THE UNIVERSITY of York

CENTRE FOR HEALTH ECONOMICS

MULTILEVEL MODELS AND HEALTH ECONOMICS

Nigel Rice and Andrew Jones

CHE Technical Paper Series 4

MULTILEVEL MODELS AND HEALTH ECONOMICS

Nigel Rice and Andrew Jones

March 1997

CENTRE FOR HEALTH ECONOMICS TECHNICAL PAPER SERIES

The Centre for Health Economics has a well established Discussion Paper series which was originally conceived as a means of circulating ideas for discussion and debate amongst a wide readership that included health economists as well as those working within the NHS and pharmaceutical industry.

The introduction of a Technical Paper Series offers a further means by which the Centre's research can be disseminated. The Technical Paper Series publishes papers that are likely to be of specific interest to a relatively specialist audience, for example papers dealing with complex issues that assume high levels of prior knowledge, or those that make extensive use of sophisticated mathematical or statistical techniques.

The series provides a maximum of 20 pages for authors. The content and its format are entirely the responsibility of the author, and papers published in the Technical Paper series are not subject to peer-review or editorial control, unlike those produced in the Discussion Paper series. Offers of further papers, and requests for information should be directed to Frances Sharp in the Publications Office, Centre for Health Economics, University of York.

THE AUTHORS

Nigel Rice is a Research Fellow in the Centre for Health Economics, University of York. Andrew Jones is Director of the Graduate Programme in Health Economics, Department of Economics and Related Studies, University of York.

Price £4.00

SUMMARY

Multilevel analyses have become an accepted statistical technique in the field of education where over the past decade or so the methods have been developed to explore the relationships between pupil characteristics and the characteristics of the schools they attend. More recently, widespread use has extended to other social sciences and health research. However, to date, little use has been made of these techniques within the health economics literature. This paper presents an introductory account of multilevel models and describes some of the areas of health economics research that may benefit from their use.

KEY WORDS - multilevel models, hierarchical models, fixed and random effects

JEL CLASSIFICATION: I1, C1

INTRODUCTION

There has been a long standing interest within the social sciences in the association between individual level characteristics and higher level factors such as community effects or societal/geographical contextual effects (see for eg., Blalock,[1] DiPrete & Forristal[2]). More recently, this interest has extended to relationships between individuals and characteristics of the institutions to which they belong, most notably in assessing 'effectiveness' or 'performance' of public institutions.[3] The investigation of such relationships has historically posed difficult analytical problems due to a lack of sufficient 'micro' level data. Ensuing studies have often relied on data aggregated beyond the basic unit of analysis, resulting in analyses that are often prone to criticisms associated with the ecological fallacy[4,5] ('macro' level associations are wrongly infered to exist at the 'micro' level). More recently there has been a general shift from placing emphasis on 'macro' level agents to incorporating both 'micro' and 'macro' level relationships and observations in empirical work. This, in part, has been due to the advancement of statistical techniques that adequately address the methodological problems imposed by the hierarchical or nested structure inherent in the requisite data. Since individuals are clustered within higher level groups (for example, households or general practices), which themselves may be clustered within further groups (geographical areas, or administrative districts), a correlation structure is imposed on the data which invalidates classical assumptions of independence and hence, common techniques such as ordinary least squares (OLS) will be inefficient. Further, a single level approach to the analysis of such hierarchically clustered data may fail to fully exploit the richness of information contained within and between the various levels being considered.

Over the past decade theoretical developments enabling nested data structures to be analysed within a generalised linear models framework have been progressed under various guises, most notably multilevel models,[6] random coefficient models[7] and hierarchical linear models[8] (hereinafter multilevel models specifically will be discussed in this paper). However, until recently a lack of appropriate software required to perform such analysis rendered these techniques inaccessible to a wide audience. Multilevel models are generally used to analyse data that fall naturally into hierarchical structures consisting of multiple macro units (contexts) and multiple micro units within each macro unit. Emphasis is placed on defining and exploring variations at each level of the hierarchy after conditioning on the set of explanatory variables of interest. The primary focus therefore lies in the estimation of the magnitude of variances operating at different levels and how such variances are related to explanatory variables. This contrasts with the more familiar

econometric approach of viewing higher level effects as nuisance parameters to either condition on or adjust for to achieve desirable properties of fixed part parameters. This is most common where it is believed, a priori, that one or more explanatory variable is correlated with the higher level effects. In such circumstances, to obtain consistent estimates of the fixed part parameters, higher level effects are often modelled in order to avoid endogeneity bias (we expand on this issue later in the paper). The explicit modelling of variances associated with the errors at each level of the hierarchy in a multilevel model is of direct interest.

Recent widespread availability of specialist software packages, a useful review of which is given by Kreft et al,[9] enabling the routine use of multilevel techniques has prompted researchers from various disciplines to adopt a hierarchical approach in their work. Today, examples of the application of these methods can be found in such diverse areas as education,[10] criminology,[11] sociology,[2] geography,[12,13] epidemiology,[14] health policy and health services research.[15,16] It is proposed that multilevel models can also offer health economists a valuable tool facilitating greater substantive insight into the processes generating research data and will compliment existing more familiar econometric techniques. A further shift in emphasis from aggregate to micro level data as predicted in a recent editorial appearing in Health Economics[17] will only serve to encourage the application of methods such as these.

This paper presents an introductory account of multilevel models, including a brief discussion of areas of health economics research that could potentially benefit from the use of these techniques.

BASIC MULTILEVEL MODEL

To illustrate the basic structure of a multilevel model we take as an example the most simplistic model consisting of two levels which may represent patients nested within provider units (see Figure 1). y_{ij} represents the response of interest which is related to a vector of explanatory variables $\mathbf{x}$ in the following manner:

$$y_{ij} = g(\beta' \mathbf{x}_{ij}) + w_{ij} \qquad (1)$$

where $w_{ij} = u_j + e_{ij}$, and $g()$ represents a link function to specify the function form of choice. The following discussion outlines a model with a linear link function, but may be generalised to nonlinear models in a straight forward way.

Subscript i (i = 1,...,N) represents individuals (patients) and subscript j (j = 1,...,M) provider units. In this specification the overall error term (w_{ij}) is decomposed into e_{ij} and u_j, where e_{ij} is the random error term for the i-th patient within the j-th hospital, assumed to have zero mean and constant variance σ_e^2. The effects of hospitals are estimated through u_j which is assumed random and again has a mean of zero and constant variance σ_u^2. We also have $\mathrm{cov}(e_{ij}, u_j) = 0$, that is, we assume that patients and hospitals are uncorrelated. This appears a reasonable assumption to make but it may be restrictive in some applications. For illustrative purposes, the indexing of the set of explanatory variables in model (1) assumes a set of variables measured at the level of the patient. However, this is for notational convenience only and often variables measured at all levels are of interest (a variable measured at level 2 is usually indexed as x_j); both in terms of their

associated parameter estimates and their effects on the estimates of variances.

For the i-th patient within the j-th hospital model (1) leads to an expectation of, $E(y_{ij}|\mathbf{x}_{ij}) = \beta'\mathbf{x}_{ij}$, and a conditional variance of $var(y_{ij}|\beta'\mathbf{x}_{ij}) = \sigma_w^2 = \sigma_e^2 + \sigma_u^2$, and hence, the overall variance is 'partitioned' into components for both hospitals and patients. The partitioning of the variance in this manner leads to the 'intra group correlation coefficient' $\rho = \sigma_u^2(\sigma_u^2 + \sigma_e^2)^{-1}$ (note, $cov(u_j + e_{i_1 j}, u_j + e_{i_2 j}) = cov(u_j, u_j) = \sigma_u^2$) which measures the strength of 'nesting' within the data hierarchy and is fundamental to the estimation procedures. In the basic formulation presented here, model (1) is often referred to as a variance components model. The analogous models found in the econometrics literature are referred to as 'error component' models (see, for example, Judge et al[23]) and in particular, examples can be found in the context of nonlinear budget sets (Hausman,[24] and Moffitt[25]).

In the presence of a non-zero intra-group correlation, estimation proceeds through the use of generalized least squares. For each of the 'higher level' groups j, we define a covariance matrix; $cov_j(y_{ij}) = V_j = \sigma_e^2 I_{n_j} + \sigma_u^2 J_{n_j}$, where I is the identity matrix, J a matrix of ones and n_j the number of individuals within the j-th group. Extending this over all higher level groups (hospitals) results in a block diagonal structure as depicted in Appendix I.

PARAMETER ESTIMATION

Various estimation routines have been developed for the analysis of hierarchical data structures, however here we concentrate on those specific to multilevel models developed by Goldstein.[6] Model estimation is based upon an iterative generalised least squares (IGLS) procedure that results in consistent and asymptotically efficient estimates of β. Using matrix notation, estimates of the fixed coefficients and their associated covariances are given in the familiar way:

$$\hat{\beta} = (X^T V^{-1} X)^{-1} X^T V^{-1} Y \qquad (2)$$
$$\text{and } cov(\hat{\beta}) = (X^T V^{-1} X)^{-1}$$

where V is the block diagonal covariance matrix. However, V is unknown and the method of obtaining starting values for β in its absence is achieved by performing an OLS estimation for the first iteration (i.e. assuming $V = I\sigma^2$). Once this is done, a vector of raw residuals $\tilde{W} = \{\tilde{w}_{ij}\} (\tilde{w}_{ij} = y_{ij} - \hat{\beta}'\mathbf{x}_{ij})$ is formed together with the cross-product matrix: $(\overset{*}{W} = \tilde{W}\tilde{W}^T)$. Each $\tilde{w}_{ij}^2$ is an estimator of $\sigma_w^2 (= \sigma_e^2 + \sigma_u^2)$ and each $\tilde{w}_{ij}\tilde{w}_{ik}$ $(j \neq k)$ is an estimator of σ_u^2; hence for the *j-th* higher level unit:

$$\hat{V}_j = \tilde{w}_j \tilde{w}_j^T$$

We write $\overset{*}{E(W)} = V$ and $\overset{**}{W} = vec(\overset{*}{W})$, where vec is the vector operator stacking the columns of $\overset{*}{W}$ underneath each other. This is then regressed on $I_{n_j} + J_{n_j}$ to obtain consistent estimates of σ_e^2 and σ_u^2.

The estimates of σ_e^2 and σ_u^2 are used to construct *V* which in turn is used in a GLS estimation of the fixed part parameters. Once the fixed coefficients are obtained, updated residuals are formed, and the random parameters estimated once again. This procedure is repeated until some specified convergence criteria is reached. At convergence, assuming multivariate Normality the estimates are maximum likelihood. For further details the reader is refered to Goldstein[6,26] and Hsiao.[27]

In general IGLS leads to biased estimates of the random parameters (in the models presented thus far: $\hat{\sigma}_e^2$ and $\hat{\sigma}_u^2$) since no account is taken of the sampling variation of the fixed part explanatory variables. This may be problematic in small samples. A restricted iterative generalised least squares (RIGLS) or restricted maximum likelihood approach involving a slight modification of IGLS leads to unbiased estimates (refer to Goldstein[28] for full details). This involves the addition of a correction term to the estimation of V at each iteration based on the set of explanatory variables and the previous estimate of V. Under multivariate Normality and assuming $V = \sigma^2 I$, applying the correction term leads to the usual formula for the unbiased estimate of σ^2 given by $\hat{\sigma}^2 = s^2/(n-p)$, where p is the number of explanatory variables. Efficient computational methods for these techniques are given by Goldstein and Rasbash.[29]

The above estimation procedures are particular to multilevel models proposed by Goldstein.[6] Other estimation procedures developed specifically for hierarchical data structures have been proposed elsewhere. Longford[7] developed a procedure based upon a 'Fisher scoring' algorithm which has been shown to be formally equivalent to IGLS[6] and forms the basis of the program VARCL.[30] Bryk and Raudenbush[8] describe the use of an EM algorithm based on an 'empirical bayesian' approach and incorporates this in the program HLM.[31] Another approach is that of Generalized Estimating Equations (GEE) by Liang and Zeger;[32] however this is principly concerned with estimates of fixed part parameters (for explanatory variables) rather than exploring the random part. A short review of these procedures is given by Goldstien[6] (p23).

RESIDUAL ESTIMATION

Often in empirical analyses, the size of the higher level effects in addition to the estimates of variances are of interest particularly in exploratory analyses, in the reporting of results and for diagnostic purposes. These are obtained in the following manner.

For the linear two level multilevel model:

$$y_{ij} = \beta' \mathbf{x}_{ij} + u_j + e_{ij}$$

the level 2 mean group effect for the *j-th* group may be estimated by the within level 2 sample mean $\overline{w}_j = \sum_{i \in j} \tilde{w}_{ij} / n_j$, where $\tilde{w}_{ij} = y_{ij} - \hat{y}_{ij}$, or it could be estimated by the overall sample mean $\overline{w} = \sum_i \sum_j \tilde{w}_{ij} / N$, where N is the total sample size. If the between level 2 unit variance were equal to zero (i.e., no systematic difference among level 2 means) then the overall mean would represent a suitable estimate for each of the level 2 effects, and if the variance within level 2 units is represented by σ^2, then the overall sample mean has a variance σ^2 / N. If this is not the case and the level 2 variance is large compared to the total variance, then an unbiased estimate for each of the level 2 means is provided by the within level 2 sample mean. However, this estimate, (with a variance of σ^2 / n_j), although being unbiased is less efficient. Most applications will consist of intermediate cases where it may be possible to do better than simply to use the 'raw' level 2 group sample means (based on mean squared errors) and the two estimators can be linearly combined in the following way to define 'shrunken' estimates of the level 2 group means (denoted $\overline{w}_{j,s}$):

$$\overline{w}_{j,s} = s_1 \overline{w}_j + s_2 \overline{w},$$

The weights s_1 and s_2 ($s_1 + s_2 = 1$) are proportional to the inverse of the respective mean squared errors of the estimators. This leads to a 'shrinkage' estimator for the level 2 group mean effects defined as:

$$\hat{u}_j = \overline{w}_{j,s} = \frac{n_j \hat{\sigma}_u^2}{\left(n_j \hat{\sigma}_u^2 + \hat{\sigma}_e^2\right)} \overline{w}_j \qquad \text{(refer to appendix II)}$$

As the number of level one units within a particular level two unit (n_j) decreases the 'shrinkage factor' increases reflecting the lack of information in the group so that the estimate becomes shrunk towards the mean estimate for the total sample (i.e. the overall population value). This is a method proposed originally by Stein[33] and expanded upon by Efron and Morris[34] and is common in the production of, for example, cancer maps.[35,36] The estimated effect for any one hospital (for example) is not solely based upon its own data, but also by taking account of the effects for other hospitals and the characteristics of the distribution they share. Higher level groups containing relatively small numbers of units can be included in the analysis and their effects adjusted according to the relative lack of information contained within the data. This contrasts with a fixed effects approach, where the relative lack of information is reflected through the associated estimated standard error. If we were to ignore the 'shrinkage factor' then the higher level estimates would be based on the mean estimate $\overline{w}_j$ alone, with little indication of the precision of a particular estimate. The approach is consistent with the assumption that higher level units belong to a population of units and that inferences are to be made about populations.

RANDOM COEFFICIENTS

It is concievable that the relationship between an explanatory variable and the response is not the same across all hospitals. Certain hospitals may have the effect of increasing the average response (for example, length of stay) of younger patients (compared to a national average) whilst

decreasing the stays of older patients. Different hospitals may treat what appear to be similar patients in a different manner, and that this difference may be related to some characteristic of the patient such that a uniform relationship across all hospitals (as assumed by (1)) is not a correct specification.

The exploration of different 'higher level effects' can be obtained by the inclusion of random coefficients (see also, Hsiao[27], chp 6 on variable coefficient models) such that the slope effect associated with an explanatory variable (x_{ij}) can be represented by:

$$y_{ij} = \beta' \mathbf{x}_{ij} + \gamma_j x_{ij} + u_j + e_{ij} \qquad (3)$$

In (3) we have three random terms, two of which are random at the hospital level, $\gamma_j x_{ij}$ and u_j. Each of these has a variance and generally there will be a non-zero covariance between them, such that:

$$\text{var}(u_j) = \sigma_u^2$$
$$\text{var}(\gamma_j) = \sigma_\gamma^2$$
$$\text{cov}(u_j, \gamma_j) = \sigma_{u\gamma}$$

leading to the conditional variance:

$$\text{var}(y_{ij} | x_{ij}) = \sigma_\gamma^2 x_{ij}^2 + \sigma_u^2 + 2\sigma_{u\gamma} x_{ij} + \sigma_e^2$$

and covariance:

$$\text{cov}(y_{ij}, y_{kj} | x_{ij}, x_{kj}) = \sigma_\gamma^2 x_{ij} x_{kj} + \sigma_u^2 + \sigma_{u\gamma}(x_{ij} - x_{kj})$$

which forms the block diagonal covariance matrix V.

The above highlights the use of random coefficients by allowing regression coefficients to vary across level 2 units. However, we can introduce more complex variance structures at any level of the hierarchy including level 1 and this may lead to substantively interesting interpretations and better model specification. A simple example of a complex level 1 variation is to make it a linear function of an explanatory variable. For example, suppose we wish to model individual alcohol consumption and we have data from a household survey consisting of a natural hierarchy of individuals within households within perhaps areas of residence. It seems sensible to assume that not only is the average alcohol consumption of males greater than females, but also that the variation in consumption is greater for males than females. We can accommodate and test this assumption by expanding the level 1 variance terms in the following manner. Suppose d_{ij} is a dummy variable representing males, then we can include an extra level 1 residual term associated with this variable, such that:

$$y_{ij} = \beta' x_{ij} + (u_j + e_{oij} + e_{1ij} d_{ij})$$

where

$$\text{var}(e_{0ij}) = \sigma^2_{e0}$$
$$\text{var}(e_{1ij}) = \sigma^2_{e1}$$
$$\text{cov}(e_{oij}, e_{1ij}) = 0$$

The contribution to the overall variance at level 1 is now a linear function of d_{ij}, leading to:
$\sigma^2_e = \sigma^2_{e0} + \sigma^2_{e1} d_{ij}$ with the necessary condition that $\sigma^2_e \geq 0$.

Specifying the level 1 variance as defined above is simply a device for obtaining the correct form of the variance function that we require. Notice that the same result could be obtained if d_{ij} was a dummy variable representing females. However, in this case the estimated variance attached to d_{ij} would be negative, although the overall variance at level 1 would still be positive. For presentational purposes this appears somewhat missleading and in the absence of prior knowledge of the likely magnitude of the variances a more flexible specification is given by constraining the variance term to be zero, whilst allowing the covariance to be estimated:

$$\text{var}(e_{0ij}) = \sigma^2_{e0}$$
$$\text{var}(e_{1ij}) = 0$$
$$\text{cov}(e_{oij}, e_{1ij}) = \sigma_{e01}$$
with
$\sigma^2_e = \sigma^2_{e0} + 2\sigma_{e01} d_{ij}$, and $\sigma^2_e \geq 0$

Constraining the variance parameter to be zero in the presence of a non-zero covariance is again a device used to obtain the required variance structure, and leads to the same estimated level 1 variance as before. This example illustrates a simple form of the level 1 variance and more elaborate variance functions including quadratic terms may be desirable and can be specified accordingly.

EXTENSIONS TO THE BASIC MODEL

The models discussed thus far represent the most basic form of a multilevel model where a continuous response is linearly related to a set of explanatory variables and the structure of the hierarchy is simple. In terms of the contributions to health and health economics research more complex multilevel models may have the most to offer. In its simplest form the hierarchy consists of two levels as specified in equation (1). However, extensions to further levels are straightforward; for example, we may be interested in the efficiency of both clinicians and provider units when assessing performance. In such a situation the hierarchy consists of patients within clinicians within provider units and a multilevel model containing three levels is required (see Figure 2):

$$y_{ijk} = \beta' \mathbf{x}_{ijk} + v_k + u_{jk} + e_{ijk} \qquad (4)$$

Interest here lies in the variance estimates σ_e^2, σ_u^2, and σ_v^2 together with fixed part parameters. The inclusion of random coefficients at any of the levels may also be of interest. Alternatively, data may consist of a series of repeated measurements on patients attending different hospitals. Again, this structure can be comfortably modelled using three levels: observations within patients within hospitals.

In the example given above of patients within clinicians within provider units its is unlikely that the hierarchy will be simple, and that, in reality, clinicians will operate in more than one provider unit. In such situations the hierachy is termed cross-classified. This occurs when individuals within a lower level cluster are grouped into a different higher level unit than peers from the same cluster (see Figure 3). Such model specifications, although being computationally more demanding,[37] can nevertheless be handled using appropriate software[38]. In Figure 3, the cross-classification occurs at level 2 and each separate cross-classification can be viewed as forming a distinct cell. The structure is then modelled using variance components (one for clinicians and one for hospitals) at level 2 and a single variance term at level 1. In notational form the variance components forming level 2 are represented by the subscripts j_1 and j_2, and classifications at the same level are grouped within parentheses. The general form of a cross-classified model can be written as follows:

$$y_{i(j_1 j_2)} = \beta' \mathbf{x}_{i(j_1 j_2)} + u_{j_1} + u_{j_2} + e_{i(j_1 j_2)}$$

The covariance structure at level 2 can be written in the following form:

$$\text{cov}(y_{i(j_1 j_2)} y_{i'(j_1 j_2')}) = \sigma^2_{u_1}$$
$$\text{cov}(y_{i(j_1 j_2)} y_{i'(j_1' j_2)}) = \sigma^2_{u_2}$$
$$\text{var}(y_{i(j_1 j_2)}) = \text{cov}(y_{i(j_1 j_2)} y_{i'(j_1 j_2)}) = \sigma^2_{u_1} + \sigma^2_{u_2}$$

In this formulation the level 2 variance is the sum of the two separate classification variances, and the covariance for two level 1 units (patients) in the same classification (cell) is equal to the variance for that classification. Therefore, two patients who share the same clinician but different hospitals have a covariance $\sigma^2_{u_1}$, and patients who share a common hospital but different clinicians have a covariance of $\sigma^2_{u_2}$. The covariance for two patients who do not share either classification is zero. The procedure allows estimates of the variances to be obtained and hence the outcomes across individuals can be 'partitioned' into that due to differences between individuals, clinicians and hospitals. In this way, the relative importance of each of the higher level effects may be assessed.

Much of health data do not conform to a simple model specification such as a continuous response and/or linear link function and further extensions of (1) to incorporate generalized linear models (refer to McCullagh and Nelder[39]) including link functions to specify logit, Poisson, negative binomial, duration (survival) and multinomial models may be specified[38].

RANDOM VERSUS FIXED EFFECTS

The literature on Panel data techniques places emphasis on the relative merits of treating higher level units as random or fixed effects (for example, see Judge et al.[23] and Hsiao[27]). For example, model (1) above may be rewritten in a form more familiar from the Panel literature:

$$E\left(y_{jt}|\mathbf{x}_{jt}\right)=\beta'\mathbf{x}_{jt} \qquad (5)$$

$$E\left(y_{jt}|\mathbf{x}_{jt},u_{j}\right)=\beta'\mathbf{x}_{jt}+\delta u_{j} \qquad (6)$$

where; $j=1,\ldots,N$, and $t=1,\ldots,T$. In model (5), the individual specific effects are specified as random effects and in (6) they are specified as a set of j-1 fixed effects to be estimated together with the set of explanatory variables $\mathbf{x}$. The choice of specification requires careful consideration and may be determined by the data generating process and/or, the type of inference sought. If individuals (and their estimated effects) are not of intrinsic importance in themselves, but moreover are assumed to be random draws from a population of such individuals and that inferences concerning population effects and their characteristics are sought then a random specification may be more suitable. However, if inferences are to be confined to the effects in the sample only, and that these effects are of substantive interest then the effects themselves are more appropriately considered fixed (see for example, Hsiao[27], p41).

Another important consideration in choosing between the two specifications is when an explanatory variable is correlated with the effects. In such circumstances, a random or fixed effects approach may lead to vastly different estimates, and again careful consideration of the model specification is warented (see for example, Hsiao[27] p41, and Hausman[40]). The difference in the two approaches can be viewed as one of whether the conditional distribution of u_j, given $\mathbf{x}$ is equal to the unconditional distribution of u_j. In the case where x is correlated with u the two approaches will yield different estimators since $E(u_j)$ will not be constant but equal to some function of x (i.e. $E(y_{jt}|\mathbf{x}_{jt},u_j)\neq E(y_{jt}|\mathbf{x}_{jt})$.

The situation can be extended to the multilevel model depicted in (1). When u_j and x_{ij} are correlated, and group sizes are relatively small, the iterative generalised least squares estimator for the parameters β will be inconsistent when the number of higher level groups becomes large. Treating the effects u_j as fixed and applying a dummy variable estimator leads to consistent estimates. However, when group sizes are large, the two estimators can be shown to be equivalent (see Blundell and Windmeijer[41]).

In the situation where an explanatory variable is correlated with the higher level effects, and the sole concern of the analyst is the consistent estimation of the parameters associated with the explanatory variables or the mean effect of the higher levels, a fixed effects specification has advantages. However, in the multilevel formulation intrinsic interest lies in the estimation and interpretation of higher level variances together with their possible modelling with random coefficients, after conditioning on the set of explanatory variables. Higher level effects are not viewed as nuisance parameters to be conditioned on, but are of central importance to the analysis. Where such correlations exist the random effects specification may be at the cost of inconsistent

parameter estimation, but has benefits both in terms of the substantive exploration of higher level effects and increased efficiency of estimation since fewer parameters are required (σ^2_w for random effects, j-1 set of dummies for fixed effects). The later is particularly important in situations where there are many higher level units, some with few level 1 observations which may occur, for example, in clinical trails data where collection is over many sites. In such cases, higher level effects may be poorly estimated and this, in part, is the motivation behind using a random specification with shrunken higher level estimates. Directly modelling the correlation structure (for example, when a higher level variance is a function of an explanatory variable) using random coefficients may be of particular interest to the analyst, but the analogous exploration of interaction terms between explanatory variables and fixed effects is likely to be prohibitive (due to loss of degrees of freedom and interpretability) in a fixed effects model. For intermediate situations where the trade off between consistency and inefficiency is unclear, application of the Hausman test[40] may shed light on the most appropriate specification if substantive considerations do not already dictate this decision.

EXAMPLE

To illustrate the use of multilevel models we use as an example an analysis of intertemporal preferences for future health effects by Cairns and van der Pol.[43] Respondents were asked to identify what future level of benefit would lead them to be indifferent between a specified benefit to be received one year in the future and the more distant delayed benefit. Each respondent was asked to provide estimates of their chosen future level of benefit for two different periods of delay. Information on individual's age, gender, long-term health, cigarette consumption, educational attainment, and whether they were parents of children under the age of ten were also obtained. For full details refer to Cairns and van der Pol.[43] From the sample data collected implied discount rates for each respondent were calculated and regressed against the set of explanatory variables. Table 1 presents the results of comparing an OLS specification and a multilevel specification of a hyperbolic discounting model. OLS standard errors have been adjusted using White's method.[44]

For a full discussion of the results and the implications drawn the reader is refered to the source[43]. However, a few points are worthy of note. First, it appears from the t ratios that although the OLS standard errors have been White corrected, they remain underestimated, and hence the significance of the coefficients are exaggerated. Secondly, the partitioning of the variance between that observed across responses within respondents and that across respondents themselves allows the intra-class correlation to be estimated. In the example here, clearly the vast majority of variation (98%) exists across individuals. This suggests that respondents vary greatly in their time preferences, but in comparison appear to be reasonably consistent in applying discount rates to different periods of delay. Thirdly, the use of multilevel analysis allows the explicit investigation of relationships operating at the level of the individual (hierarchy of responses within individuals). A priori, it is expected that individuals will vary greatly in terms of their time preferences and how heavily they discount future health benefits. The significance of the coefficients attached to the variance of delay and the covariance between delay and the constant supports this view. Further, the negative coefficient of the covariance indicates that, in general, the discount rate falls more rapidly with respect to increasing delay for individuals with relatively high short run discount rates compared to those with relatively low short run discount rates.

	OLS Analysis		Multilevel Analysis	
	Coefficient	t-value	Coefficient	t-value
Fixed effects:				
Constant	0.5224	13.05	0.4691	11.98
Age	-0.0047	2.84	-0.0027	1.63
Age squared	0.000043	2.57	0.000024	1.45
Smoker	0.0328	2.52	0.0251	2.03
Delay	-0.0047	4.88	-0.0038	7.91
Health	-0.0230	1.76	-0.0199	1.62
Education	-0.0210	1.86	-0.0156	1.41
Children	0.0146	1.39		
Random effects:				
Level 1:				
σ_e^2			0.0011	3.84
Level 2:				
σ_u^2 (variance of constant)			0.0482	13.48
σ_γ^2 (variance of Delay)			0.0001	9.26
$\sigma_{u\gamma}$ (covariance between constant and Delay)			-0.0018	11.43

Table 1: comparison of OLS and multilevel estimates for hyperbolic discounting model.

Model: $y_{ij} = \alpha + \beta x_{ij} + \gamma d_{ij} + (u_j + v_j d_{ij} + e_{ij})$

where y_{ij} is the response: implied discount factor for a specified period of delay,

x_{ij} set of explanatory variables conditioned upon.

d_{ij} variable representing period of delay (termed Delay).

u_j represents individual heterogeneity (variance in response across individuals after conditioning on x_{ij} is estimated).

$v_j d_{ij}$ represents a random coefficient specifying the level 2 variance is a function of the period of delay (estimated as a variance associated with v_j and a covariance between v_j and u_j).

e_{ij} is the level 1 error term (estimated as a variance)

A further advantage of applying a multilevel specification to these data is that the observed heterogeneity across individuals is modelled whilst preserving degrees of freedom. Due to the lack of multiple responses illicited from individuals, a fixed effects specification would be prohibitive in this application. The above example provides an illustration of the type of output and interpretation from a multilevel model. Below we outline some potential areas of application of these techniques to health economics.

APPLICATIONS TO HEALTH ECONOMICS

The potential areas of application of multilevel techniques to health economics research are many. Perhaps its most obvious application is in furthering our understanding of medical practice variations. It is well known that these occur at all levels from individual clinicians to total health care systems across countries (for a useful review of the literature see Phelps[45], Phelps and Mooney[46], or Andersen and Mooney[47]) and additional evidence of its existence are unsurprising. Of particular interest is the welfare loss associated with small area variations in treatment patterns.[48] However, the key to directing policy instruments towards producing greater health care gains through the elimination of variations is a fuller appreciation and understanding of the magnitudes, location, characteristics and causes of medical practice variations[47]. Methods of empirical analyses that can address such issues through a greater descriptive or explanatory examination of the differences in the way health care is practiced and administered can potentially offer further insight into the process of informing policy. The ability to partition unexplained variation between the different levels of the hierarchy and to explore the characteristics of the variation observed may provide important information concerning the origins and mechanism through which variations in medical practice occur, their economic consequences, and how to guide policy to reduce the associated welfare losses.

In studies of health it has been argued that there is a distinct geography of lifestyle behaviour[14,49,50] in that the contextual aspect of an individuals place of residence should not be ignored when exploring individual health behaviour and health outcomes. It is argued that variations in health behaviour are influenced by space and locality and that the interaction of these with individual level characteristics are crucial to a fuller understanding of health behaviour mechanism[14,49,50,51]. It has been proposed that this is particularly important in studies examining health related behaviours such as smoking, drinking, diet, sexuality and drug misuse as "these 'lifestyles' might be expected to result from both individual predisposition and geographically based cultural influence"[14]. Substantive interest lies in exploring the extent to which contextual effects of areas play a role in determining individual health behaviour once individual level characteristics (pre-disposition, socio-economic characteristics etc.) have been adequately controlled for. The partitioning of variation into individual and area effects together with the explicit modelling of this variation allows such insights to be made, and the use of multilevel models is becoming a favoured statistical tool in this area.

Such considerations have important implications for an equitable allocations of resources. Funding formulae based extensively upon small area deprivation indicies as proxies for ill health (for example, Jarman[52] and Townsend[53]) assume that the relationship between health needs (often measured as expressed demand) and deprivation is the same across all areas. This is unlikely to be the case, and area inequalities may remain after controlling for the compositional factors of populations. Adopting a multilevel approach Congdon[16] concluded this to be the case in an analysis of small areas (electoral wards) within administrative districts (53 local authorities in London and East Anglia). For both death and long term illness in middle age he found a significant area effect associated with deprivation (in particular, the rate of unemployment and housing tenure) after controlling for the general relationship between deprivation and morbidity observed across all areas. Similar area effects have been observed in an analysis relating general practitioner consultation rates and socio-economic characteristics of patients[54]. Such findings highlight the potential need for some local discretion when allocating resources. The recent resource allocation formula in England adopted a multilevel approach[55] to examine area variations in hospital utilisation.

The introduction of an internal market within the NHS has forced the health sector to become more concerned with issues of performance which has meant that both cost and quality are relevant concerns of providers when meeting purchasers needs. Performance or effectiveness have been assessed in different ways to suit different audiences and include performance indicators intended to allow the reader to draw comparisons between areas or institutions, either directly or through reference to a mean, and efficiency measures that are more concerned about the influences of economies of scale and scope in specialisation. Various quantitative techniques have been used to measure performance and efficiency ranging from very crude age-sex standardised rates (for example, Scottish Office Clinical Outcome Indicators[55]) to more sophisticated models grounded on economic theory such as efficient frontier estimation (see for example, Wagstaff[57]). The use of multilevel models has created interest and debate in the measurement of performance[3] and with a little thought and ingenuity may also prove useful in frontier estimation, particularly when there are competing levels at which inefficiency may occur.

The techniques are also likely to be of great value in panel studies of aggregate hospital level repeated observations. The repeated nature of panel data can be viewed in a multilevel framework with measurement occassions within hospitals forming the first level and hospitals or provider units the higher level groups leading to the familiar panel data variance components random effects model.[36]

Although well established econometric techniques (fixed and random effects models) may also be used to estimate performance emphasis in the multlevel model is placed on the estimation of the distribution of the population of higher level effects and their interpretation, which in itself may lead to greater substantive insight. Given a sufficiently long panel, changes in inefficiency over time could be explicitly investigated, thereby avoiding the somewhat restrictive assumption of constant inefficiency. The multilevel approach has advantages over some conventional repeated measures or panel data techniques; often conventional models require a balanced design - broadly, that the measurement occasions are the same for all higher level units of interest. These restrictions are severe and provided measurement occasions are missing at random do not apply in the multilevel specification.

The methods should also have useful applications in evaluation studies (both observational and RCTs) where data on cost and effects are collected over multiple sites. In such circumstances, even with randomisation, it can be expected that the site of treatment may have an impact on the outcome regardless of treatment the patient receives. This may result from various sources, and is likely to be due to differences in medical practice as administered by individual clinicians or clinical management and resources dictated by provider units. However, it may also be due to differences in subpopulations from which each site recruits and hence although patients may be randomised to treatments at each site, they may not be representative of the general population. Pooling data over sites without regard to such site-specific differences may lead to incorrect inference (for example, inefficient parameter estimates). The inclusion of site as a level in a multilevel analysis will ensure that the clustering effects within sites will be adequately controlled for.

DISCUSSION AND CONCLUSIONS

Multilevel models are applicable to data arranged in hierarchies consisting of multiple macro units with multiple micro units within each macro unit. The approach allows exploration of variation

arising at different levels of the hierarchy and modelling of the correlation structure inherent in such data sets and leads to efficient parameter estimates. Estimates of higher level effects (for example, hospital effects), adjusted for sample size, may be obtained through residual estimation and used as measures of efficiency. Hierarchies formed by the clustering of individuals into well defined groups can be seen to extend beyond provider units or geographical localities to include repeated measures. Extensions to accommodate generalised linear models allows a flexible approach to model specification and offers researchers more appropriate tools to address the types of data commonly encountered in health economics and more general health services research. The potential application of these methods to health economics are many and include the exploration of medical practice variations, examining inequalities in resource allocation, the analysis of economic evaluation data collected alongside clinical trials, and the estimation of performance measures. The careful application of these methods may provide researchers with a valuable tool placing emphasis on analysis of variation, and should compliment the use of more familiar techniques that are available in the well established panel data literature.

REFERENCES

1 Blalock, H. Contextual-effects models: theoretical and methodological issues. Annual Review of Sociology 1984; **10**: 353-72.

2 DiPrete, TA., Forristal, JD. Multilevel Models: Methods and Substance. *Annual Review of Sociology* 1994; **20**: 331-57.

3 Goldstein, H., Spiegelhalter, D.J. League tables and thier limitations: statistical issues in comparisons of institutional performance (with discussion). *Journal of the Royal Statistical Society A* 1996; **159**: 385-444.

4 Robinson, W.S. Ecologial correlations and the behaviour of individuals. *Am. Social. Rev.* 1950; **15**: 351-7.

5 Selvin, H. Durkheim's suicide and problems of empirical research. *American Journal of Sociology.* 1958; **63**: 601-19.

6 Goldstein, H. *Multilevel Statistical Models*. London: Edward Arnold, 1995.

7 Longford, N.T. *Random coefficient models*. Oxford: Clarendon Press, 1993.

8 Bryk, A.S., Raudenbush, S.W. *Hierarchical linear models*. Newbury Park: Sage, 1992.

9 Kreft, I.G., de Leeuw, J., van der Leeden, R. Comparing five different statistical packages for hierarchical linear regression: BMDDP-5V, GENMOD, HLM, ML3 and VARCL. *American tatistician* 1994; **48**: 324-335.

10 A special issue of the Journal of Educational and Behavioral Statistics has been devoted to Multilevel Models. Kreft, I.G.G (ed.) *Journal of Educational and Behavioral Statistics* 1995; **20**: No 2.

11 Herbert, D. Neighbourhood incivilities, crime, place and multilevel models. *Area* 1994; **26**: 81-84.

12 Jones, K. *Multilevel Models for Geographical Research.* Geobooks, Norwich, 1991.

13 Duncan, C., Jones, K., Moon, G. Do Places Matter? A multilevel analysis of regional variations in health-related behaviour in Britain. *Soc. Sci. Med.* 1993; **37**: 725-733.

14 Von-Korff, M., Koepsell, T., Curry, S., Diehr, P. Multilevel analysis in epidemiologic research on health behaviours and outcomes. *American Journal of Epidemiology*. 1992; **132**: 1077-82.

15 Rice, N., Leyland, A.H. Multilevel Models: applications to health data. *Journal of Health Services Research and Policy* 1996; **1**: 154-64.

16 Congdon, P. The impact of area context on long term illness and premature mortality: An illustration of multilevel analysis. *Regional Studies* 1995; **29**: 327-344.

17 Phelps, C. E. Perspectives in Health Economics. *Health Economics* 1995; **4**: 335-353.

18 Aitkin, M., Longford, N. Statistical modelling issues in school effectiveness studies. *J. R. Statist. Soc. A*. 1986; **149**: 1-43.

19 Paterson, L., Goldstein, H. New statistical methods for analysing social structures; an introduction to multilevel models. *British Educational Research Journal* 1991; **17**: 387-93.

20 Goldstein, H., Rasbash, J., Yang, M., Woodhouse, G., Pan, H., Nuttal, D., Thomas, S. A multilevel analysis of school examination results. *Oxford review of Education*. 1993; **19**: 425-33.

21 Goldstein, H. Statistical information and the measurement of education outcomes (editorial). *Journal of the Royal Statistical Society, A* 1992; **155**: 313-15.

22 Wolfson, M.C. Social Proprioception: Measurement, Data, and Information from a Population Health Perspective. In Evans, R.G. Borer, M.L. Marmor, T.R. (eds.) *Why are some people healthy and others not? The determinants of Health of Populations*. New York, Aldine De Gruyter.

23 Judge, GG. Griffiths, WE. Hill, RC. Lee, TC. *The theory and practice of econometrics.* New York: Wiley, 1980

24 Hausman, JA. The econometrics of nonlinear budget sets. *Econometrica* 1985; **53**: 1255-82.

25 Moffitt, R. The econometrics of kinked budget constraints. *Journal of Economic Perspectives* 1990; **4**: 119-39.

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

27 Hsiao, C. *Analysis of Panel Data*. Cambridge: Cambridge University Press, 1986.

28 Goldstein, H. Restricted unbiased iterative generalized least-squares estimation. *Biometrika* 1989; **76**: 622-3.

29 Goldstein, H., Rasbash, J. Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalised least squares. *Computational Statistics and Data Analysis* 1992; **13**: 63-71.

30 Longford, NT. *VARCL - software for variance components analysis of data with hierarchically nested random effects (maximum likelihood).* Princeton, NS; Educational Testing Service 1988.

31 Bryk, AS. Raudenbush, SW. Seltzer, M. Cogdon, R. *An introduction to HLM: Computer program and user's guide (2nd Edn.).* Chicago: University of Chicago; Dept. of Education.

32 Liang, K. Zeger, SL. longitudinal data analysis using generalised linear models. *Biometrika* 1986; **73**: 45-51.

33 Stein, C. *Inadmissibility of the usual estimator for the mean of a multivariate normal population.* Procedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol 1, Berkeley: University of California Press, 1955: 197-206.
34 Efron, B., Morris, C. Stein's paradox in statistics. *Scientific American* 1977; **236**: 119-27.
35 Clayton, D., Kaldor, J. Empirical Bayes estimates of age-standardised relative risks for use in disease mapping. *Biometrics* 1987; **43**:671-681.
36 Langford, I.H. A multilevel log-linear model of childhood leukemia mortality. *Health & Place* 1995; **1**: 113-119.
37 Rasbash, J., Goldstein, H. Efficient analysis of mixed hierarchical and cross-classified random structures using a multilevel model. *Journal of Educational Statistics* 1994; **19**: 337-50.
38 Rasbash, J. Woodhouse, G. *MLn Command Reference* 1995. London: Institute of Education.
39 McCullagh, P. Nelder, J. *Generalized linear models (2nd edition).* London, Chapman and Hall 1989.
40 Hausman, JA. Specification tests in econometrics, *Econometrica* 1978; **46**: 1251-71.
41 Blundell, R. Windmeijer, F. Correlated cluster effects and simultaneity in multilevel models. *Mimeo: London, Institute for Fiscal Studies.*
42 Baltagi, BH. *Econometric analysis of panel data.* New York: Wiley, 1995.
43 Cairns, JA. van der Pol, MM. *Saving future lives: a comparison of three discounting models.* Presented to Fifth European Workshop on Econometrics and Health Economics, Barcelona, 1996.
44 White, H. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica*, 1980; **48**: 817-38.
45 Phelps, CE. Diffusion of information in medical care. *Journal of Economic Perspectives* 1992; **6**: 23-42.
46 Phelps, CE. Mooney, C. *Variations in medical practice use, causes and consequences*, in: RJ Arnauld, RF Rich and WD White, eds., Competitive approaches to health care reform (Washington, DC: The Urban Institute Press) 139-78.
47 Andersen, T.F., Mooney, G. *The Challenges of Medical Practice Variations.* London: MacMillan Press 1990.
48 Phelps, CE. Welfare loss from variations: further considerations. *Journal of Health Economics* 1995; **14**: 253-60.
49 Duncan, C., Jones, K., Moon, G. Health -related behaviour in context: a multilevel modelling approach. *Social Science and Medicine*, to be published.
50 Jones, K., Moon, G., *Health, Disease and Society*. London: RKP, 1987.
51 Blaxter, M. *Health and Lifestyles*. London: Tavistock/Routledge, 1990.
52 Jarman, B. Identification of underprivileged areas. *British Medical Journal* 1983; **286**: 705-9.
53 Townsend, P., Beattie, A. Phillimore, P. *Health and deprivation: inequality and the north.* London: Croom Helm, 1988.
54 Carr-Hill, R.A., Rice, N., Roland, M. Socio-economic determinants of rates of General Practice consultations. *British Medical Journal* 1996; **312**: 1008-13.
55 Carr-Hill, R.A., Sheldon, T.A., Smith, P., Martin, S., Peacock, S., Hardman,G. Allocating resources to health authorities: development of methods for small area analysis of inpatient services. *BMJ* 1994; **309**: 1046-1049.
56 Scottish Office. *Clinical outcome indicators - 1993.* Edinburgh: Clinical Resource and Audit Group, 1994.
57 Wagstaff, A. Estimating efficiency in the hospital sector: A comparison of three statistical cost frontier models. *Applied Economics* 1989; **21**: 659-672.

Figure 1: Two level model

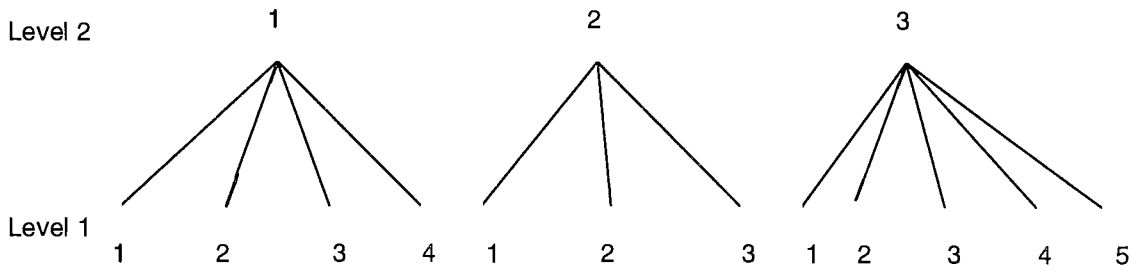


Figure 2: Three level model

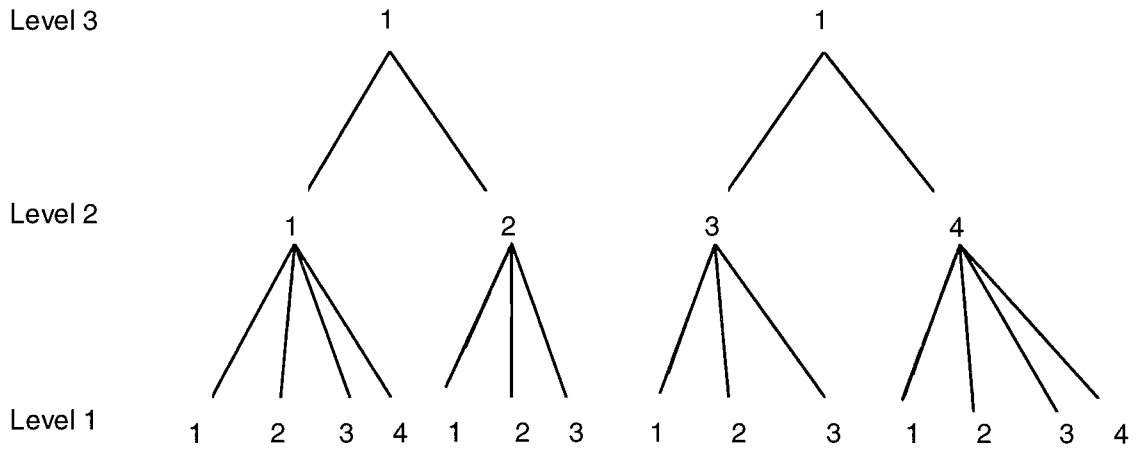
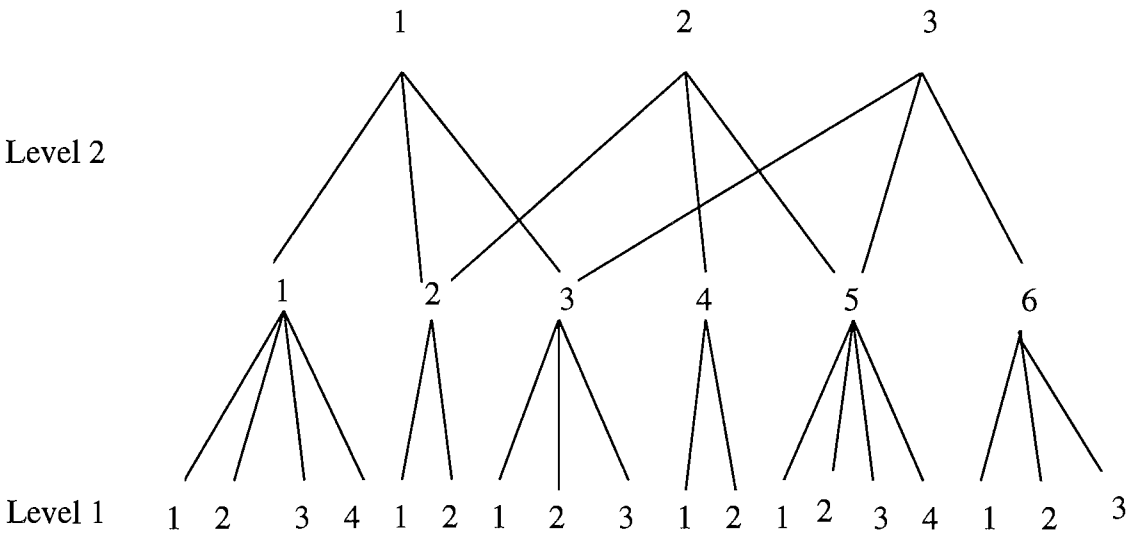


Figure 3: Three level cross-classified model



Appendix I: Covariance structure

Block diagonal covariance structure:

$$\mathrm{cov}(y_{ij}) = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & A_M \end{pmatrix}$$

$$\text{where } A_j = \begin{matrix} \sigma_e^2+\sigma_u^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ \sigma_u^2 & \ddots & & \vdots \\ \vdots & & \ddots & \sigma_u^2 \\ \sigma_u^2 & \cdots & \sigma_u^2 & \sigma_e^2+\sigma_u^2 \end{matrix} \quad , \; j = 1,\cdots,M$$

Appendix II: Shrinkage Estimators

i) MSE of $\overline{w}_j$

$$= E\left[\overline{w}_j - \mu_j\right]^2$$
$$= E\left[\overline{w}_j - E(\overline{w}_j)\right]^2 + \left[E(\overline{w}_j) - \mu_j\right]^2$$
$$= E\left[\overline{w}_j - E(\overline{w}_j)\right]^2 + [0]^2$$
$$= \sigma_w^2 / n_j$$

ii) MSE of $\overline{w}$

$$= E\left[\overline{w} - \mu_j\right]^2$$
$$= E\left[\overline{w} - E(\overline{w})\right]^2 + \left[E(\overline{w}) - \mu_j\right]^2$$
$$= E\left[\overline{w} - \mu\right]^2 + E\left[\mu - \mu_j\right]^2$$
$$= E\left[\overline{w} - \mu\right]^2 + E\left[\overline{w}_j - \mu_j\right]^2 + E\left[\overline{w}_j - \mu\right]^2$$
$$= \sigma_B^2 \sum_j n_j^2 \Big/ N^2 + \sigma_w^2 / N + \sigma_B^2$$

ignoring negative powers of N, we have: $MSE(\overline{w}) \approx \sigma_B^2$, an acceptable assumption for large sample sizes.

We have:

$$s_1 = n_j / \sigma_w^2$$
$$s_2 = 1 / \sigma_B^2$$

and

$$s_1 + s_2 = 1$$
$$\Rightarrow \frac{\sigma_w^2 + \sigma_B^2 n_j}{\sigma_w^2 \sigma_B^2} = 1$$
$$\Rightarrow s_1 = \frac{\sigma_B^2 n_j}{\sigma_w^2 + \sigma_B^2 n_j}, s_2 = \frac{\sigma_w^2}{\sigma_w^2 + \sigma_B^2 n_j}$$

Therefore:

$$\overline{w}_{j,s} = \frac{\sigma_B^2 n_j}{\sigma_w^2 + \sigma_B^2 n_j} \overline{w}_j + \frac{\sigma_w^2}{\sigma_w^2 + \sigma_B^2 n_j} \overline{w}$$

The unbiased but inefficient estimator $\overline{w}_j$ is 'shrunk' towards the biased but efficient estimator $\overline{w}$. The result is a biased estimator with a mean squared error smaller than that of $\overline{w}_j$; that is:

$$MSE(\overline{w}_{j,s}) = E\left[\overline{w}_{j,s} - \mu_j\right]^2 \approx \frac{1}{\sigma_B^{-2} + n_j \sigma_w^{-2}}$$ (Longford[7], p17 (1995))

and so is smaller than both σ_w^2 / n_j and σ_B^2, and approches σ_w^2 / n_j as σ_B^2 becomes very large and vica versa.

In the multilevel specification $e_{ij} \approx N(0, \sigma_e^2)$ and $u_j \approx N(0, \sigma_u^2)$, and hence:

$$\hat{u}_j = \frac{n_j \hat{\sigma}_u^2}{(n_j \hat{\sigma}_u^2 + \hat{\sigma}_e^2)} \overline{w}_j$$

so $\overline{w}_j$ is shrunk towards zero when the within unit sample size n_j is small and when the level 2 variance, as a proportion of the total variance, is small.

Appendix III: Information about software

The following software packages are available for multilevel analyses:

MLn
Contact address: The Multilevel Models Project, 11 Woburn Square, London WC1A 0SN.
E-mail: temsmya@ioe.ac.uk
Web site: http://www.ioe.ac.uk/hgoldstn/home.html
North American mirror site: http://www.medent.umontreal.ca/multilevel/

***BUGS** (Bayesian inference Using Gibbs Sampling)*
Contact address: MRC Biostatistics Unit, Institute of Public Health, Robinson Way, Cambridge CB2 2SR. E-mail: david.spiegelhalter@mrc-bsu.cam.ac.uk
Web site: http://www.mrc-bsu.cam.ac.uk/project/bugs

HLM
Contact address: Iec ProGAMMA, P.O. Box 841, 9700 AV Groningen, The Netherlands.
E-mail: gamma.post@gamma.rug.nl
or
Scientific Software Inc., 1525 East 53rd St., Suite 906, Chicago, Ill. 60615, U.S.A.

MIXOR and MIXREG
Contact address: Ann Hohmann, Ph.D., M.P.H., NIMH Services Research Branch, 5800 Fishers Lane, Room 10C-06, Rockville MD 20857, U.S.A.
or
Don Hedeker, Div. of Epidemiology & Biostatistics (mc 922), School of Public Health, University of Illinois at Chicago, 2121 West Taylor, Room 510, Chicago, IL 60612-7260, U.S.A. E-mail: hedeker@uic.edu
These programs are available by transfer from the web site
http://www.ioe.ac.uk/hgoldstn/mixreg.html

VARCL
Contact address: Iec ProGAMMA, P.O. Box 841, 9700 AV Groningen, The Netherlands.
E-mail: gamma.post@gamma.rug.nl

There is a discussion list devoted to multilevel models; to join this send the message
join multilevel firstname(s) lastname
to mailbase@mailbase.ac.uk