MULTILEVEL MODELLING NEWSLETTER

Produced through the Multilevel Models Project:

c/o Math, Statistics & Computing Department
Institute of Education, University of London,
20 Bedford Way, London WC1H 0AL
e-mail: rprosser%uk.ac.lon.educ.isis@ac.uk
Telephone: 01-636-1500 ext 473

Vol. 1 No. 2 May, 1989

EDINBURGH MULTILEVEL CONFERENCE PROGRAM ANNOUNCED

The *International Conference on Applications of Multilevel Methods in Educational Research* will be held in Edinburgh on August 12th and 13th, 1989. The conference will share results from studies of schools, schooling and educational reform and, from the vantage point of this experience, it will examine the potential uses of multilevel methods and the conceptual issues involved in the design and interpretation of multilevel analyses.

The conference will concentrate on non-technical aspects of multilevel methods; participants will not need to be familiar with computing or statistical aspects of multilevel techniques. The tentative program outline is as follows:

Saturday, August 12, 1989

Effects of School Organization on Teacher Commitment and Self-Efficacy

V. Lee	The effects of school organization on teachers' satisfaction and self-efficacy
A. Bryk	The organization of teachers' work: its effects on teacher commitment and engagement in U.S. high schools
B. Rowan	Aggregate measures of teachers' perceptions of school climate: a multilevel analysis

Effects of School Organization and Practice on Pupil Achievement

M. Lockheed	Effects of school organization and practice in developing countries
I. Plewis	Using multilevel models to link educational progress with curriculum
S. Raudenbush and D. Willms	The stability of school effects

Reliability and Validity from a Multilevel Perspective

S. Jacobsen	The relationship between kindergarten screening measures and grade 3 achievement
L. Paterson	Trends in attainment in Scottish schools
N. Longford	Searching for multivariate outcomes in education

Sunday, August 13, 1989

Multilevel Research on Performance Indicators

M. Aitkin	Pragmatic decision-making with available admininstrative data
C. Fitz-Gibbon	A-levels in the U.K.: School effects
R. Bosker	Indicators off school performance in the Netherlands

Social Stratification and Schooling Outcomes

D. Raffe	Assessing the impact of a decentralized initiative: the British TVEI
A. Gamoran	Curriculum differentiation and achievement: a multilevel model
C. Garner	Neighbourhood effects on educational attainment

WORKSHOPS & CONFERENCES

Edinburgh Conference ... cont'd from p.1

For those with experience in multilevel modelling, there will be an opportunity to discuss applications, experiences, and problems at two informal sessions planned for the afternoons of the 14th and 15th.

The conference is being organized jointly by the Centre for Educational Sociology (CES) and by Stephen Raudenbush (Michigan State University) and Douglas Willms (CES and University of British Columbia). There is no conference fee, but a charge of £30 will be made to cover part of the cost of lunch and refreshments on both days and of pre-circulated papers.

The Edinburgh International Festival starts on August 13, so participants are advised to arrange accommodation as early as possible. A copy of the 1989 Edinburgh Accommodation Register will be sent to those who request it.

To register, contact the conference administrator at the Centre for Educational Sociology, Edinburgh University, 7 Buccleuch Place, Edinburgh EH8 9LW Scotland. The Fax no. is 031 668 3263, the telephone no. is 031 667 1011 ext 6803, and the e-mail address is ekjc18@uk.ac.edinburgh

TORONTO WORKSHOP: 1990

Planning is under way at the Ontario Institute for Studies in Education (OISE) in Toronto, Canada for a workshop in multilevel modelling. This training session for data analysts, conducted by the staff of the Multilevel Models Project, will be held from March 22 to March 24 1990. The tentative timetable is as follows:

22 (eve.) Registration, social gathering, and introductory seminar.

23 (a.m.) Introduction to *ML3* and hands-on practice with example data sets.

23 (aft.) Analysis of more extensive data sets including data supplied by partcipants.

23 (eve.) Analysis of longitudinal and repeated measures data.

24 (a.m.) Diagnosis of common problems. Interpretation of analyses of participants' data sets.

24 (aft.) Wrap-up session.

For those who can remain in Toronto, computers and assistance will be available throughout the weekend.

A more detailed outline will be available in mid-June 1989. Final details will be ready to circulate to participants by November 1989. The registration fee for the workshop will be about US$400. Arrangement for accommodation and meals will be announced later.

Inquiries may be directed to Professor D. F. Burrill, Department of MECA, OISE, 252 Bloor Street West, Toronto, CANADA M5S 1V6; or by e-mail to MLEVEL@UTOROISE on BitNet.

THREE-LEVEL WORKSHOP NOTICE

On October 18-20, 1989, the Multilevel Models Project will offer a workshop on three-level modelling at the Institute of Education of the University of London. This training session, designed primarily for social scientists, will feature discussion of theory as well as hands-on work with the new *ML3* program. The emphasis will be on data analytic practice rather than mathematical statistics. The first day will be a review of the basics of two-level analysis for those who are new to the field. An optional evening session will be arranged for participants who wish to begin an analysis of data they bring.

No fee is charged for attending the Project's workshops, and each participant receives a copy of the Project's software. Further information and registration forms are available from Bob Prosser at the Institute of Education.

WANTED

Comprehensive review of *Multilevel Analysis of Educational Data* (see p. 3) for the October edition. "Reward" offered: a copy of the book. Please direct enquiries to Bob Prosser at the masthead address. Thank you.

NEW LITERATURE

Multilevel Analysis of Educational Data IS NOW AVAILABLE

Multilevel Analysis of Educational Data, edited by Darrell Bock, is the proceedings of an international conference on multilevel analysis held in April, 1987. Published in hardcover by Academic Press, it sells for *£30*.

The work presented is "theoretical" and "applied," and there are numerous data analysis examples. A comprehensive review would be most welcome (see p. 2), but for now the range of the topics discussed will be indicated simply with a listing of the chapter titles.

- Some Applications of Multilevel Models to Educational Data *Donald Rubin*
- Empirical Bayes Methods: A Tool for Exploratory Analysis *Henry Braun*
- A Hierarchical Item-response Model for Educational Testing *Robert Mislevy & Darrell Bock*
- Difficulties with Bayesian Inference for Random Effects *Charles Lewis*
- Multilevel Aspects of Varying Parameters in Structural Models *Bengt Muthén & Albert Satorra*
- Models for Multilevel Response Variables with an Application to Growth Curves *Harvey Goldstein*
- Multilevel Models: Issues and Problems Emerging from their Recent Application in British Studies of School Effectiveness *John Gray*
- Toward a More Appropriate Conceptualization of Research on School Effects: A Three-level Hierarchical Linear Model *Tony Bryk & Steve Raudenbush*
- Quantitative Models for Estimating Teacher and School Effectiveness *Steve Raudenbush & Tony Bryk*
- Multilevel Investigations of Systematically Varying Slopes: Issues, Alternatives and Consequences *Leigh Burstein, Kyung-Sung Kim, & Ginette Delandshere*
- Profile Predictive Likelihood for Random Effects in the Two-level Model *Murray Aitkin*
- Fisher Scoring Algorithm for Variance Component Analysis of Data with Multilevel Structure *Nick Longford*

Peter McCullagh, Hariharan Swaminathan, Robert Tsutakawa, and Paul Holland provided discussions of the papers, and Darrell Bock contributed an additional paper entitled Measurement of Human Variation: A Two-stage Model.

VIEWPOINT

A QUESTION OF NOTATION

You have probably noticed (and perhaps become frustrated by) a diversity in the notation used in expressing multilevel models. Consider subscripting, as a simple example. An examination of the articles in *Multilevel Analysis of Educational Data* (to choose a cross section of the literature close at hand) suggests that about half the authors like person i to belong to group j, and the other half prefer person j to be part of group i. (In one exposition, the preference seems to switch midway!)

A second area of potential confusion is the naming of random parameters in models with coefficients that vary at more than one level. A general notation would provide two pieces of information for each random variable and random parameter, namely, the identifier(s) of the associated explanatory variable(s) and the level.

While notation may not be the most serious obstacle encountered by newcomers to the field as they learn about basic two-level modelling, our experience in presenting introductory workshops to social scientists indicates that the "slightest" fuzziness can lead to serious misinterpretations. Further, it seems to us that as three- (and higher-) level modelling becomes more common, and as complex possibilities such as random cross classifications are used, easing notational burden will become increasingly important.

What do you think about this matter? We'll print your views in the October edition.

SOFTWARE

ML2 V2.0 IS NOW BEING SHIPPED

Several new features have been added to the *ML2* software since the report in the previous issue of this newsletter. Six of these will be noted here:

- a listwise deletion option;
- a command for summarizing the grouping in a multilevel data set and another for creating a group size variable;
- a facility for merging data from a group-level file with data from an individual-level file;
- a command for extracting group-level information in a compact form; and
- a record-splitting command useful in preparing to conduct a multivariate or repeated measures analysis.

Version 2.0 comes with a new manual which contains several examples, improved indexing of commands, and an expanded discussion of basic theory.

ML2 is designed for 286/ 386 computers with at least 560Kb of *available* RAM and a hard disk. A graphics card is needed to use the high resolution plotting facilities. A coprocessor version is available, and the program can be sent on a 5.25 inch (360 Kb) or a 3.5 inch diskette.

Registered users of *ML2* will be sent new disks and manuals automatically. (A registered user is one who has paid £50 or US$80.) Enquiries about ordering *ML2* may be sent to Bob Prosser at the address on the masthead.

ML3 IS UNDERGOING TESTING

ML3, the Multilevel Models Project's three-level software is in the preliminary stages of testing, and Version 1.0 should be ready for distribution in mid-summer. The operation of the program is very similar to that of *ML2*, and a very wide variety of random term covariance structures can be fitted. In addition some new features will be available:

- a macro facility;
- plotting operations which reflect the multilevel structure of the data; and
- a command for performing a weighted analysis.

Please direct enquiries to Jon Rasbash at the address on the front page.

MULTILEVEL WORK AT UCLA

Jan de Leeuw

Several projects are in the works at the University of California at Los Angeles. Jan de Leeuw, Ita Kreft, and Kyung-Sung Kim are planning a comparison of five multilevel modelling programs–HLM, VARCL, GENMOD, *ML2*, and BMDP-5V–with respect to (a) classes of models which can be fitted; (b) computational precision; (c) ease of operation; (d) output; (e) options; and (f) speed.

GENMOD is Bill Mason's mainframe program based on the EM algorithm for fitting random coefficient hierarchical models (Wong & Mason, 1989), and BMDP-5V is a program for fitting mixed models based on work by Jennrich & Schluchter (1986).

A second study will examine the stability and prediction properties of random coefficient regression models, using cross-validation, the bootstrap, and asymptotic expansions. Visitors Nick Longford and Joop Hox will be involved in these projects.

A program called Multipath is being developed to perform multilevel path analysis using variables on various measurement levels. Part of the computational core using the Lindstrom & Bates (1988) algorithm has been prototyped in APL, and part of the command line interface has been written in C on the Macintosh.

It is expected that VARCL and GENMOD will be incorporated into a new series of BMDP software. At present, Macintosh (68000 and 68020/ 030) versions of these programs are available. The VARCL interface is being improved with windows and menus, and a new manual is planned.

References

Lindstrom, M. J., & Bates, D. M. (1988). Newton Raphson and EM algorithms for linear mixed-effects models for repeated measures data. *JASA, 83,* 1014-1022.

Jennrich, R. I., & Schluchter, M. D. (198). Unbalanced repeated measures models with structured covariance matrices. *Biometrics, 42,* 805-820.

Wong, G. Y., & Mason, W. M. (1989). *Ethnicity, comparative analysis, and a generalization of the hierarchical normal linear model for multilevel analysis* (Report 89-138). Ann Arbor: University of Michigan, Population Studies Center.

DEVELOPMENTS

HLM2 HAS NEW FEATURES, AND HLM3 V1.0 IS NOW AVAILABLE

Tony Bryk

The HLM2 and HLM3 programs provide a user-friendly interactive environment for fitting two- and three-level hierarchical linear models, respectively. The software has been specifically designed to facilitate cross-level (e.g., "slopes-as-outcomes") modelling. Within-unit effects can be either fixed or random, and different between-unit models can be specified for each random effect. Hypotheses about the random parameters can be tested using χ^2 values provided. Users can obtain a residual file for graphing and analysis with other software. Listwise and pairwise deletion options are available for handling missing values of individual-level variables.

Mini and mainframe versions are available for several machine/ OS combinations–the IBM 3090 (VM/CMS), H-P 9000 (UNIX V), and VAX 8650 (VMS). The software is written in standard FORTRAN 77 and should require only minor changes to port to other systems. An MS-DOS version of HLM2 is available that can handle up to 300 groups, five within- and seven between-group explanatory variables. The latter requires a machine with 640K of RAM, a hard disk, and a coprocessor.

Version 2.0 of HLM2 includes the following new features:

- a *V known* routine for applications in which the sampling variance matrix for the within-unit model is assumed known;
- an Aitkin accelerator to reduce computation time (see Laird, Lange, & Strom, 1987);
- a general linear hypothesis testing module for fixed effects and likelihood ratio testing for dispersion components;
- improved formatting of output; and
- restructured "fixing" of within-unit slopes permitting easy inclusion of fixed-effect cross-level interactions.

The V known routine can be used in synthesizing results across a set of research studies and in studying correlates of diversity (Raudenbush & Bryk, 1985, 1987). This routine will handle multiple random effects per unit and accept a general dispersion sampling matrix.

HLM3 has all the basic features of HLM2 including a similar interactive user interface. Heterogeneous or homogeneous dispersion matrices can be specified, and two-level problems in which the first stage outcome is multivariate can be modelled.

Ordering HLM2 and HLM3

Until July 1, 1989 the software can be purchased by writing to: Richard Congdon, University of Chicago, 5835 S. Kimbark Avenue, Chicago, Illinois 60637 (USA).

Future distribution of the programs will be handled by Scientific Software, 1369 Neitzel Road, Mooresville, Indiana 46158-9312 (USA).

Work in Progress

Michael Seltzer working with Wing Wong and Tony Bryk at the University of Chicago has been developing a *data augmentation* strategy for HLM2 (Tanner & Wong, 1987). Data augmentation is a very general computational algorithm which can be employed to generate empirical posterior distributions for HLM. These distributions will provide a better basis for hypotheses testing on small samples, and for examining the sensitivity of effect estimates when the latent random effects are non-normal.

Frank Jenkins has recently completed his PhD work at Michigan State under Steve Raudenbush on "Multilevel Path Modeling." His thesis represents a significant extension of the basic two-level model to include a full path analysis model at level 1.

References

Laird, N., Lange, N., & Strom, D. (1987). Maximum likelihood computations with repeated measures: application of the EM algorithm. *JASA, 82,* 97-105.

Raudenbush, S., & Bryk, A. (1985). Empirical Bayes meta-analysis. *Journal of Educational Statistics, 10,* 75-98.

Raudenbush, S., & Bryk, A. (1987). Examining correlates of diversity. *Journal of Educational Statistics, 12,* 241-269.

Tanner, M., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *JASA, 82,* 528-540.

STORAGE AND RETRIEVAL OF MULTILEVEL DATA

Pieter van den Eeden & Pjotr Koopman

The Ministry of Education in the Netherlands plans to modify the national system of secondary education to make it more comprehensive. The plans call for changes to be phased in over ten years beginning in 1989 with continuous monitoring of the implementation. This evaluation requires collection of multilevel data throughout the implementation period. Storage of information on 20,000 students in three cohorts is needed, and at least five hierarchical levels will be maintained.

The governmental committee which will be organizing the evaluation (the CGE) has recently received a report on a study conducted by by Pjotr Koopman and Pieter van den Eeden concerning systems for storage and retrieval of the evaluation data. The two primary questions addressed were

- Which data base management (DBM) system is appropriate for storing and managing longitudinal educational data in a data base in which units of several levels are stored? and
- Which DBM system is appropriate for representing several hierarchical structures in preparation for different multilevel analyses.

Two important requirements were set for the storage/ retrieval system. The first requirement is that the data must be easily transformable into files that can be read by multilevel modelling software such as *ML2*, HLM, and VARCL. A further necessity is the capability to create multiple hierarchical structures with the data. For example, a three-level analysis of pupils, schools, and cohorts can be nested in two ways. The first is a structure of pupils nested within cohorts and a nesting of these cohorts in schools. This hierarchy is suited for questioning change and stability of the educational processes within specific schools. The second involves pupils nested within schools and schools within cohorts. This is useful for examining change and stability between schools over time. The choice of hierarchical structure depends on the theoretical research questions of interest.

A relational database is ideal for meeting the storage requirements. The data from each different level can be stored efficiently and reliably in a separate table–essentially a flat file. Relations among tables can be pointed out explicitly using information within the tables themselves. The key variable of one table is just an ordinary variable in the related table. This property permits retrieval of any type of hierarchical structure. It is possible to distribute data from higher level units to constituent lower level units. Further, it does not matter whether this data is obtained at the higher level directly or aggregated from lower level units. The file management capabilities of statistical packages such as SPSSX and SAS are seen to be inadequate for this purpose.

The second requirement is that investigators in different parts of the country must have ready access to the data as soon as it is stored in the system. Computer networks provide part of the answer. In the Netherlands, the European Academic and Research Network (EARN) permits file transfer as well as remote login. Modern relational databases are accessible to multiple users with read permission at the same time. Tables and their subparts can be screened so that each user has access to only those sections of the data he/ she is permitted to use.

Users would, of course, have to learn the language of the database program selected. This is not seen as a major problem, however. A new standard is emerging for database programs–structured query language (SQL)–and the popular systems such as SIR, SQL/ DS, ORACLE, and dBase IV are all SQL-oriented.

The study of data storage and retrieval systems concluded with a recommendation of a specific database package–SIR. This program was preferred to its closest competitor, SQL/ DS, for two reasons. Firstly, it is easier to update from flat files (as would be done several times during the evaluation). Secondly, the PQL syntax of SIR is bears a strong resemblance to that of SPSSX, a statistical package that is widely used in the Netherlands. Although statistical packages like SAS and SPSSX are viewed as inadequate for managing multilevel data, SIR-SAS and SIR-SPSSX interfaces permit easy access.

The report is currently available in Dutch. If there is sufficient interest, an English translation or summary can be considered. For further information, please contact Pjotr Koopman, Centre for Educational Research of the University of Amsterdam, S.C.O., Grote Bickersstraat 72, 1013 KS Amsterdam, The Netherlands.

MODELLING g-LEVEL DATA WITH A $(g-1)$ LEVEL MODEL

Harvey Goldstein

Introduction

With the increasing use of multilevel statistical models for highly structured data sets with many levels, the question arises as to whether software designed to analyse a g-level data set can be used to model a $(g+1)$ or higher level data set. As we show below, the general answer is no, but there are certain situations where such an approach can provide useful insights into the data structure.

We consider the case of $g = 2$ with a fixed small number of level 2 units. This allows us to discuss the issues while avoiding the use of unduly cumbersome notation to deal with the general case. In fact we can illustrate the problem by thinking of the concrete example of students (level 1) nested within years (level 2) nested within schools (level 3). The students are from successive year cohorts and the response variable is, say, the score derived from grades obtained on the 16+ school leaving examination in England and Wales (Inner London Education Authority [ILEA], 1987). We first consider the full 3-level model specification and then study possible models which can be fitted when the coefficients of the linear model are allowed to be random only at levels 1 and 2. Goldstein (1987, chapter 4) provides an account of how to specify and analyse such models.

The 3-level Model

The basic multilevel model can be written in the following form (Goldstein, 1989)

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{Ze}$$

with $E(\mathbf{e}) = \mathbf{0}$ and $E([\mathbf{Ze}][\mathbf{Ze}]^T) = \mathbf{V}$. In this expression, $\mathbf{X}$ is a matrix of explanatory variables which define the fixed part of the model, and $\mathbf{Z}$ is a matrix of explanatory variables whose coefficients vary randomly across all or some of the level 1, level 2 or level 3 units, and thus define the random part of the model. In order to discuss the model structure, that is the structure of $\mathbf{V}$, we need deal only with a single level 3 unit. We shall consider the case where all level 3 units contain no more than two level 2 units, but place no restriction on the number of level 1 units. Thus in terms of our substantive example, we suppose that the schools have data on two cohorts of students in consecutive years.

We shall deal first with the simple variance components model, that is with a simple variance parameter only at each level. The structure for $\mathbf{V}$ is as follows, for the level 3 unit indexed by t.

$$\mathbf{V}_t = \sigma_e^2 \mathbf{I}_{n_t} + \bigoplus_{j=1}^{j=2} \sigma_u^2 \mathbf{J}_{n_{tj}} + \sigma_v^2 \mathbf{J}_{n_t} \qquad (1)$$

where n_t is the number of level 1 units in the level 3 unit, n_{tj} is the number of level 1 units in the jth level 2 unit and $\mathbf{J}_n$ is the $(n \times n)$ matrix with every element equal to 1. The first term on the right hand side of equation 1 represents the level 1 contribution to the total variation, the second the level 2 contribution and the third the level 3 contribution.

The 2-level Model

Since we now are modelling the data as a 2-level structure, level 2 now consists of both years for the school. Thus the old level 3 unit becomes the new level 2 unit. Retaining the simple form for the level 1 contribution in equation 1, a general form for the structure of $\mathbf{V}$ is given by

$$\sigma_e^2 \mathbf{I} + \mathbf{Z}\Omega_2\mathbf{Z}^T \qquad (2)$$

where Ω_2 is the $(r \times r)$ covariance matrix of the random coefficients associated with the columns of $\mathbf{Z}$.

Note first that we require just two random parameters to be defined by $\mathbf{Z}$, and there are three ways of doing this. We can define two variances, a variance and a covariance or two covariance parameters. The

Iterative Generalised Least Squares (IGLS) algorithm (Goldstein, 1986) allows any of these to be defined. The specification of a covariance parameter when one or both the corresponding variance parameters is absent is to be regarded for present purposes purely as a device to achieve the desired structure for $\mathbf{V}$.

It is clear from the structure shown in equation 1 that, without loss of generality, we require the elements of $\mathbf{Z}$ to be either 0 or 1, and to be equal within each (old) level 2 unit. Thus there are only 3 possible column vectors for $\mathbf{Z}$. These three are the constant vector containing 1's, the vector which contains elements equal to 1 for the first year cohort and 0 for the second and the vector which contains all entries equal to 0 for the first year cohort and 1 for the second. We also note that the first vector is simply the sum of the other two. Thus there are only three independent parameters which can be defined, two variances and a covariance. If we specify $\mathbf{Z}$ as the matrix which contains these three column vectors, namely

$$\mathbf{Z} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \end{pmatrix} \tag{3}$$

then there is no pair of elements of the (3×3) covariance matrix of the three coefficients random at level 2 which will lead to the structure in equation 1.

We now explore two models which can provide useful information nevertheless. As was mentioned above, the arguments used here extend in a straightforward fashion to any number of units at each level and any number of levels.

A Pseudo 3-level Model

We shall refer to the column vectors in equation 3 respectively as $\mathbf{Z}_1$, $\mathbf{Z}_2$, and $\mathbf{Z}_3$. The full covariance matrix of the random parameters defined at level 2 is written as follows

$$\begin{pmatrix} \sigma^2_{u_1} & & \\ \sigma_{u_{21}} & \sigma^2_{u_2} & \\ \sigma_{u_{31}} & \sigma_{u_{32}} & \sigma^2_{u_3} \end{pmatrix} \tag{4}$$

Our first model sets all the off-diagonal covariance terms in expression 4 to zero. This yields the structure, for the tth level 3 unit

$$\begin{aligned} \mathbf{V}_t &= \sigma^2_e \mathbf{I}_{n_t} + [\mathbf{Z}_{2t}\ \mathbf{Z}_{3t}]\ \Omega_2\ [\mathbf{Z}_{2t}\ \mathbf{Z}_{3t}]^T + \mathbf{Z}_{1t}\ \Omega_3\ \mathbf{Z}^T_{1t} \\ &= \sigma^2_e \mathbf{I}_{n_t} + \bigoplus_{j=2}^{j=3} \sigma^2_{u_j} \mathbf{J}_{n_{tj}} + \sigma^2_{u_1} \mathbf{J}_{n_t} \end{aligned} \tag{5}$$

where

$$\Omega_2 = \begin{pmatrix} \sigma^2_{u_2} & \\ 0 & \sigma^2_{u_3} \end{pmatrix}, \quad \text{and} \quad \Omega_1 = \sigma^2_{u_1}$$

Clearly, this is equivalent to the expression in equation 1 if and only if $\sigma^2_{u_2} = \sigma^2_{u_3}$. Goldstein (1987, appendix 3.1) shows how linear constraints can be incorporated into the model, which would therefore allow a true 3-level model to be specified. Unless this is done, expression 5 would need to be interpreted as a model which fitted different between-school variances for each year with a covariance between year 1 and year 2, these parameters respectively being estimated by $\sigma^2_{u_1} + \sigma^2_{u_2}$, $\sigma^2_{u_1} + \sigma^2_{u_3}$, and $\sigma^2_{u_1}$. In fact, this formulation constrains the covariance to be non-negative and this can be removed by including the variances of the coefficients of

$\mathbf{Z}_2$ and $\mathbf{Z}_3$ and their covariance. If in practice we find that, approximately, $\sigma^2_{u_2} = \sigma^2_{u_3}$, then we can regard expression 5 as a reasonable approximation to a 3-level model.

A Saturated 2-level Model

In the previous section we referred to the model obtained by using the variances and covariance of the coefficients of $\mathbf{Z}_2$ and $\mathbf{Z}_3$. In the general case with many true level 2 units (years), we can extend this accordingly, but the number of random parameters correspondingly increases, so that with p level 2 units we obtain $p(p+1)/2$ random parameters. In effect, such a model treats each level 2 unit as a category defined by dummy variables whose coefficients vary randomly across level 2 units. Thus, in the example, there is one such coefficient associated with each year. In the 3-level formulation these levels are treated as random, so providing a parsimonious and efficient model. The resulting level 2 variation can then be further structured, for example by fitting a linear trend.

Random Coefficients

The generalisation of the above models to the case where $\mathbf{Z}$ contains further discrete or continuous explanatory variables is straightforward. For example, if we have a 'pretest' or 'school intake' score for each student, we might wish to model its coefficient as varying randomly from year to year and from school to school; that is at both levels 2 and 3. If we denote this variable by $\mathbf{W}$ and the random term in its coefficient by w, then the pseudo 3-level model of equation 5 becomes, for a level 3 unit and dropping the subscript t

$$\mathbf{V} = \sigma^2_e \mathbf{I} + [\mathbf{Z}_2\ \mathbf{Z}_3\ \mathbf{W}_2\ \mathbf{W}_3]\ \Omega^{(1)}_2\ [\mathbf{Z}_2\ \mathbf{Z}_3\ \mathbf{W}_2\ \mathbf{W}_3]^T + [\mathbf{Z}_1\ \mathbf{W}]\ \Omega^{(1)}_3\ [\mathbf{Z}_1\ \mathbf{W}]^T$$

where

$$\Omega^{(1)}_2 = \begin{pmatrix} \sigma^2_{u_2} & & & \\ 0 & \sigma^2_{u_3} & & \\ \sigma_{u_2 w_2} & 0 & \sigma^2_{w_2} & \\ 0 & \sigma_{u_2 w_3} & 0 & \sigma^2_{w_3} \end{pmatrix}, \quad \Omega^{(1)}_3 = \begin{pmatrix} \sigma^2_{u_1} & \\ \sigma_{u_1 w} & \sigma^2_{w} \end{pmatrix}, \quad \text{and } \mathbf{W}_j = (\text{diag } \mathbf{W})\mathbf{Z}_j \quad (j = 2, 3).$$

Conclusion

The arguments presented for modelling 3-level data with a 3-level model rather than attempting to approximate the data with a 2-level model apply generally to $(g+1)$-level data. It is only in the case where the number of units at an intermediate level is very small, that such an approximation becomes feasible. In addition, if the units can also be associated with recognisable data categories as in the example, then some useful information can be extracted from a g-level model.

In the practical situation where $(g+1)$-level software is unavailable, a preliminary examination of the data may indicate where two or more levels can be merged. Consider, for example, the case of three level data at student, classroom and school levels. If the number of schools is not too large, we can fit a separate term for each school and study the 2-level model with classrooms at level 2. If the between-classromm variation is small we might then feel justified in ignoring this level. We should note that, because classrooms are not linked in any well defined way across schools, we cannot apply the procedures discussed above.

References

Inner London Education Authority, (1987). *Actual and predicted examination scores in schools.* London: ILEA Research and Statistics Division.

Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalised least squares. *Biometrika, 73*, 43-56.

Goldstein, H. (1987). *Multilevel models in educational and social research.* London: Griffin; New York: Oxford University Press.

"CENTERING" PREDICTORS IN MULTILEVEL ANALYSIS: CHOICES AND CONSEQUENCES

Steve Raudenbush

In a classic article, Cronbach and Webb (1975) alerted psychologists and sociologists to the dangers which arise from ignoring the hierarchial structure of data collected in social groups such as classrooms and schools. Using a forerunner of today's multilevel methods, the authors reanalyzed data from a well-known aptitude-by-treatment interaction study. The reanalysis showed the original findings–that the optimal mode of mathematics instruction depends upon the aptitude of the student–were fallacious. Cronbach and Webb reasoned that a true aptitude-by-treatment interaction would be manifest by treatment group differences in the magnitude of aptitude-achievement regression slopes within classrooms; i.e., a treatment that works better for low than for high aptitude students would attenuate the relationship of achievement to aptitude within the classrooms employing that treatment. Similarly, a treatment that worked better for high than for low aptitude students would increase the aptitude-achievement relationship. However, the multilevel analysis showed no such effects on within-classroom slopes. Rather, the interaction apparent in the original analysis resulted from different between-classroom slopes in the two treatments. This between-class interaction was poorly estimated and uninterpretable. By failing to separate the within- and between-group regressions, the original investigators missed a crucial feature of their evidence and drew an unwarranted conclusion.

The tools available to Cronbach and Webb were crude by today's standards. Researchers now have the means to disentangle effects occurring at each level of a hierarchy. Unfortunately, many users of the powerful new multilevel methods don't achieve this goal. The chief cause is a failure to understand the crucial importance of *centering.* Three cases illustrate this point: the estimation of *contextual effects,* the estimation of *cross-level interactions,* and the study of *random regresison slopes.*

Contextual Effects

Contextual or "compositional" effects involve, for example, the effects on pupil achievement of sharing membership in a classroom with students of high aptitude or motivation. Such effects are presumed to make a contribution independent of the contribution of one's own aptitude or motivation. In actual research practice, the effects of "mean aptitude" or "mean achievement" are typically open to a variety of interpretations: they may reflect effects of better instructional practice occurring in such contexts, the normative effects of a peer group, or they may serve as proxies for unmeasured or poorly measured student characteristics. Contextual effects have been found with such high frequency, however, that they ought not be ignored (see Willms' 1986 review).

The classic formulation of the *contextual effects model* involves an equation of the form

$$Y_{ij} = \beta_{0j} + \beta_w X_{ij} + \gamma \overline{X}_j + u_j + e_{ij} \qquad (1)$$

where Y_{ij} might be the achievement of child i in school j, X_{ij} is prior individual aptitude, $\overline{X}_j$ is class mean aptitude, and u_j is a random effect associated with school j. The test of the contextual effect is the test of the significance of γ.

As Aitkin and Longford (1986) point out, however, a model of the form of Equation 1 will often suffer from high collinearity and therefore poor precision. However, by reformulating the model with X_{ij} "centered" around its school mean, we have the model

$$Y_{ij} = \mu_j + \beta_w (X_{ij} - \overline{X}_j) + \beta_b \overline{X}_j + u_j + e_{ij} \qquad (2)$$

Equation 2 is really just a reparameterization of Equation 1, which can be seen by re-writing Equation 1 as

$$Y_{ij} = [\mu_j - \beta_w \overline{X}_j] + \beta_w X_{ij} + (\beta_b - \beta_w)\overline{X}_j + u_j + e_{ij},$$

from which it becomes clear that the contextual effect, γ of Equation 1, is equivalent to $\beta_b - \beta_w$ of Equation 2. The great advantage of the formulation in Equation 2 is the person-level variable, $X_{ij} - \overline{X}_j$, is orthogonal to the group-level variable, $\overline{X}_j$. In Equation 2, the test of the contextual effect can be accomplished by means of testing the difference between β_b and β_w. If the a difference between these two parameters is found, Equation 2 is retained as the model. If predictors are added to the Equation at both levels of aggregation, their effects will be adjusted for the within and between group effects of X. This is precisely the kind of adjustment that

Cronbach and Webb sought and that their predecessors failed to achieve.

Example 1

Using the HLM program's centering option, I estimated the model given by Equation 2 where Y is math achievement, X is socioeconomic status (SES), and $\overline{X}_j$ is school mean SES. The data are from the sophomore cohort of the High School and Beyond survey and invovle 158 schools and about 3000 students. The results were

Coefficient	Estimate	S.E.
β_b	5.32	.46
β_w	1.49	.14

To test the significance of the contextual effect, I employed the general linear hypothesis provision of HLM which enables one to test any contrast involving the fixed effects in a multilevel model. In this case, the test involves, in essence, the formula

$$(\hat{\beta}_b - \hat{\beta}_w)/S$$

where $S = [\mathrm{Var}(\hat{\beta}_b) + \mathrm{Var}(\hat{\beta}_w) - 2\mathrm{Cov}(\hat{\beta}_b, \hat{\beta}_w)]^{0.5}$. The test statistic took on the value of 62.69, to be compared with a χ^2 variate with one degree of freedom, indicating that the within- and between-school effects of SES are indeed different. Thus two parameters are needed to represent the relationship between X and Y, and by formulating the model as in Equation 2, the two effects in question are orthogonal, maximizing precision in estimation.

Cross-Level Interactions

Inferences of the type sought by Cronbach and Webb require comparison of within-group slopes. The study they criticized was based on biased estimates of within-group slopes; the failure to distinguish the between- and within- group slopes led to distorted estimates that were consistent estimates of neither. Actually, using the within-group centering option of HLM allows consistent estimation of within-group slopes even when the between-group model is misspecified! By using centering, the within-group predictors are orthogonal to all between-group predictors and so cannot be biased by a failure to include the appropriate between-group model. Of course the within-group model must be approriately specified.

Example 2

I reanalyzed the US High School data mentioned above with and without including mean SES as a predictor and with and without within-school centering. We shall assume that the model including mean SES is "correctly specified" based on the analysis in Example 1. (This does not imply that the model is "good" but just that the model with mean SES is better than the model without it.) The results are tabulated below. Note that the within-group slope estimate is "on the money" under centering even when mean SES is removed. Without centering, the within-group estimate is pulled (inappropriately) toward the "between group" estimate. Again the results are based on maximum likelihood estimation with variance components via HLM. In each case the parameter estimated is β_w.

	Uncentered	Centered
With mean SES	1.53	1.49
Without mean SES	1.71	1.49

Example 3

David Raffe at the Centre for Educational Sociology, University of Edinburgh, has been evaluating the effectiveness of an innovative educational intervention implemented at a number of sites. Each site has both program and non-program children. The foci of interest include not mainly the average effectiveness of the innovation but rather the extent to which the effect of program participation varies from site to site. They key issue here is to estimate consistently the effect of participation at each site–even though it may be impossible to incorporate most of the important "site" level predictors of the outcome. By working hard on specifying the student-level model and by centering the program participation variable within each site, the analysis can obtain a good estimate of the varying effect of the treatment across sites. Site-level variables ommitted from the model are orthogonal to the site-by-treatment interaction and therefore fail to bias its estimation. This feature is not present without centering within each site.

Random Slopes

Suppose now that the regression coefficients of Equations 1 and 2 are random. In Equation 1, the predictor, X, has not been centered within its school. This makes interpretation of the intercept problematic, because the intercept can be viewed as

$$\beta_{0j} = [\mu_j - \beta_{wj}\overline{X}_j]$$

Hence, the intercept is adjusted for the effect of X, but the coefficient of adjustment varies randomly from school to school. This also means that the variance of the intercept depends on the variance of the slope, and a hidden covariance between the two is also induced. A superior formulation uses the centered X. The the intercept is unambiguous–it's just μ_j and "separate" from the slope. By adding variables to predict the intercept and slope, one can monitor the reduction in variance associated with each. The analysis still adjusts for the school's mean SES, but now the coefficient of adjustment is β_b (which is appropriate given the model).

Conclusion

Judicious use of a centering option (centering around each "group's" mean) is an important tool for the multilevel data analyst. Such centering is not always desirable. For example, when slopes are fixed and no contextual effect is present, the best option is to center around the grand mean. Then each group's intercept is appropriately adjusted for the covariates. Sometimes, however, within-group centering can allow one to avoid the dangers so clearly described by Cronbach and Webb nearly 15 years ago.

References

Aitkin, M. & Longford, N. (1986). Statistical modeling issues in school effectiveness studies. *Journal of the Royal Statistical Society, Series A, 149*, 419-461.

Cronbach, L.J. & Webb, N. (1975). Between-class and within-class effects in a reported aptitude by treatment interaction: Reanalysis of a Study by G.L. Anderson. *Journal of Educational Psychology, 67*, 717-714.

Willms, J.D. (1986). Social class segregation and its relationship to pupils' examination results in Scotland. *American Sociological Review, 51*, 224-241.

CONTRIBUTORS

Thanks very much to the people who provided articles for this issue.

Tony Bryk

Department of Education
University of Chicago
5835 S. Kimbark
Chicago, IL 60637 U.S.A.

Jan de Leeuw

Department of Psychology
U.C.L.A.
405 Hilgard Ave.
Los Angeles, CA 90024

Pjotr Koopman

S.C.O.
Grote Bickersstraat 72
1013 KS Amsterdam
THE NETHERLANDS

Steve Raudenbush

Department of Education
Michigan State University
East Lansing, Michigan 48824
U.S.A.

Pieter van den Eeden

Fakulteit der Sociaal-Kulturele Wetens
Vrije Universiteit
Koningslaan 22-24
1075 AD Amsterdam
THE NETHERLANDS

The **Multilevel Modelling Newsletter** is published by the Multilevel Models Project at the Institute of Education (University of London). We welcome your suggestions and ideas, and we request reports on your work, other news, and questions regarding analysis issues. In particular, we would invite comments and/ or queries concerning the three longer reports in this edition.

Many thanks to the readers who returned the questionnaire in the January issue; we will try to follow up on as many of your suggestions as we can.

The newsletter is distributed free of charge (in January, May, and October), but reproduction and distribution are costly. We would ask that if you are able, to please send a contribution towards these expenses (say, £3 or US$5). Thank you.