MULTILEVEL MODELLING NEWSLETTER

The Multilevel Models Project
Mathematical Sciences
Institute of Education, University of London
20 Bedford Way, London WC1H 0AL, ENGLAND
Web sites: http://www.ioe.ac.uk/multilevel/
http://multilevel.ioe.ac.uk/
Enquiries about newsletter to Ian Plewis
E-mail: i.plewis@ioe.ac.uk
Tel: +44 (0) 20 7612 6688 *Fax: +44 (0) 20 7612 6572*

Vol. 12 No. 2 **December, 2000**

Forthcoming Workshop

25 - 27 April 2001, a three-day introductory workshop to multilevel modelling using *MLwiN* will take place in the Institute of Education, University of London.
Enquiries to Amy Burch at Mathematical Sciences, Institute of Education, 20 Bedford Way, London WC1H OAL. Tel: +44 (0) 20 7612 6688, Fax: +44 (0) 20 7612 6572, e-mail: a.burch@ioe.ac.uk.

Third International Conference on Multilevel Analysis

The Third International Conference on Multilevel Analysis will be held in Amsterdam on April 9 - 10, 2001, followed on April 11 by a one-day course taught by Harvey Goldstein and William Browne (Institute of Education, University of London) about bootstrap and MCMC applications in *MLwiN*.

The conference will be about all aspects of statistical multilevel analysis: theory, software, methodology, and innovative applications. It is organised by Joop J. Hox (University of Utrecht) and Tom A.B. Snijders (University of Groningen). The conference will be in an informal style, with much room for discussion.

There are two invited presentations:

Harvey Goldstein (Institute of Education, University of London) and William Browne (Institute of Education, University of London) "Fitting models to complex data involving hierarchical, crossed and multiple membership structures"

Sophia Rabe-Hesketh (Institute of Psychiatry, University of London) "Generalising the generalised linear mixed model"

aLCD analysis OF LARGE AND COMPLEX DATASETS

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

You are invited to submit abstracts for contributed presentations. Abstracts of 10 to 50 lines of text can be submitted until January 1, 2001. Those who submit an abstract will be notified before January 15 about the acceptance for presentation at the conference. Abstracts should be sent as an ASCII file (not in a different word processor format!!) by email to multi.level@ppsw.rug.nl.

Information about the conference, including a registration form, is available at the web site, www.gamma.rug.nl/multilev.html. A registration form is also available at this web site. Further questions can be addressed to the email address multi.level@ppsw.rug.nl.

Social Science Methodology in the New Millennium: Fifth International Conference on Logic and Methodology

The above conference, organised by Research Committee 33 of the International Sociological Association, was held in Cologne, Germany between 3 and 6 October. There were four sessions devoted to multilevel modelling, two organised by Joop Hox and Cora Maas and two organised by Dick Wiggins. In addition, there were some papers in other sessions of the conference that used multilevel modelling techniques. The following papers were presented:

1. BELL, J.F., and DEXTER, T. *Using Multilevel Models to Assess the Comparability of Examinations*. bell.j@ucles.org.uk

2. BENTLER, P.M. and LIANG, J. *General Two-Level Mean and Covariance Structure Models: Maximum Likelihood via EM Type Algorithms*. bentler@ucla.edu

3. BLIEN, U., and BRIXY, U. *A Multilevel Analysis of the Development of Organisations under the Influence of Context Effects. A Test of the CIGLS Estimator with Panel Data*. uwe.blien@iab.de

4. BOKER, S. *Multilevel Modelling of Dynamical Systems: Random Coefficients and Order Parameters*. sboker@nd.edu

5. DELPRATO, M.A. *Determinants of the Educational Achievement for Elementary Schools Using Multilevel Analysis Techniques*. delpratom@hotmail.com

6. FIELDING, A. *Generalised Linear Mixed Models for Ordered Responses in Multilevel and Other Complex Data Structures in the Social Sciences*. a.fielding@bham.ac.uk

7. FRICK, U. and REHM, J. *Length of Stay in a Psychiatric Hospital as a Function of Patient and Organisational Characteristics - A Multilevel Analysis*. uli.frick@bkr-regensburg.de

8. HEATH, A. and ANDERSEN, R. *Individual and Neighbourhood Class Effects on Voting in Britain*. robert.andersen@sociology.ox.ac.uk

9. HOX, J.J. *Multilevel Multivariate and Structural Equation Models: Some Missing Links*. j.hox@fss.uu.nl

10. JOSHI, H., WIGGINS, R. and VERROPOULOU, G. *Migration, Family Diversity and Children's Well-Being: A Multilevel Analysis of the Second Generation of the 1958 British Cohort*. hj@cls.ioe.ac.uk

11. MAAS, C.J.M. and HOX, J.J. *Robustness of Multilevel Parameter Estimates*. c.maas@fss.uu.nl

12. PLEWIS, I. *Modelling Variation in Parameters in Explanatory Models for Longitudinal Data*. i.plewis@ioe.ac.uk

13. RAUDENBUSH, S. *Assessing the Quality of Environments in which Children Develop*. rauden@umich.edu

14. STOEL, R. and van den WITTENBOER, G. *Longitudinal Multilevel Models and Latent Growth Curve Models in the Presence of Missing Data and Autocorrelation*. reinoud@educ.uva.nl

15. YANG, M. and WOODHOUSE, G. *Multivariate Analysis of Examination Data*. m.yang@ioe.ac.uk

Multivariate Behavioral Research: Special Issue

Multivariate Behavioral Research, the main journal of the Society of Multivariate Experimental Psychology, is going to dedicate a special issue to the topic of multilevel data analysis early in 2001. Steve Reise and Naihua Duan (UCLA) are the guest editors. The issue will include articles on the analysis of multi-site intervention trials by Carvajal, Baumler, Harrist, & Parcel and by Livert, Rindskipf, Saxe, & Stirratt. Hoesksma and Knol as well as Plewis describe the analysis of longitudinal data and testing developmental hypotheses. McArdle presents an application of MLM to the prediction of college grades. Krull and MacKinnon describe the analysis of mediated effects in MLM contexts, and Rabe-Hesketh, Toulopoulue, & Murray present methodological issues that arose in their analysis of cognitive function in schizophrenic patients and their relatives.

Ordinal responses in cross-classified and other complex structures

As part of his Visiting Research Fellowship at the Multilevel Models project, Antony Fielding of the University of Birmingham has written an *MLwiN* macro ORDCAT. This extends the quasi-likelihood procedures for ordinal responses in hierarchical structures available in the distributed *MLwiN* macro MULTICAT to cross-classified and weighted random effects. The set up for the estimation in these structures is similar to that outlined in the *MLwiN* user guide for linear response models. The macro and accompanying user notes can be downloaded from www.bham.ac.uk/economics/staff/tony.htm. Papers dealing with methods and applications are also available at that site.

Using ordinal multilevel models to assess the comparability of examinations

John F Bell and Trevor Dexter

University of Cambridge Local Examinations Syndicate

Introduction

Ordinal models to investigate the comparability of different syllabuses for the same subject and for the same type of examination are described. Comparability is defined in terms of candidates at the same level of prior attainment having the same probability of obtaining a grade regardless of the syllabus (Bell and Dexter, 2000). This requires assumptions about progress being the same for both qualifications. The data have a multilevel structure with the examination candidates grouped in centres (usually schools or colleges). This leads to the use of multilevel regression models taking the result of the examination as the dependent variable, the measure of prior or concurrent outcomes as one of the explanatory variables and dummy variables for the syllabuses under consideration.

The results of school examinations in the UK are expressed as a series of grades, seven in this case. Hence, a choice has to be made as to how the grades should be used as a dependent variable. There are three choices. Firstly, the grades can be converted into points and analysed as a continuous variable using a linear multilevel model. Secondly, the grades can be treated as an ordinal variable and a proportional odds model can be used. Finally, a series of binary variables can be created and analysed using logistic regression. The last two options are considered here.

If the examination grade is considered as an ordinal variable (e.g. Fielding, 1999) then this type of response can be fitted using the proportional odds model. Although this model solves the problem of floor and ceiling effects by fitting a series of s-shaped logistic curves, it assumes identical log-odds ratios for each grade, i.e., the shape of relationship between the probability of obtaining a grade and the outcome measure is the same for each grade. In this example, the use of the proportional odds model will be considered and the adequacy of this assumption investigated. There are other models that have been proposed for ordinal data such as continuation odds models. However, these are less suitable here because the categories used to develop the models are based on more than one grade boundary. This means that the parameters of the model are difficult to interpret. For example, the probability of obtaining a grade C is determined by the choice of the D/C boundary and the choice of the C/B boundary while the probability of obtaining at least a grade C is only determined by the choice of the D/C boundary.

The data come from an approximate 10% random sample of centres that entered candidates for GCSE in a particular subject taken from a linked database of Key Stage 3 (KS3) assessment results and GCSE results. In England (and Wales and Northern Ireland), pupils in state-maintained schools are tested in English, mathematics and science at fourteen years of age (which is also described as the end of KS3). These pupils usually go on to take their GCSE examinations at age 16 at the end of KS4. There are five different syllabuses for the subject under consideration and the sample included 43,366 candidates nested within 460 centres. It should be recognised that the analysis described in this paper was carried out to demonstrate methodology and is not intended to be definitive. Potentially important explanatory variables that could have significant implications have not been included in the model. It is for this reason that the examination and the syllabuses have not been identified. The ordinal regression models were fitted using MIXOR (Hedeker and Gibbons, 1996). Full details of this theory applied to comparability analyses are given in Bell and Dexter (2000).

The assumption about the relationships between the variables for each grade can be investigated graphically or by fitting a series of logistic regression models and seeing how the parameters vary. Cook and Weisberg (1999) proposed the use of a smoother to visualise the mean function when the response is binary. This technique can also be applied to ordinal data. It is possible to plot the estimated logits for each level of the total KS3 score for the five GCSE syllabuses for the example subject and then fit a series of lines using the LOWESS smoother (Fig. 1). This plot ignores the multilevel structure and plots where separate lines were fitted for samples of centres were also considered. However, it is not clear from this type of plot whether the proportional odds assumption holds. It is difficult to be certain that the lines represent curves that have the same shape. Only curves for at least a grade A, at least a grade C and at least a grade F have been plotted. The ‘at least grade A’ curves are the group at the bottom right of the graph and the ‘at least grade F’ curves form the group toward the top left. The analysis is restricted to these three grades for two reasons. Firstly, the plot is easier to follow with fewer curves plotted. Secondly, in the process of setting the grade boundaries, only the boundaries for grades A, C and F are set by the awarding committee and the remainder are calculated from these boundaries using a set of rules. Note that there were not many candidates with very low KS3 scores, which means that the extreme left-hand ends of the curves have not been predicted very well. This plot would suggest that the relationships between total KS3 score and the probabilities of obtaining a grade vary between syllabuses, although it might possibly be only one of the syllabuses. If this were the case, it would suggest that there is a difference in the nature of the syllabuses.

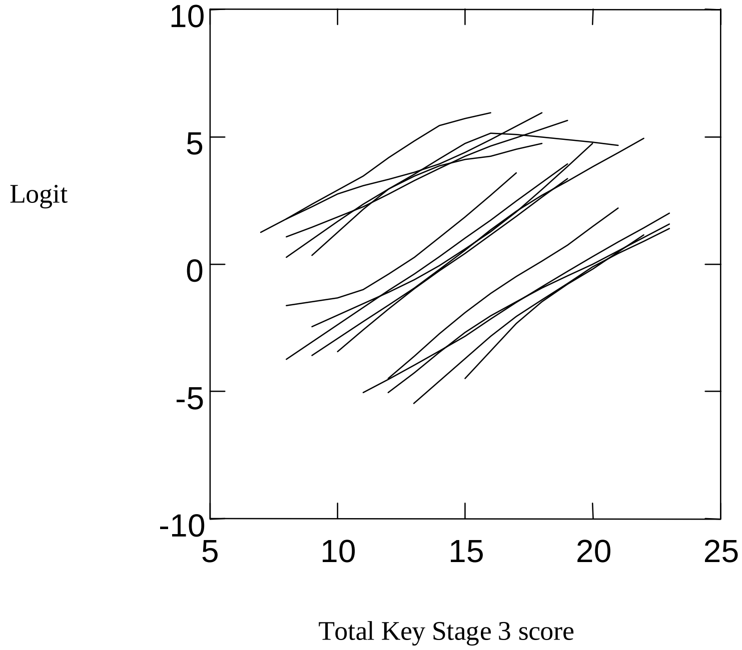

Figure 1: Logit transformations of the estimated probabilities

In Table 1, the estimates from an ordinal model and a series of binary models are presented. These models were fitted using the program MIXOR (Hedeker and Gibbons, 1996) which means that the random centre parameter is a standard deviation rather than a variance. An inspection of the total KS3 parameters and the syllabus parameters indicate that the proportional odds model is not appropriate for this dataset because the values vary from grade to grade. In addition to fitting the models considered below, random slopes models were considered. The random slope parameter for the total KS3 score was significant but too small to be of any substantive importance.

Table 1: Comparison of proportional odds and logistic regression models

Parameter	Ordinal		At least Grade A		At least Grade C		At least Grade F	
	Est.	s.e.	Est.	s.e.	Est.	s.e.	Est.	s.e.
Fixed								
Intercept	-4.42*	0.08	-15.16*	0.19	-10.86*	0.13	-2.54*	0.24
KS3 total	0.74*	0.00	0.77*	0.01	0.78*	0.01	0.52*	0.01
Syll 1	-0.34*	0.05	0.00	0.13	-0.20	0.11	-0.73*	0.21
Syll 2	-0.38*	0.05	0.01	0.12	-0.24*	0.11	-0.91*	0.20
Syll 3	0.21*	0.06	0.46*	0.16	-0.02	0.13	0.02	0.22
Syll 4	-0.25*	0.05	-0.24	0.14	-0.22*	0.11	-0.81*	0.06
Random								
Centre	0.73*	0.16	0.89*	0.04	0.86*	0.03	0.81*	0.06
Thresholds								
F	1.09*	0.04						
E	2.55*	0.04						
D	4.20*	0.04						
C	5.84*	0.04						
B	7.98*	0.04						
A	9.97*	0.05						
A*	11.92*	0.05						

* significant at 0.05

Although there is evidence to suggest that the proportional odds model was not appropriate for this set of data, the evidence from the exploratory plots suggests that the binary models are not satisfactory either. A second series of logistic regression models was fitted this time with terms representing the interaction between syllabus and KS3 score. A significant interaction was found only for syllabus 1(Table 2). There are differences between the syllabuses. Differences between syllabuses can be corrected by changing the boundary marks for the grades only if there are no interactions between the syllabus and the national test results. In this instance, the difference between syllabus 1 and the others could be caused by differing measurement characteristics.

One feature of this analysis is the difference between grade F and the other two grades. It should be recognised, however, that for this subject the vast majority of pupils achieve at least a grade F.

Table 2: Final series of logistic regression analyses

Parameter	At least Grade A Est.	s.e.	At least Grade C Est.	s.e.	At least Grade F Est.	s.e.
Fixed						
Intercept	-14.97*	0.02	-10.65*	0.14	-2.20*	0.26
KS3 total	0.76*	0.01	0.76*	0.01	0.49*	0.01
Syll 1	-0.82	0.42	-1.18*	0.29	-1.91*	0.42
Syll 2	0.01	-0.12	-0.24*	0.11	-0.92*	0.20
Syll 3	0.47*	0.16	-0.01	0.12	0.01	0.22
Syll 4	-0.24	0.14	-0.21	0.11	-0.80*	0.21
KS3*syll1	0.05*	0.02	0.06*	0.02	0.09*	0.03
Random						
Centre	0.89*	0.04	0.85*	0.04	0.81*	0.04

* significant at 0.05

Using the methodology given in Snijders and Bosker (1999), the explained proportions of the variation for the final series of models were calculated (Table 3). The models do not explain a high proportion of the variance in this example data set. It should be noted that, in most cases, the candidates would have attended the same centre for three years prior to KS3 tests. This means that some of the centre effect is included the KS3 total score.

Table 3: Partition of variation for the final series of binary models

Proportion of variation:	Grade A	Grade C	Grade F
Explained	0.54	0.55	0.37
Unexplained at the centre level	0.10	0.09	0.12
Unexplained at the candidate level	0.36	0.36	0.51

With this approach to comparability, it is impossible to determine whether any difference that is found indicates a difference in standard between awarding bodies or syllabuses or both, or whether it indicates that the candidates made more progress on one syllabus compared with another. This approach should be thought of as a screening process to identify potential problems. To investigate comparability further it is necessary to carry out costly studies involving expert judges.

References

Bell, J.F., and Dexter, T. (2000). *Using multilevel models to assess the comparability of examinations*. Paper presented at Fifth International Conference on Social Science Methodology, October 3 - 6, 2000. http://www.leeds.ac.uk/educol/documents/00001528.doc

Cook, R., and Weisberg, R.D. (1999). *Applied regression including computing and graphics*. New York: John Wiley and Sons.

Fielding, A. (1999). Why use arbitrary points scores? Ordered categories in models of educational progress. *Journal of the Royal Statistical Society, A – Statistics in Society*, **162** (3): 303-328.

Hedeker, D., and Gibbons, R.D. (1996). MIXOR: a computer program for mixed-effects ordinal regression analysis. *Computer Methods and Programs in Biomedicine*, **49**: 157-176.

Snijders, T.A.B., and Bosker, R.J. (1999). *Multilevel Analysis. An introduction to basic and advanced multilevel modelling*. London: Sage Publications.

Review of ' Modeling Longitudinal and Multilevel Data: Practical Issues, Applied Approaches, and Specific Examples, Edited by T. D. Little, K. U. Schnabel & J., Baumert. Pp viii & 297, 2000, Mahwah, NJ: Lawrence Erlbaum Associates.

Antony Fielding

University of Birmingham

This reviewer, being interested in extensions of multilevel modelling methodology and newer areas of application, was immediately attracted by the title of this volume. As he read it, however, came a growing realisation that the title was somewhat misleading. He expected to see a rich exploitation of the developing methods for the analysis of hierarchically structured data and associated software to longitudinal studies. In this he was disappointed. Most of the modelling frameworks proposed for the situations and applications discussed fall within the structural equation modelling (SEM) tradition. Thus, for instance, practically every chapter makes reference to or uses standard SEM software such as LISREL, AMOS, and MX. Except for a few earlier chapters there is scant attention to hierarchical random effects features embodied in multilevel software such as MLwiN, HLM, SAS Proc Mixed, or econometrics software for dealing with panel data. More appropriate titles for the book might have been Structural Equation Models for Longitudinal Data or Latent Variable Models in Longitudinal Analysis. With titles such as these the reviewer's prior excitement might have been dampened a little, or at least he might have got excited for different reasons.

The book starts off well enough in raising these expectations. The initial chapter by the editors refers to hierarchical linear modelling (HLM) and SEM as the unequal twins. We are led to expect there may be greater exploitation of both in comparative analyses than actually ensues (apart from a limited number of brief references in isolated chapters). Chapter 2 by Joop Hox is an introduction to the different modelling frameworks of these twins with contrasts that could be exploited by longitudinal researchers. This is an excellent chapter written in a very clear way. The essence of longitudinal data - that it may be construed to be hierarchically structured as a series of time observations within individuals - is emphasised. It is this feature that makes it clearly multilevel and which is not fully exploited in the chapters that follow. Hox also discusses newer developments in which latent variable models operate at various levels in a hierarchy of units. This idea is not explored much in the rest of the book.

Chapters 3-7 discuss a range of applications. They mostly deal with latent growth curves of varying degrees of complexity. The structural equation tradition is emphasised in the main. Here individuals have different patterns of growth governed by different values of intercepts (levels), slopes and other characterising quantities. These latter are treated as unobservable latent variables on individuals within a structural model. The multilevel modelling tradition would treat them as randomly varying coefficients with a distribution. There are a few comparisons of these but almost all the applications discussed utilise the latent variable conceptualisation. From side comments in a few places there may be a lack of recognition by some authors of the capacity of multilevel models to deal with random regression coefficients. Occasionally there are comments that structural models and multilevel models reach broadly the same general conclusions. However, there is no attempt anywhere to recognise that multilevel modelling methodology may perhaps yield more efficient estimates.

Another irritating feature of many of these chapters is the standard insistence that models and applications must always be accompanied by a path diagram. The latter idea started off its history as a very illuminating and clear way of representing the way that variables, both manifest and latent, influenced each other. As models become more complex this illumination disappears in the morass of confusing detail that must necessarily be imposed on the diagram. It then ceases to make the situation clearer and often makes it foggier.

Chapters 8 –14 deal with a variety of topics within the longitudinal analysis tradition. Chapter 8 is a discussion of a model where developmental changes over time are conceived of as changes of state in a latent class variable. This idea was one originally developed by the pioneers of latent class analysis, Lazarsfeld and Henry. This is not referenced. Chapter 9 is partly a thinly veiled attack on classical hypothesis testing which is strictly unnecessary for its central theme, sequential evaluation and testing of nested structural models. Chapters 10, 11, and 12 deal with a variety of interesting approaches to the endemic problems of longitudinal analysis: attrition and missing data. These have rather broader applicability than being tied to a specific modelling tradition. Chapter 12 is nearer the multilevel modelling tradition in that it handles attrition and missingness through full information maximum likelihood on the raw data. Chapter 13 is about attempts to reconcile and integrate different pieces of software with each other. Chapter 14 is a methodological departure from traditional structural equation software through fitting covariance structures to summarised covariance matrices. Maximum likelihood estimation and fits to raw data are proposed as a possibly more fruitful alternative.

The papers in this book are interesting and mostly well written. Those interested in structural and latent variable modelling will find a range of interesting applications and new developments. The chapters on missing data offer some new insights to longitudinal researchers wrestling with these problems. Those interested in multilevel modelling per se will be a little disappointed that the contents do not fully reflect the title. However, the data used in the book's applications are made fully available on the publisher's web page. Perhaps some of those disappointed researchers may be able to demonstrate elsewhere the scope for full multilevel modelling of these data.

Some Recent Publications Using Multilevel Models

Dryler, H. (1999). The Impact of School and Classroom Characteristics on Educational Choices by Boys and Girls: A Multilevel Analysis. *Acta Sociologica*, **42** (4): 299-318.

Fielding, A. (1999). Why use arbitrary points scores? Ordered categories in models of educational progress. *Journal of the Royal Statistical Society, Series A*, **162**: (3), 303-328.

Fielding, A. (1999). *Hierarchical Random Effects Models for Ordered Category Responses*, in Statistical Modelling: Proceedings of the 14th International Workshop, edited by H. Friedl, A Berghold, G. Kauerman, 508-512, Technical University Graz, Graz, Austria.

Leyland, A. H., Langford, I. H., Rasbash, J., & Goldstein, H. (2000) Multivariate spatial models for event data. *Statistics in Medicine*, **19**: 2469-2478. http://www.msoc-mrc.gla.ac.uk/Reports/PDFs/Published-Papers/Sep00-PRESMPV.pdf

Leyland, A. H., & McLeod, A. (2000). Mortality in England and Wales, 1979-1992. *An introduction to Multilevel Modelling using MLwiN*. MRC Social and Public Health Sciences Unit, Glasgow, Occasional paper no. 1. http://www.msoc-mrc.gla.ac.uk/Reports/PDFs/Occasional-Papers/OP-001.PDF

Kuo, M., Mohler, B., Raudenbush, S. L., & Earls, F. J. (2000). Assessing Exposure to Violence Using Multiple Informants: Application of Hierarchical Linear Model. *J Child Psychol. Psychiat.* **41** (8): 1049-1056.

Rosel, J., Oliver, J.C., Jara, P., & Caballer, A. (2000). A multilevel time-series model for the incidence of AIDS cases in Spain. *Health and Place*, **6:** 309-317.

Yang, M., Heath, A., & Goldstein, H. (2000). Multilevel models for repeated binary outcomes: attitudes and vote over the electoral cycle. *Journal of Royal Statistical Society, A*, **163** (1): 49-62.

Please send us your new publications in multilevel modelling for inclusion in this section in future issues.

MLwiN Clinics in London Spring 2001

Tuesday 16 January 2001
Tuesday 13 February 2001
Tuesday 13 March 2001
Tuesday 10 April 2001
Tuesday 8 May 2001

at

Multilevel Models Project
11 Woburn Square, London WC1A 0SN

Contact Min Yang for appointments
Tel: +44 (0) 20 7612 6682/6688
Email: m.yang@ioe.ac.uk

If you would like to be added to the e-mail list, please send your e-mail address stating that you wish to subscribe to the newsletter, to a.burch@ioe.ac.uk