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. 2 No. 1 January 1990

MULTILEVEL MODELLING INTO THE 1990S

The start of a new decade seems like a fitting time to do a bit of informal stocktaking concerning recent advances in multilevel modelling and to make a few predictions and suggestions about future developments. The growth of theory and practical experience has been rapid during what might be called the "new age of multilevel modelling" (defined here, a bit arbitrarily, as the period since the appearance of Mason, Wong, & Entwisle (1984), Bryk & Raudenbush (1985), Aitkin & Longford (1986), Goldstein (1986), and de Leeuw & Kreft (1986)).

Perhaps the most notable achievement to date is the most obvious one. *It is now a relatively straightforward matter to perform analyses that are far more sophisticated than those advocated even a decade ago,* thanks to efficient, friendly software that is relatively easy to obtain. Methodologists are sorting out interpretational concerns (e.g., the issue of centering), and applied researchers are grappling with relationships between substantive theory and complex statistical models. Statisticians are generalizing the well-established two/ three-level continuous-response models–witness, for example, the methodology now available for discrete data.

"The word" is speading, and the community of practitioners is growing at a satisfying rate. While much of the developmental work continues to be within the field of education, researchers in a variety of other fields are starting to apply the techniques. A communications professor, for example, has recently conducted a study dealing with news ideology of articles, nested within newspapers.

Here are my own predictions for developments over the next five years.

- Software for multilevel analysis will become much more accessible through development of new user interfaces. Interactive help systems will facilitate model specification. An existing major software package will offer a *general* multilevel procedure.
- Diagnostic procedures analogous to those available for single level models will be implemented.
- An increasing number of major educational (and other) studies will be conceptualized within a multilevel framework and designed to facilitate construction of sophisticated models. Database software designed specifically for large-scale users of multilevel modelling will be written.
- Educational jurisdictions will more frequently use multilevel modelling in comparing institutions and trying to determine what makes them effective for different students. There will be no shortage of controversy. Some particularly effective methods for presenting analysis results to lay audiences will emerge.
- Wider use of multilevel modelling will be made in the analysis of survey data.
- Graduate courses and textbooks focussing on multilevel modelling will be readily available.

Readers of the *Newsletter* are invited to submit their views about probable/ needed developments in the years ahead, and we will publish them in a future issue.

Bob Prosser

THEORETICAL

MULTILEVEL MODELS FOR DATA WITH A NON-NORMAL DISTRIBUTION

Nick Longford

Normal distribution of the random terms in multilevel analysis is an important restrictive assumption. Much of the observational data in the social sciences are inherently discrete, and in the extreme, binary (e.g., Yes/ No responses to survey questions). For such data the normal linear multilevel analysis is not appropriate not only because of the violation of the assumption of normality, but also because we usually wish to use a nonlinear scale such as the logit for binomial data, logarithm for Poisson data, etc. It is therefore desirable to have an extension of the multilevel methods for a wider class of distributional assumptions, which would at the same time be an extension of the methods for regression analysis of *independent* non-normally distributed data.

In a standard interpretation of the (normal) two-level models we consider varying within-group regressions. Using the familiar GLIM terminology, a natural extension for the *generalized linear models* is to consider within-group regressions given by a specified error distribution (or functional relationship of the variance of an observation on its mean) and link function (such as binomial and logistic):

$$\begin{aligned} \mathrm{E}(Y_{ij}|\mathbf{b}_j) &= \mathrm{h}(\mathbf{X}_{ij}\mathbf{b}_j) \\ \mathrm{Var}(Y_{ij}|\mathbf{b}_j) &= c\mathrm{H}(\mathrm{E}(Y_{ij}|\mathbf{b}_j)). \end{aligned} \qquad (1)$$

Here h and H are the link and variance functions, respectively, $j\ (= 1, \ldots, N_2)$ and $i\ (= 1, \ldots, n_j)$ are the group and elementary unit identifiers, respectively, and c is a scale parameter (in many cases equal to 1). Next we assume that the vectors of within-group coefficients $\mathbf{b}_j$ are normally distributed:

$$\mathbf{b}_j \sim \mathrm{N}(\beta, \Sigma) \qquad (2)$$

The parameterization for Σ is subject to the analyst's choice. In particular several variances within Σ can be set to zero. The so-called contextual models assume a normal regression model for $\mathbf{b}_j$, but these group-level regressors can be absorbed within $\mathbf{X}_{ij}$ in (1). Thus (1) and (2) have the same flexibility and generality as the normal two-level models. In fact, the normal models are a special case with $\mathrm{h}(z) = z$, $\mathrm{H}(z) = 1$, and $c = \sigma^2$. They simultaneously generalize the multilevel models to non-normal assumptions and the GLIM models for non-independent outcomes. Extension of the two-level model to more levels is straightforward and fully analogous to the normal case.

Maximum Likelihood Estimation

Exact maximum likelihood estimation with a model given by (1) and (2) and its multilevel extensions is a formidable task. For the simplest pattern of variation in (2) which corresponds to straight variance components or the compound symmetry models, Anderson and Aitkin (1985) applied the EM algorithm for analysis of binary outcomes and discussed the associated computation problems. For larger data sets and more complex patterns of variation a large number of multivariate integrals would have to be evaluated numerically at each iteration, and the number of iterations could be substantial.

An approximate method for fitting the compound symmetry submodel of (1) and (2) for the logistic regression with binary data has been proposed by Williams (1982). Morton (1987) designed a moment-like estimation procedure for the three-level model with Poisson-logarithmic assumptions. These and other methods are applicable only to a very specific situation.

Goldstein (1989) and Longford (1988) have constructed algorithms for approximate maximum likelihood estimation of generalized multilevel models (1) and (2). These algorithms are essentially reweighted versions of tha algorithms for normal multilevel analysis (IGLS, EM, or Fisher-scoring). Goldstein (1989) derived his algorithm by a linearization procedure, while Longford's (1988) uses approximate integration of the conditional quasilikelihood. The quality of the approximation requires further research, but the implementation is relatively easy. The binomial/ logistic, Poisson/ logarithmic, and gamma/ reciprocal combinations of error distribution and link function are implemented in the VARCL software. Full implementation of this extension, however, would be realized only with installation of a module similar to the $OWN in GLIM in which the analyst would have comlete flexibility in the selection of the error distribution and link function.

DEVELOPMENTS

FITTING LOGLINEAR ML MODELS

Harvey Goldstein

A simple loglinear 2-level model can be written as

$$p_{ij} = \exp(\beta\mathbf{X} + u_j) + e_{ij}$$

This is the standard single level formulation with the addition of a simple random term, u_j, varying across level 2 units. The $\mathbf{X}$ matrix might consist of a set of dummy variables defining main effects and interactions in a contingency table or, for example, a set of explanatory variables for individuals with a binary (0,1) response denoting presence or absence of an attribute.

As an illustration, suppose that in a study of student achievement, a dichotomous variable for performance (high / low scoring) is examined in relation to to a binary absenteeism variable (often / seldom absent). The following 2 x 2 table shows the frequencies for the jth school:

	Often Absent	Seldom Absent
High Scoring	n_{11j}	n_{12j}
Low Scoring	n_{21j}	n_{22j}

An additive model can be written

$$p_{ij} = \exp(\beta_{0j} + \beta_1 X_{1ij} + \beta_2 X_{2ij}) + e_{ij} \qquad (1)$$

where X_{1ij} and X_{2ij} are dummy variables for the row and column classifications. Here, $\beta_{0j} = \beta_0 + u_j$ where β_0 is a fixed parameter and u_j is random as before. There are four level 1 units (cells) per level 2 unit, indexed by i and

$$p_{ij} = n_{ij}/n_j$$

where n_j is the total number of students in the jth school. Equation (1) implies that there is no relationship between performance and absenteeism within each school. It also assumes that only the overall mean proportion (β_{0j}) varies from school to school. Generally this will be unrealistic since we would expect both the proportion with high scores and the proportion often absent to vary also. Thus a more realistic model is as follows

$$p_{ij} = \exp(\beta_{0j} + \beta_{1j} X_{1ij} + \beta_{2j} X_{2ij}) + e_{ij} \qquad (2)$$

where all three coefficients are random at level 2, with corresponding variances and covariances to be estimated.

If we now make the usual statistical assumptions that the proportions have a multinomial distribution and that the coefficients are distributed normally across level 2 units, we can obtain estimates of the required parameters, that is the variances and covariances of the coefficients and the mean proportions. A goodness of fit test can be obtained and tests of significance and confidence intervals constructed.

Both VARCL and *ML3* allow such models to be fitted readily, as well as corresponding logit linear models for binary response data. VARCL also allows a *gamma link function* to be specified, thus for example, allowing models in which the response variable is a variance. *ML3* can also accommodate such models.

In VARCL the user selects the link function she wishes to use from a menu and is then prompted for information on the response and explanatory variables. In *ML3* the procedure involves the user defining new explanatory variables which are functions of the fixed predictor, and which change from iteration to iteration. These can be specified conveniently using a new *macro facility* which allows automatic updating of explanatory variables and options at each iteration. In addition *ML3* allows the user to specify a distribution for the proportions which is not multinomial: the variances and covariances of the proportions are required only to be inversely proportional to the total number of cases in the level 2 unit. In this way extra-multinomial variation can be handled. This is sometimes the case, for example, with spatial data.

Full details will be incorporated into new documentation for *ML3*, and the theory is given in a paper which has been submitted for publication.

With these extensions to the multilevel model, most kinds of complex survey data can now be modelled efficiently and informatively and a number of survey practitioners are showing considerable interest in this possibility.

PROJECT

FUTURE PLANS FOR THE MULTILEVEL MODELS PROJECT

The Economic and Social Research Council (ESRC) of Great Britain has recommended that funding be awarded to the Multilevel Models Project for a three year extension beginning in March 1990.

The basic work of the new project will be that of the current one, namely development of multilevel methodology in theoretical and practical directions and dissemination of information about the technique. Instructional workshops will be conducted for researchers, and the *Newsletter* will continue to be published regularly. The writing of a practical book on multilevel data analysis is proposed. An additional user-support dimension will be the exploration of methods to add a knowledge enhancement/ sophisticated help facility to *ML3*.

Proposed areas for further theoretical development include the following:

- measurement error in explanatory variables;
- multilevel time series;
- robust estimation;
- optimal design for multilevel studies; and
- random cross-classifications.

Each of these will be discussed briefly.

Measurement Error

The present methods involve the assumption that the explanatory variables are measured without error. This assumption is often violated leading to bias in parameter estimates. While the basic theory for measurement error in level 1 explanatory variables has been developed (see Goldstein, 1986), it has not yet been implemented or applied. It also needs extending to deal with measurement error in higher level explanatory variables. This problem occurs, for instance, if an aggregate variable is computed from partial data in a higher level unit, e.g., when the mean pretest score for a class is based on only a sample of the students in the class.

Multilevel Time Series

In some longitudinal studies–human growth studies, for example–the interval between repeated measurements may be very short. This results in violations of the assumption of independence of level 1 random variables across level 1 units. It appears that embedding time series models within a multilevel framework is a useful generalization of traditional time series analysis. Multilevel models which include parameters for the autocorrelations are being developed.

Robust Estimation

Apart from the provision of flexible procedures in *ML3* for studying model residuals, little work has been done in on the issue of robustness of estimates to violations of assumptions and the presence of outliers. This is an important area to explore. A promising approach seems to lie in applying some of the procedures used with ordinary linear models during the part of *ML3*'s (iterative generalized least squares) estimation cycle in which the residuals from the fixed part of the model are computed.

Optimal Design

Some users who are familiar with multilevel analysis are planning studies with a view to conducting multilevel analyses of the data. While general suggestions can be given, it is not clear at the moment how to design a study which permits an optimal analysis of the data. At a basic level, work is needed on the relative proportions of units at the different levels for efficient estimation. Because random as well as fixed parameters are estimated, there are considerations additional to those which have been studied for single level models.

Random Cross-classifications

An example of data with a randomly cross-classified structure would be information about students within cells that are defined by the crossing of identifiers of their schools and residential neighbourhoods. (Students are thus nested simultaneously within units of two types at the same level.) Special models are needed to estimate the contribution to variance at a particular level from the different types of units at that level.

It appears at the moment to be difficult to develop an estimation algorithm which is fast and uses computer memory efficiently. The plan for the new project is to implement routines for fitting the simplest variance components models, to carry out data analyses using these, and to continue the search for an efficient general algorithm.

NEWS

Several practical and interpretational issues will receive attention in the new project. Two of these are boundary value problems and complex level 1 covariance structures.

Boundary Value Problems

Experience with data sets has raised numerous problems and possibilities associated with estimation of the random parameters. First, when the number of such parameters at any level is large relative to the number of units at that level, the response surface in the neighbourhood of the GLS or ML estimates may be relatively flat. Typically, a response surface of this type seems to have little effect on the fixed coefficient estimates but raises problems in the interpretation of the random parameter estimates. One computational variation explored a bit in the current project–selective "freezing" and "thawing" of some individual parameters' estimates during "early" iterations–appears to avoid some instability problems but requires more extensive investigation before recommendations can be made to researchers.

Complex Level 1 Covariance Structures

It is possible to specify coefficients in a model as varying randomly across level 1 units. In this way, different variances for boys and girls, say, or different types of schools can be accommodated. This allows a more precise model specification and provides a flexible procedure for studying how the between-individual variation changes as a function of further variables. The interpretation of estimates from such models will be considered in more depth in the new project.

Researchers' suggestions of problems and issues along the way will shape the course of the project as well. It promises to be an exciting three years.

MULTILEVEL CAUSAL AND FACTOR MODELS PROJECT: AN UPDATE

Rod McDonald

A brief account of this project was given in the January 1989 issue of the *Newsletter.* Funding for the project, running from April 1988 to March 1989 at Macquarie University has been provided by an Australian Commonwealth Research Grant, mainly covering the appointment of a Research Fellow, Pius Lam, who is primarily responsible for computer programming. Prue Parker, a Ph.D. student, is exploring the applicability of the theory.

Goldstein and McDonald (1988) contains the general foundations for the project, while McDonald and Goldstein (1989) contains a detailed treatment of the two-level model for causal relations with latent variables on which all the present research is being concentrated. In this model–the *Reticular Action Model* (RAM)–a simple path-graphic representation of causal relationships between observed variables or factors defined on level 1 and level 2 sampling units (e.g., students and schools) translates directly into a simplest possible model for two-level linear structural relations. McDonald and Goldstein (1989) shows that the case of a balanced sampling design (say equal numbers of students per school) yields simple sufficient statistics for fitting the resulting covariance structures, as well as a suitable discrepancy function and large-sample test of the fit of the model. The unbalanced case remains much more complex, requiring estimation of the mean vectors even when these are not in any way constrained by the model.

The work is being carried out along the following lines: A program for the RAM model for ordinary single-level data minimizes the likelihood-based discrepancy function by a quasi-Newton technique using numerical gradients and sparse matrix methods considered desirable in the two-level programs. A program for the two-level RAM model (as in McDonald and Goldstein (1989) for balanced data) has been developed on the same basis. The corresponding program for unbalanced data also uses numerical gradients and a quasi-Newton minimization method applied to a function of the likelihood. It is expected that these programs will guide the development of counterparts which employ analytic gradients.

From work done so far, it seems clear that for applications to emperical data the programs will need to be supplied with good guessed values of the parameters. Devices for obtaining good starting values will be tested soon. Extension of the theory to cover missing data and the incorporation of fixed explanatory variables will probably constitute the last phase of the project.

BOOK REVIEW

Analysis of Complex Surveys.
Edited by C. J. Skinner, D. Holt & T. M. F. Smith (1989). Wiley. £38.50. Hardcover. 306 pages.
Dougal Hutchison

This book has its origins in a conference funded by the SSRC in Southampton in 1975 and research projects taking place since then. One's first fear is that this might be simply a somewhat belated collection of conference papers, but, fortunately, this turns out to be quite unjustified, though a reference to the Ministry of Education does make one fear for its age. All the papers have undergone extensive rewriting, references have been updated, substantial additional work has taken place, there is an index, and perhaps most important, there has been a serious attempt to produce a unified conceptual framework and overview for the contributions.

All statisticians working in social or educational research will be familiar with the use of complex surveys, and the standard texts and techniques to allow for the fact that we were not using SRS sampling, with its inherent justification of the iid assumption. However much of the development took place in government and market research, where the main interest was in estimating descriptive parameters of a population from a usually much smaller sample, while what we tended to be interested in was exploration of relationships, typified by regression analyses. The effects on regression analysis were not covered in the standard texts, and while the authors' methods could presumably be extended to calculate sampling errors for regression coefficients, their hearts didn't seem really to be in it. Besides, weren't we interested in the error relating to the model, rather than that relating to sampling?

Such model-related topics are covered by the authors in this book. As they argue, the standard Design Effects are aimed more to compare survey designs, whereas by the time we have reached the analysis stage, the survey has already taken place. Further, where there is (for example) heteroscedasticity, the iid assumptions are breached even for an SRS sample. The authors generalise the design effect or deff to a misspecification effect (or meff), which takes account of the model-based error assumptions as well.

The book is divided into three sections, each with an introduction. Sections A and B deal with Aggregated Analysis, i.e., where the model is defined at the population level, and the population and survey structure are viewed basically as a nuisance, while Section C extends the model to include not just the survey variables, but also variables used in the survey design which are related to these survey variables and other explanatory variables. Each section contains chapters dealing with continuous and discrete variables separately, and the first two sections also have chapters dealing specifically with multivariate aspects.

Section A deals with Standard Errors and Significance Tests, and can perhaps be thought of a welcome update extending the Cochran/Kish tradition to include model-based error. Section A contains chapters on:

- Domain Means, Regression and Multivariate Analysis *C. J. Skinner*
- Chi-squared Tests For Contingency Tables *J. N. K. Rao & D. R. Thomas*
- Measures of Association for Contingency Tables *E. A. Molina*

Section B covers Point Estimation and Bias. Of particular interest are the discussions on non-response effects and ignorable and uninformative sampling schemes, and the vexed question of whether or not to weight the sample according to selection probabilities when carrying out analyses. This section contains chapters on :

- The Effect of Selection on Regression Analysis *G. Nathan & T. M. F. Smith*
- Multivariate Analysis *T. M. F. Smith & D. J. Holmes*
- Selection Based on the Response Variable in Logistic Regression *A. J. Scott & C. J. Wild*

Section C, the one dealing with Disaggregated Analysis, ought to be of particular interest to readers of this Newsletter, since the description 'Disaggregated Analysis' corresponds basically to to multilevel modelling. It is always of interest to produce a synthesis and framework linking areas generally considered separately. The attempt to reconcile different approaches to the same problem and to view a

continued on page 7

SOFTWARE DEVELOPMENTS

ML3 FOR VAX IS NOW READY

In response to requests from users with large data sets, an implementation of *ML3* has been developed for DIGITAL VAX computers. The program can be used on any model of VAX provided the machine is running under the VMS operating system. The VAX was chosen for the Multilevel Models Project's entry into the mainframe world because of its popularity, but implementations on other computers will be considered in the future.

The price for a single copy is £500 (US$800). For further details about the program and information about ordering, please contact Bob Prosser at the address on the masthead.

NEW DISTRIBUTION POLICY FOR ML3

The Multilevel Models Project has updated its distribution policy for the PC implementation of *ML3*, effective January 1, 1990. During the program's developmental phase, researchers who requested the software were supplied with a copy and asked to "register" by paying a nominal (£50) fee. Those who did received upgrades as they were produced. Participants in the Project's (free) workshops received copies at no charge. This distribution approach introduced many people to multilevel analysis and provided important feedback, facilitating testing and improvement of the program.

The new distribution policy can be described as "three-level." (1) Participants in the Project's workshops held in London, will continue to receive copies of the software at no charge. (2) The *regular* single-copy price–which covers basic telephone support–is £200 (US$320). (3) Researchers affiliated with a university will receive a discount of 40% on their orders. Substantial discounts will be given to multi-copy users, e.g., instructors of statistics courses.

Significant improvements in the program will be introduced in 1990, and these will be available as optional extensions. High-resolution graphics similar to *ML2*'s will be available in several months. The capacity to take into account measurement error in explanatory variables will be introduced, as will a facility for weighting.

People who registered as users under the former policy (before December 31, 1989) will be offered a special price if they wish to purchase these.

SOFTWARE COMPARISON AT AERA

Session 53.02 (April 20, 1990 at 10:35 am) of the annual meeting of the American Educational Research Association will feature a presentation of the report by Ita Kreft et al. comparing four programs for multilevel analysis. Discussion by developers will follow.

Book Review continued from page 6

familiar area from a new perspective leads to new insights and to questioning of assumptions previously overlooked. This section contains chapters on:

- Multilevel and Multivariate Models in Survey Analysis *H. Goldstein & R. Silver*
- Regression Models for Stratified Multi-Stage Cluster Samples *D. Pfefferman & L. Lavange*
- Logistic Models for Contingency Tables *D. Holt & P. D. Ewings*

This book raises a number of rather disconcerting questions. How many of today's accepted results from social science research would still survive if analysed by the more reliable methods advocated in this book? If the answer is not many, how can we avoid making any more comparable errors? It would obviously be very important to persuade the major package distributors to include these methods in their products. Specialised programs are valuable and welcome, but despite continuing attempts to make them more accessible, they are still only usable by specialist statisticians and by researchers with a degree of training. Yet there is very little training at first or higher degree level in the mathematical and statistical aspects of research methods in this country.

The authors have extended our treatment of 'error' from people-sampling error only to include misspecification error as well. What happens when the next extension comes? This is by no means an idle question. The Assessment of Performance Unit science project found that when they took account of item sampling as well as people sampling, the sampling errors were trebled–sufficient change to play havoc with anyone's conclusions.

The Analysis of Complex Surveys is valuable and interesting, and should prove important reading for professionals in the field.

CENTERING: A POSTSCRIPT (?)

Ian Plewis

There is, I think, a measure of agreement between Steve Raudenbush, Nick Longford and myself, as a result of Raudenbush's response to the other two pieces in the October 1989 issue of the *Newsletter* (see also May 1989). In particular, we all agreed on the importance of linking model specification to the research question to hand, rather than seeing it as an unconnected technical problem which can always be solved in the same way. Nevertheless, I would like to clarify the point I made about modelling strategies or, more modestly, modelling tactics. Raudenbush correctly states that the two sets of equations (1) and (2) are equivalent, i.e.

$$Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + e_{ij} \qquad (1a)$$

$$\beta_{0j} = \theta_{00} + \theta_{01}\overline{X_j} + u_{0j} \qquad (1b)$$

and

$$Y_{ij} = \beta_{0j} + \beta_1(X_{ij} - \overline{X_j}) + e_{ij} \qquad (2a)$$

$$\beta_{0j} = \theta_{00} + \theta_{01}\overline{X_j} + u_{0j} \qquad (2b)$$

However, by themselves, the level one equations, (1a) and (2a), are not of course the same. I believe that, before attempting to model level two (and above) variation in the intercept, we need to establish that it exists. After all, in the school effectiveness literature, for example, school differences are often small and it has been argued that individual (or level one) effects masquerade as context (or level two) effects in poorly specified models. Thus, we ned to specify our individual level model as completely as possible. If we fail to eliminate level two variation in the intercept with our level one model, we might then adopt Raudenbush's group-centered approach. Note, however, that there will be data sets for which, given possibly more than one explanatory variable in (1a), equation (1b) will be redundant because β_{0j} does not vary. On the other hand, (2b) cannot be omitted from (2), even when there is no variation in the intercept, because $\theta_{01} = \beta_1$. It is in this sense that Raudenbush's specification does seem to prejudge the issue. This may not matter for researchers familiar with multilevel models, but it could confuse some who are less experienced.

CONTRIBUTORS

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

Dougal Hutchison

National Foundation for Educational Research, Slough, U.K.

Nick Longford

UCLA, Los Angeles, CA , USA

Rod McDonald

Macquarie University, North Ryde, NSW, AUSTRALIA

Ian Plewis

Thomas Coram Research Unit London

References

Anderson, D.A., & Aitkin, M. (1985). Variance component models with binary response: Interviewer variability. *Journal of the Royal Statistical Society, Series B, 47,* 203-210.

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

Goldstein, H. (1989). *Non-linear multilevel models.* Manuscript submitted for publication.

Goldstein, H. & McDonald, R.P. (1988). A general model for the analysis of multilevel data. *Psychometrika, 53,* 45-467.

Longford, N.T. (1988). A quasi-likelihood adaptation for variance component analysis. *Proceedings of the Statistical Computing Section, American Statistical Association,* 137-142.

McDonald, R.P. & Goldstein, H. (1990). Balanced versus unbalanced designs for linear structural relations in two-level data. *British Journal of Mathematical and Statistical Psychology, 42,* 215-232.

Morton, R. (1987). A generalized linear model with nested strata for extra-Poisson variation. *Biometrika, 74,* 247-258.

Williams, D.A. (1982). Extra-binomial variation in logistic linear models. *Applied Statistics, 31,* 144-148.