University of Bristol. He holds a PhD in sociology from the University of California, Berkeley, and has been a visiting researcher at institutions in Canada, Mexico, Sweden, Spain, and Italy. His methodological research has concentrated on the use of multilevel models in analysing comparative survey data. In his substantive research, he studies social trust, political trust, and public attitudes towards environmental degradation and protection. He has also used qualitative comparative-historical methods to investigate the political origins of economic globalisation. His work has appeared in journals such as the *European Sociological Review*, *American Journal of Sociology*, *Political Science Research and Methods*, and the *British Journal of Political Science*. He teaches environmental policy and politics, political economy, and research methods in Bristol's School of Geographical Sciences.

Course description

Many surveys spanning multiple countries--or many regions within a single country--are now being fielded multiple times over the course of years or even decades. Examples include the European Social Survey, International Social Survey Programme, EU Statistics on Income and Living Conditions, and (across states) the U.S. and German General Social Surveys. The range of topics that can be studied using data from these surveys is extremely broad: from health to religiosity to social and political attitudes and behaviours. This course will show students how to analyse these comparative longitudinal survey data (CLSD) using multilevel models that exploit any or all of three different kinds of variation: differences between countries, change within countries over time, and variation across individuals.

We will begin by considering the structure of CLSD, and then what fixed effects and random effects (multilevel) models each reveal about the variation between and within groups in data characterised by clustering. We will see how CLSD can be understood as doubly hierarchical (or clustered), and therefore how we can analyse them with models partitioning between and within effects. We will also consider the capabilities of societal growth curves, and the insights that can be gained from models with random (country-specific) slopes. The course will emphasise the use of graphical analysis throughout, and note some risks that analysts of CLSD need to avoid.

Software

In the lab sessions we will use the open-source R software and environment for statistical computing, including some easily installable add-on packages (lme4, MCMCglmm, plm, WDI). Students are welcome to work on their own laptop, ideally with R pre-installed (from https://cran.r-project.org/).

Prerequisites

Basic understanding of multilevel models, and some experience of designing and fitting them. Students unfamiliar with multilevel/random effects models, but who are strongly familiar with other techniques for panel data analysis, are also welcome to join. Students unfamiliar with multilevel/random effects models are encouraged to enrol first in the course "Practical multilevel modelling with Stata" taught by Leonardo Grilli and Carla Rampichini in Week 1 of the summer school. Students are not required to have any prior familiarity with R.

Schedule

6 July, 2017

Time	Topic
09.00-10:45	Lecture: - understanding and characterising the structure of CLSD - examples of applications - fixed effects and random effects (multilevel) models - between and within relationships - the REWB specification
10:45-11:00	Break
11:00-13.00	Lab: - merging survey and macro data - mean-centering - graphing CLSD - fitting fixed effects and random effects models to two-level data - using the lme4 package

7 July, 2017

Time	Topic
09.00-10:45	Lecture: - cross-sectional and longitudinal relationships in CLSD - societal growth curves - random slopes - cautionary notes - generalized linear mixed models
10:45-11:00	Break
11:00-13.00	Lab: - fitting multilevel models to CLSD - working with European Values Survey data - using the MCMCglmm package

References

Course largely based on:

Bell, Andrew, and Kelvyn Jones. 2015. "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data." *Political Science Research and Methods* 3[1]: 133–153. doi:10.1017/psrm.2014.7.

Enders, Craig K., and Davood Tofighi. 2007. "Centering Predictor Variables in Cross-Sectional Multilevel Models: A New Look at an Old Issue." *Psychological Methods* 12(2): 121–138.

Fairbrother, Malcolm. 2014. "Two Multilevel Modeling Techniques for Analyzing Comparative Longitudinal Survey Datasets." *Political Science Research and Methods* 2[1]: 119–140. doi: 10.1017/psrm.2013.24.

Schmidt-Catran, Alexander W., and Malcolm Fairbrother. 2016. "The Random Effects in Multilevel Models: Getting Them Wrong and Getting Them Right." *European Sociological Review* 32[1]: 23-38. doi: 10.1093/esr/jcv090.

Shor, Boris, Joseph Bafumi, Luke Keele, and David Park. 2007. "A Bayesian Multilevel Modeling Approach to Time-Series Cross-Sectional Data." *Political Analysis* 15: 165–181.