POLICY: SOCIOLOGICAL METHODS & RESEARCH is a quarterly journal devoted to sociology as a cumulative empirical science. The objectives of **SMR** are multiple, but emphasis is placed on articles that advance the understanding of the field through systematic presentations that clarify methodological problems and assist in ordering the known facts in an area. Review articles will be published, particularly those that emphasize a critical analysis of the status of the arts, but original presentations that are broadly based and provide new research will also be published. Intrinsically, **SMR** is viewed as a substantive journal but one that is highly focused on the assessment of the scientific status of sociology. The scope is broad and flexible, and authors are invited to correspond with the editors about the appropriateness of their articles.

MANUSCRIPTS should be submitted in quadruplicate to the attention of J. Scott Long, **SOCIOLOGICAL METHODS & RESEARCH**, Department of Sociology, Ballantine 744, Indiana University, Bloomington, IN 47405; BITNET to JSLONG@IUBACS. Simultaneous submission to another publication is acceptable only if **SMR** is given first refusal rights if the paper is accepted for publication. Articles should be typewritten, double-spaced, with footnotes, references, tables, and charts on separate pages and should follow the current journal style (which is based upon that followed by the Statistical, Psychological and Sociological Associations). An abstract of no more than 175 words, as well as a brief biographical paragraph describing each author's current affiliation, research interests, and recent publications, should accompany the manuscript. If the paper requires use of software written by the author, two copies of the software must be submitted with the manuscript. Enclose a stamped self-addressed postcard for acknowledgment of manuscript receipt. If the author's BITNET address is included with the submission, receipt will be acknowledged via BITNET. **Manuscripts will not be returned to the author;** therefore, please do not send your original copy.

SOCIOLOGICAL METHODS & RESEARCH (ISSN 0049-1241) is published four times annually—in August, November, February, and May—by Sage Publications, Inc., 2455 Teller Road, Thousand Oaks, CA 91320. Telephone: (805) 499-0721; fax/order line: (805) 499-0871.

Subscriptions: Regular institutional rate $159.00 per year; single issue, $41.00. Individuals may subscribe at a one-year rate of $57.00; single issue, $18.00. Add $6.00 for subscriptions outside the United States. Orders from the U.K., Europe, the Middle East, and Africa should be sent to the London address (below). Orders from India should be sent to the New Delhi address (below). Noninstitutional orders must be paid by personal check, VISA or MarsterCard.

Second class postage paid at Thousand Oaks, California.

This journal is abstracted or indexed in **Family Resources Database, Sociological Abstracts, Human Resources Abstracts, Social Sciences Citation Index, Social Research and Abstracts, Communication Abstracts, Sage Public Administration Abstracts, Current Contents, Automatic Subject Citation Alert, Sociology of Education Abstracts, Health Instrument File, SMR Database,** and **Social Planning/Policy and Development Abstracts,** and is available on microfilm from University Microfilms, Ann Arbor, Michigan.

Back Issues: Information about availability and prices of back issues may be obtained from the publisher's order department (address below). Single-issue orders for 5 or more copies will receive a special adoption discount. Contact the order department for details. Write to London office for sterling prices.

Inquiries: Address all correspondence and permissions requests to SAGE PUBLICATIONS, Inc., 2455 Teller Road, Newbury Park, California 91320. Inquiries and subscriptions from the U.K., Europe, the Middle East, and Africa should be sent to SAGE PUBLICATIONS Ltd, 6 Bonhill Street, London EC2A 4PU, United Kingdom. From India, write to SAGE PUBLICATIONS INDIA Pvt. Ltd, P.O. Box 4215, New Delhi 110 048 India. Other orders should be sent to the Thousand Oaks office.

Advertising: Current rates and specifications may be obtained by writing to the Advertising Manager at the Newbury Park office (address above).

Claims: Claims for undelivered copies must be made no later than twelve months following month of publication. The publisher will supply missing copies when losses have been sustained in transit and when the reserve stock will permit.

Change of Address: Six weeks' advance notice must be given when notifying of change of address. Please send old address label along with the new address to insure proper identification. Please specify name of journal. POSTMASTER: Send address changes to: Sociological Methods & Research, c/o 2455 Teller Road, Thousand Oaks, CA 91320.

PRINTED ON ACID-FREE PAPER

This special issue of SMR *is about the analysis of data collected at different levels of observation, such as groups and individuals within these groups, and about the methodological problems that are present when natural experimentation and observations nested within existing social groups are the object of study. The methodological problems are summarized in the term* multilevel problems. *A multilevel problem is a problem that inquires into the relationships between a set of variables that are measured at a number of different levels of a hierarchy. This article discusses some traditional approaches to the analysis of multilevel data and their statistical shortcomings. The random coefficient linear model is presented, which resolves many of these problems, and the currently available software is discussed. Next, some more general developments in multilevel modeling are discussed. The authors end with an overview of this special issue.*

Multilevel Analysis Methods

JOOP J. HOX
University of Amsterdam

ITA G. G. KREFT
California State University, Los Angeles

The relationship between individual and society has been a prominent problem in sociology ever since Durkheim's research into suicide (Durkheim 1898/1951). The general concept is that the behavior of individuals is influenced by the social contexts to which they belong and that the properties of a social group are influenced by the individuals who make up that group. This general conception has led to much empirical research into the interaction between variables that describe properties of individuals and variables that describe properties of social groups (cf. Boyd and Iversen 1979; Roberts and Burstein 1980; Van den Eeden and Huettner 1982; Blalock 1984). It has also led to a major philosophical debate as to whether contexts have any autonomous existence or whether their properties can be fully reduced to the properties of the individuals defining the context. This debate has not led to a firm conclusion, and we will not

---

AUTHORS' NOTE: *The names are in alphabetical order. We thank Edith de Leeuw, Scott Long, and Pieter van den Eeden for their helpful comments and Scott Long for his invaluable assistance.*

SOCIOLOGICAL METHODS & RESEARCH, Vol. 22, No. 3, February 1994 283-299

go into it here (for a discussion in terms of contextual analysis, see Hauser 1970). Instead, we will focus on the methodological problems that arise when we want to model relationships between individual and contextual variables, and we will discuss some analysis techniques that have come into use only recently.

It is important to note that individuals and the social contexts in which they live can be viewed as a hierarchical system of individuals and groups, with individuals and groups defined at separate levels of this hierarchical system. Variables can be defined at all distinct levels of this hierarchical system. Lazarsfeld and Menzel (1961) give a typology to describe different types of variables, defined at different levels. The lowest level is usually formed by individuals, although test items or repeated measurements in longitudinal designs can function as first-level observations too. Galtung (1969), for instance, defines roles as the lowest level nested within individuals.

When variables from different levels are analyzed at one single, common level, it becomes an important problem to identify the proper level to which all variables must be aggregated or disaggregated for the statistical analysis. Also, if the researcher is not very careful in the interpretation of the results, a danger exists of committing the fallacy of analyzing the data at one level and making inferences to another (higher or lower) level. This fallacy is known as the "ecological" fallacy or the Robinson effect, after Robinson (1950). Robinson describes the mathematical relation between ecological (i.e., aggregate) and individual correlations and concludes that both are almost certainly not equal, either in magnitude or in sign. Alker (1969) presents a broader typology of possible fallacies. These fallacies also affect more complicated models (Hannan 1971; Burstein 1978).

The discussion about inferential fallacies, which focused on conditions and correcting procedures to permit inference from one level to another, has given way after 1980 to a discussion about the appropriate procedures to investigate cross-level hypotheses, or multilevel problems (cf. Iversen 1991). Often problems of a multilevel nature will concern samples of individuals within existing real-life settings, for example students or parents nested in schools (Muthén and Goldstein in this issue), employees nested in industries (Kreft and de Leeuw in this issue), or respondents nested in interviewers (Hox in this issue). Other examples are longitudinal data, which have a nested structure of repeated measurements within subjects (Bryk and Raudenbush 1987), complex surveys with respondents nested within sampling units (Goldstein and Silver 1989), and data from factorial surveys (Hox, Kreft, and Hermkens 1991).

We start with an overview of the special problems encountered when analyzing grouped data, followed by an introduction of the hierarchical linear model or random coefficient (RC) model and the available software. For more detailed descriptions of the random coefficient model within the context of specific applications, we refer to the articles in this special issue by Hox, Kreft and de Leeuw, Snijders and Bosker, and Goldstein. Next we discuss some recent developments, one of which (multilevel covariance structures) is presented in more detail in the articles by Muthén and McDonald.

OVERVIEW OF MULTILEVEL PROBLEMS IN DATA ANALYSIS

The main problems that prohibit the application of traditional single-level statistical models to multilevel data can be summarized with the keywords *dependency*, *random effects*, *hierarchical nesting*, and *cross-level interactions*.

Dependency. Observations within a group that are close in time and/or space are expected to be more similar than observations in different groups. The amount of covariation between the scores of observations sharing the same context can be expressed by the intraclass correlation (Hays 1973). When ordinary significance tests are used, treating the individual as the unit of analysis, the important assumption of independence of residual error terms is violated. Even small values of the intraclass correlation have been shown to lead to Type I errors that are much larger than the nominal alpha level. For example, Barcikowski (1981) shows that an intraclass correlation of 0.01 in four groups of 25 observations results in a Type I error rate of 0.17 instead of the nominal 0.05 alpha level.

Random effects. An effect in ANOVA is said to be fixed when inferences are to be made only about the treatments actually included. An effect is random when the treatment groups are sampled from a population of treatment groups and inferences are to be made to the population of which these treatments are a sample. Random effects need random effects ANOVA models (Hays 1973). Multilevel models assume a hierarchically structured population, with random sampling of both groups and individuals within groups. Consequently, multilevel analysis models must incorporate random effects.

Hierarchical nesting. In ANOVA, when treatments or groups occur only within levels of another factor, the first factor is said to be nested within the second. The same nesting occurs with hierarchical data, with individuals nested within groups. Nested designs call for analysis models that handle two sources of variation: between group variation and systematic variation between individuals within groups. ANOVA models for nested and repeated measures designs can be used to analyze multilevel data, but complications arise when the group sizes are unequal or the time periods between repeated measures vary. Multilevel models handle these complications in a natural way.

Cross-level interactions. Cross-level interactions are interactions between explanatory variables defined at different levels of the hierarchy. An important subclass of the multilevel problem is the question of how several individual and group variables jointly influence one individual outcome variable. Traditionally, such questions were answered by executing a single analysis of variance or multiple regression type of model with one dependent variable at the lowest (individual) level and explanatory variables at all levels of the hierarchy. The goal of the analysis is to determine the direct effect of individual- and group-level explanatory variables, as well as to determine if characteristics of the context are moderating individual-level relationships (cf. Cronbach and Webb 1975). Analyzing cross-level interactions requires that variables defined at different levels of the hierarchy are combined in a single statistical model. As we will show, disaggregating all higher level variables and performing a single-level analysis implies unacceptable simplifications, leading to inefficient parameter estimates and downwardly biased precision estimates. Multilevel models are designed to solve these problems.

PAST APPROACHES TO MULTILEVEL ANALYSIS

Assume that we have a dependent variable Y_{ij} at the lowest level, with $j = 1 \dots K$ higher level units or groups, and $i = 1 \dots n_j$ lower level units or individuals in each group. We have P ($p = 1 \dots P$) explanatory variables X_{pij} at the individual level, and Q ($q = 1 \dots Q$) explanatory variables Z_{qj} at the group level.

In the past, two approaches have been popular to analyze this type of hierarchically nested data using traditional multiple regression models. The first approach has been to disaggregate all higher level explanatory variables Z_{qj} to the lowest level, followed by an ordinary least squares regression analysis. The regression equation, predicting Y_{ij} by the X_{pij}s and Z_{qj}s is given by:

$$\underline{Y}_{ij} = \beta_0 + \beta_1 X_{1ij} + \dots + \beta_P X_{Pij} + \beta_{P+1} Z_{1j} + \dots + \beta_{P+Q} Z_{Qj} + \underline{\varepsilon}_{ij}. \quad (1)$$

In equation (1), β_0 is the regression intercept, and β_p ($p = 1 \dots P + Q$) are the $P + Q$ regression slopes. We use the convention adopted in Kreft, de Leeuw, and Kim (1990), where we underline random variables in model equations. In the ordinary multiple regression equation (1), there is only one stochastic error component ε_{ij}, which is assumed to be independent, normal, and homoscedastic. However, because the individuals are sampled within groups, all unmodeled group variation will enter the residual error term at the group level. Thus we may expect a nonzero covariance between the residual error terms of individuals making up a group, which violates an important assumption on which standard errors and significance tests in ordinary multiple regression and analysis of variance are based. The lack of independence results in inefficient estimates and a Type I error rate that is much higher than the nominal alpha level. Ordinary multiple regression also assumes that all regression coefficients are equal in all groups, but offers no procedure to test this assumption. This restriction does not apply if analysis of covariance (ANCOVA) is used, with the lower level explanatory variables X_{pij} as covariates. ANCOVA is a fixed-effects model that allows different intercepts and slopes in the separate groups. If there are many groups, this leads to a large number of estimated parameters. This is not very efficient nor parsimonious, and the assumption of fixed-slope effects is often not realistic. Also, ANCOVA does not model group-level explanatory variables.

A second way to analyze multilevel data is the "slopes-as-outcomes" approach, first discussed in Burstein, Linn, and Capell (1978). In this approach, separate fixed-effects regression models are fitted within each group, using the lower level explanatory variables X_{pij} as predictors. Next, the within-group coefficients β_{pj} ($p = 0 \ldots P$, where β_{0j} is the intercept) are, in turn, used as dependent variables to be predicted by the higher level explanatory variables Z_{qj}. As de Leeuw and Kreft (1986) point out, using ordinary multiple regression to estimate the regression coefficients in both steps is inconsistent because, in the first step, we view the regression coefficients β_{pj} as fixed coefficients to be estimated by within groups regressions, whereas in the second step, we view them as random variables to be estimated by a between-groups regression. If the usual model assumptions are true at the lowest level, they will not be true at the second; the error structure will generally be quite different from the error structure assumed by the linear model. As a result, significance tests based on the usual standard errors are badly misleading (de Leeuw and Kreft 1986; Bryk and Raudenbush 1992).

In sum: Ordinary least squares (OLS) models applied to hierarchical data produce unstable estimates of the parameters and downwardly biased estimates of their precision because the assumptions on which OLS is based are badly violated. The assumptions underlying ANCOVA are more realistic, but still too restricted. For hierarchical data, we need hierarchical models and specialized estimation methods.

RANDOM COEFFICIENT MODELS FOR HIERARCHICAL DATA

Statistically, multilevel models are an extension of variance component models (e.g., Searle, Casella, and McCulloch 1992), where the parameters are not used to model the expected values of the outcomes, but rather to model variances and covariances. Variance components in a one way classification are formulated as in (2):

$$\underline{Y}_{ij} = \gamma_0 + \underline{\delta}_j + \underline{\varepsilon}_{ij}, \tag{2}$$

where $\underline{Y}_{ij}$ is an observation of the *i*th person in the *j*th context, $\underline{\delta}_j$ is the effect of the *j*th context, and $\underline{\varepsilon}_{ij}$ is the individual error term. In this simple case, with no explanatory variables, the scores for $\underline{Y}_{ij}$ may be thought of as a sum consisting of a general mean γ_0, a random effect $\underline{\delta}_j$ associated with a particular context, and an individual error term $\underline{\varepsilon}_{ij}$. Thus the variance component model distinguishes individual error variation and error variation of contexts or groups.

The multilevel models discussed in this special issue are different from the one-factor random effects model (2), because the last one is too general. RC models are mixed models that allow fixed effects to be estimated next to the random effects. In equation (3) we introduce such a mixed model, where we have $P + 1$ explanatory variables X_{pij} ($p = 0 \ldots P$, the intercept β_0 is modeled by $X_0 = 1$) with random regression coefficients $\underline{\beta}_p$.

$$\underline{Y}_{ij} = \Sigma\, \underline{\beta}_{pj} X_{pij} + \underline{\varepsilon}_{ij}, \tag{3}$$

where the β are random coefficients with a fixed and a random component, defined as a function of higher level variables Z_{qj} ($q = 0 \ldots Q$, with $Z_0 = 1$):

$$\underline{\beta}_{pj} = \Sigma \gamma_{pq} Z_{qj} + \underline{\delta}_{pj}. \tag{4}$$

The usual assumptions are that the X_{ij} and Z_j are fixed, that the individual level errors ε_{ij} are independently and identically distributed with expectation 0 and variance σ_ε^2 within each group, and that the errors $\underline{\delta}_{pj}$ are independent from the $\underline{\varepsilon}_{ij}$ and have a multivariate normal distribution with expectation 0 and (co)variances $\omega_{pp'}$, collected in covariance matrix Ω. In this extension of the variance components model, the total variance, usually divided into a single within and a single between part, is divided into a single within part (the variance σ_ε^2 of $\underline{\varepsilon}_{ij}$), and a number of between parts (the variances ω_{pp} of the $\underline{\delta}_{pj}$), one for each coefficient. By substituting (4) into (3) we obtain

$$\underline{Y}_{ij} = \Sigma\, \Sigma\, \gamma_{pq} X_{pij} Z_{qj} + \Sigma\, \underline{\delta}_{pj} X_{pij} + \underline{\varepsilon}_{ij}, \tag{5}$$

where γ_{pq} is defined as all fixed components of the random coefficients of the model, including the intercept, and δ_{pj} is defined as all random components of the random coefficients in the model. Note that the researcher has the choice of setting the slope variation ($\underline{\delta}_{pj}$, $p > 0$) to zero. Doing so, the slope is no longer random over contexts, but fixed. If there are no explanatory variables, the RC model (5) simplifies to the variance component model (2). In this model, we can define the intraclass correlation coefficient as $\rho_i = \omega_{00}/(\omega_{00} + \sigma_e^2)$, where ω_{00} is the variance of the $\underline{\delta}_{0j}$ and σ_e^2 is the variance of the $\underline{\varepsilon}_{ij}$. In the general random coefficient model, we also have covariances ω_{pq} between the error terms $\underline{\delta}_{pj}$ and $\underline{\delta}_{qj}$, and the intraclass correlation is no longer simple to interpret.

The difference in definition of the regression coefficients in the fixed models compared to the random model is that the first are conceived as representing always the same (fixed) treatments, whereas the regression coefficients in the random case are conceived as a random sample of a population distributed as (γ_{00}, ω_{00}) for the intercept, and as (γ_{pp}, ω_{pp}) for the slopes. Each variance component ω_{pp} is a variance in its own right and is a component of the variance in $\underline{Y}$.

The purpose of RC models is to estimate the expected values γ_{pp} and variances ω_{pp}. Computational methods for the joint estimation of random and fixed effects by empirical Bayes' estimation methods can be found in the literature. For a review of the statistical theory, we refer to Mason, Wong, and Entwisle (1984), de Leeuw and Kreft (1986), Goldstein (1987), and Bryk and Raudenbush (1992). These investigators all work with basically the same model, which deals with the problem of analyzing nested data collected under nonexperimental conditions. For examples of multilevel modeling and the RC model, we refer to the contributions of Hox and Kreft and de Leeuw in this issue.

DESCRIPTION OF THE AVAILABLE SOFTWARE: GENMOD, HLM, ML3, AND VARCL

The model for the analysis of hierarchically nested data is known under several names. It is a linear model, hence the name hierarchical linear model (HLM, in Bryk and Raudenbush 1992). The effects estimated are not fixed, but random, reflecting the random sampling of the groups in the model; hence another name used for this type of model, random coefficient model (de Leeuw and Kreft 1986). The error variance of the model is decomposed in variance at the individual level and variances at higher levels of the hierarchy, which leads to another name for this model, variance component models (Longford 1990).

Currently, there are four specialized programs available to analyze random coefficient regression models: HLM (Bryk, Raudenbush, Seltzer, and Congdon 1988), GENMOD (Mason, Anderson, and Hayat 1988), ML3 (Prosser, Rasbash, and Goldstein 1991), and VARCL (Longford 1990).[1] There are some differences between the models that can be fitted by the programs, but the various approaches have much in common and generally lead to the same conclusion. Most programs can deal with up to three levels (HLM3 [which is in beta testing], ML3, VARCL3—VARCL9 can analyze random intercept models with 9 levels). All programs use maximum likelihood (ML) estimation to decompose the variance. Asymptotic standard errors are available for the estimated parameters, which can be used for hypothesis testing. All programs allow to fix the variance of a specific regression slope at zero, which assumes that this slope is fixed. In addition, a deviance can be calculated from the value of the likelihood function, which can be used to test the difference in fit between two nested models, in a manner analogous to the chi-square test between two nested covariance structure models. This likelihood ratio test is especially useful for testing differences in the random structure of the model that are the result of fixing certain regression slopes across groups.

Computing the ML estimates involves complex nonlinear expressions in the parameters, which must be solved by an iterative procedure. The major difference between the programs is in the choice of the criterion that is optimized and the choice of the algorithm that is used to optimize the criterion.

ALGORITHMS

The four programs use three types of algorithms: iterative generalized least squares (IGLS), scoring, and EM algorithm.

IGLS (Goldstein 1987) views the likelihood function as depending on two sets of parameters: the fixed regression coefficients γ_j and the variance components Ω_j and σ_j. If the variance components were known, the regression coefficients could be estimated easily by weighted least squares. If the regression coefficients were known, the variance components can be estimated fairly simply as well. IGLS alternates these two minimizations iteratively. IGLS is used by ML3.

The second algorithm, proposed for multilevel models by Longford (1987) and de Leeuw and Kreft (1986), is the classical method of (Fisher) scoring. This is the Newton-Raphson method, applied to the likelihood function with a convenient first-derivative approximation to the second derivatives. The scoring algorithm is used by VARCL.

The third algorithm is the EM algorithm of Dempster, Laird, and Rubin (1977). The EM algorithm is based on bounding the likelihood function with a more convenient minorization. In each iteration, the minorization is maximized, and each of these steps increases the likelihood function as well. Multilevel programs using the EM algorithm are GENMOD and HLM.

Theoretically, a comparison of the algorithms should reflect the fact that EM typically has (very) slow linear convergence, which is global, (i.e., which occurs from any starting point). Both iterative generalized least squares and scoring have fast linear convergence and, for models that fit very well, scoring will have almost quadratic convergence. Both methods have global convergence as well. The Newton-Raphson method is truly quadratic, but it may diverge from the starting point when this is poorly chosen. Only the EM algorithm used by GENMOD and HLM produces estimates that cannot have values outside permissible boundaries; ML3 (based on IGLS) and VARCL (scoring) adjust inadmissible estimates. From our experience so far (Kreft et al. 1990), it seems that in practice the differences between the programs are less clear than the theory suggests.

FULL AND RESTRICTED ML

It is well-known that full ML estimation (FML), which treats the fixed coefficients as known, often produces downwardly biased variance estimates (Harville 1977). Restricted ML (REML) adds bias correction terms to the ML estimates, which amounts to treating the fixed coefficients as quantities that have uncertainty built in, when the

random parameters are computed. For small data sets, the different estimation procedures may have different results. In theory, REML should produce less biased estimates, especially in small data sets. In practice, it appears that with small data sets, results may differ even among programs that use the same estimation method (Kreft et al. 1990). For literature on RIGLS (which is comparable to REML), we refer to Goldstein (1987). A practical distinction is that with FML, the overall deviance can be used to test differences between nested models that differ in either their fixed or their random components, whereas with REML, the deviance can only be used to test differences in the random components. GENMOD and HLM use REML, VARCL uses FML, and ML3 offers both.

RECENT DEVELOPMENTS

The random coefficient model is essentially a multilevel version of the familiar multiple regression model. The wide range of applications of the random coefficient model attests to its versatility. However, the model has some serious limitations. First, it is design oriented, assuming fixed regressors. In hierarchical data, the regressors are often sampled in the same way as the dependent variables, and thus this assumption is not appropriate. The second limitation is that it deals with a single dependent variable, predicted by several explanatory variables. In general, there is often more than one dependent variable. One could apply it to each dependent variable separately, but this is not completely satisfactory because it ignores the relations between them. Thus a multivariate extension is desired and, preferably, the capability to fit multilevel path models. A third limitation is that the available programs cannot handle missing data. A fourth limitation is the linearity of the existing techniques, and a fifth is the assumption of normality of the residuals.

All of these limitations restrict the applicability of the currently available techniques to sociological research data. For two of these limitations, multivariate extensions and nonnormal models, solutions are emerging that are sufficiently well developed to be applied in substantive research. The development of multilevel generalized linear models will be treated below. The development of multilevel structural equations models is discussed by Muthén and by McDonald in this issue, and we will mention them only briefly here.

A major restriction of the multilevel models developed so far is that it applies only to models with a normal error distribution. An appropriate model for nonnormal data is the generalized linear model (McCullagh and Nelder 1989). This class of models is defined by three components: a linear regression equation, a specific probability distribution for the dependent variable, and a link function that links the predicted values to the observed data. Generalized linear models (GLIM; Numerical Algorithms Group 1985) can be used to analyze, among others, logit and probit models, loglinear models, and models for counts. Multilevel generalized linear models have been described by Wong and Mason (1985), Longford (1988), Mislevy and Bock (1989) and Goldstein (1991). Longford's (1988) VARCL program implements several generalized linear models, including the logistic link function for dichotomous and binomial data, the logarithmic function for Poisson data, and the reciprocal link function for gamma distributed data. It uses a quasi-likelihood procedure, and the algorithm is based on an iteratively reweighted least squares procedure, which is carried out together with the standard iterative procedure for the variance components. Comparable analyses can be carried out with Goldstein's program, ML3. With VARCL, analysts simply specify one of the available error distributions, and the program will carry out an analysis using the canonical link function for that distribution. With ML3, analysts have to use a macro to calculate the nonlinear transformation function and control the iterations. This makes ML3 both more flexible and more complicated than VARCL. A full implementation of generalized multilevel models would give access to a wide variety of probability distributions and link functions, comparable to the GLIM package. Enterprising analysts could develop ML3 macros for many of these models; an impressive macro library is already available from the developers of ML3.

Another extension of the linear multilevel model is the extension to structural equation modeling as introduced by both Muthén and McDonald in this issue. Applications with existing software for this type of modeling with, for instance, LISREL (Jöreskog and Sörbom 1989), EQS (Bentler 1985), or LISCOMP (Muthén 1987) require serious ad hoc programming with unusual setups.

OVERVIEW OF THIS ISSUE

This special issue starts with an article by Hox, who applies the random coefficient regression model to the analysis of interviewer effects in survey research. He shows how a multilevel analysis solves several methodological problems that have worried researchers and discusses how the multilevel model relates to other models used in this field.

Kreft and de Leeuw analyze the gender gap in earned incomes. Their article focuses on the interpretation of model parameters in the random coefficient regression model. They show how useful a detailed inspection of predicted values (posterior means) can be in the interpretation of a set of models.

Snijders and Bosker discuss the problem of how to define "explained variance" in a multilevel model. They show that explained variance is more complex than in single-level analysis and cannot be used the same way as in traditional linear models. Still, certain analogies do exist, which are not only useful as an indication of the explanatory power of a model, but also can be used to diagnose certain types of model misspecification.

Goldstein's article presents a model that applies when the individuals do not fall into mutually exclusive groups, but are included in a two-way factorial nested design. His model includes models for other, more complicated designs. Goldstein shows how data on parents and schools located in different neighborhoods can be viewed as a factorial design and can be analyzed using existing software for multilevel analysis.

Muthén presents the principles of multilevel modeling applied to the analysis of covariance structures with latent variables. His example concentrates on a common factor model, but the methods are applicable to general path models with latent variables. He describes a consistent estimator that simplifies the application of existing software for this purpose.

McDonald discusses a general hierarchical model for path analysis with latent variables (based, like Muthén's, on general theory given earlier by Goldstein and McDonald). He describes the application of a program specifically developed to implement such a hierarchical path analysis and gives a constructed example.

All articles in this issue deal with the multilevel problem in their own way. Hox's and Kreft and de Leeuw's contributions are applications of the random coefficient regression model for hierarchically structured data. The Snijders and Bosker article deals with a specific problem related to this model, the interpretation of explained variance. Finally, Goldstein, Muthén and McDonald present a new type of multilevel model. Goldstein still stays within the framework of the RC linear model, whereas Muthén and McDonald show a different approach. Their models are not RC models. The main difference between the RC models presented by Kreft et al. and Hox and the models presented by Muthén and McDonald is that the latter fit fixed coefficients and random predictors, whereas the RC models fit random parameters and fixed regressors. Of course it would be possible to extend the RC model to include random regressors in an RC pathmodel, but at the time of this writing such models are not available.

NOTE

1. Program Information (PC-versions):

GENMOD is written by Hermalin and Anderson at the Population Studies Center, University of Michigan, from instructions provided by Wong and Mason. It is available from William M. Mason, Department of Sociology, UCLA, 405 Hilgard Avenue, Los Angeles, CA 90024. Two diskettes, manual $20.

HLM, Version 2.20 is written by Bryk, Raudenbush, and Congdon. The manual is written by the same authors, with Seltzer. The program is available from Scientific Software, Inc., 1525 East 53rd Street, Suite 906, Chicago, IL 60615. The program and manual together are $300.

ML3, Version 2.3, is software for two- or three-level analysis written by Rasbash. The manual is by Prosser, Rasbash, and Goldstein. The program is based on theoretical work by Goldstein. It is available from the Multilevel Models Project, Institute of Education, 20 Bedford Way, London WC1H 0AL, UK. The program and manual are $475 (regular version) or $570 (extended memory version).

VARCL was initiated by Aitkin and Longford and is written by Longford. The program is distributed by ProGamma, P.O.B. 841, Groningen, The Netherlands. The price for

manual, software, and code was at the time of this writing $250 for individuals and departments and $800 for institutions.

REFERENCES

Alker, H. R. 1969. "A typology of fallacies." Pp. 69-86 in *Quantitative Ecological Analysis in the Social Sciences*, edited by M. Dogan and S. Rokkan. Cambridge, MA: M.I.T. Press.

Barcikowski, Robert S. 1981. "Statistical Power with Group Mean as the Unit of Analysis." *Journal of Educational Statistics* 6:267-85.

Bentler, P. M. 1985. *Theory and Implementation of EQS: A Structural Equations Program*. Los Angeles: BMDP Statistical Software.

Blalock, Hubert M., Jr. 1984. "Contextual-Effects Models: Theoretical and Methodological Issues." *Annual Review of Sociology* 10:353-72.

Boyd, Lawrence H., Jr. and Gudmund R. Iversen. 1979. *Contextual Analysis: Concepts and Statistical Techniques*. Belmont, CA: Wadsworth.

Bryk, Anthony S. and Stephen W. Raudenbush. 1987. "Applying the Hierarchical Linear Model to Measurement of Change Problems." *Psychological Bulletin* 101:147-58.

———. 1992. *Hierarchical Linear Models*. Newbury Park, CA: Sage.

Bryk, Anthony S., Stephen W. Raudenbush, Michael Seltzer, and Richard T. Congdon. 1988. *An Introduction to HLM: Computer Program and Users' Guide*. Chicago: University of Chicago.

Burstein, Leigh. 1978. "Assessing Differences Between Grouped and Individual-Level Regression Coefficients." *Sociological Methods & Research* 7:5-28.

Burstein, Leigh, Robert L. Linn, and I. Capell. 1978. "Analyzing Multi-Level Data in the Presence of Heterogeneous Within-Class Regressions." *Journal of Educational Statistics* 4:347-89.

Cronbach, Lee J. and Noreen Webb. 1975. "Between-Class and Within-Class Effects in a Reported Aptitude × Treatment Interaction: Re-Analysis of a Study by G. L. Anderson." *Journal of Educational Psychology* 67:717-24..

de Leeuw, Jan and Ita G. G. Kreft. 1986. "Random Coefficient Models for Multilevel Analysis." *Journal of Educational Statistics* 11:57-85.

Dempster, Arthur, Nan Laird, and Donald Rubin. 1977. "Maximum Likelihood for Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society* B39:1-38.

Durkheim, Emile. 1951. *Suicide*. Glencoe, IL: Free Press. (Original work published 1898)

Galtung, Johan. 1969. *Theory and Methods of Social Research*. New York: Columbia University Press.

Goldstein, Harvey. 1987. *Multilevel Models in Educational and Social Research*. New York: Oxford University Press.

———. 1991 "Non-Linear Multilevel Models, with an Application to Discrete Response Data." *Biometrika* 78:45-51.

Goldstein, Harvey and R. Silver. 1989. "Multilevel and Multivariate Models in Survey Analysis." Pp. 221-35 in *The Analysis of Complex Surveys*, edited by C. Skinner, D. Holt, and F. Smith. New York: Wiley.

Hannan, Michael T. 1971. *Aggregation and Disaggregation in Sociology*. Lexington, MA: D. C. Heath.

Harville, David. 1977. "Maximum Likelihood Approaches to Variance Component Estimation and Related Problems." *Journal of the American Statistical Association* 72:320-40.

Hauser, Robert M. 1970. "Context and Consex: A Cautionary Tale." *American Journal of Sociology* 75:645-63.

Hays, William L. 1973. *Statistics for the Social Sciences*. New York: Holt, Rinehart & Winston.

Hox, Joop J., Ita G. G. Kreft, and Piet L. J. Hermkens. 1991. "The Analysis of Factorial Surveys." *Sociological Methods & Research* 19:493-510.

Iversen, Gudmund R. 1991. *Contextual Analysis*. Newbury Park, CA: Sage.

Jöreskog, K. G. and D. Sörbom. 1989. *LISREL 7. A Guide to the Program and Applications*. Chicago: SPSS Inc.

Kreft, Ita G. G., Jan de Leeuw, and Kyung-Sung Kim. 1990. *Comparing Four Different Statistical Packages for Hierarchical Linear Regression: GENMOD, HLM, ML2, and VARCL* (CSE Technical Report #310). Los Angeles, CA: University of California Press.

Lazarsfeld, Paul F. and H. Menzel. 1961. "On the Relation Between Individual and Collective Properties." Pp. 422-40 in *Complex Organizations*, edited by A. Etzioni. New York: Holt, Rinehart & Winston.

Longford, Nick T. 1987. "A Fast Scoring Algorithm for Maximum Likelihood Estimation in Unbalanced Mixed Models With Nested Random Effects." *Biometrika* 74:817-27.

———. 1988. *A Quasilikelihood Adaption for Variance Component Analysis*. Princeton, NJ: Educational Testing Service.

———. 1990. *VARCL. Software for Variance Component Analysis of Data With Nested Random Effects (Maximum Likelihood)*. Princeton, NJ: Educational Testing Service.

Mason, William M., A. F. Anderson, and N. Hayat. 1988. *Manual for GENMOD*. Ann Arbor: University of Michigan, Population Studies Center.

Mason, William M., George Y. Wong, and Barbara Entwisle. 1984. "Contextual Analysis Through the Multilevel Linear Model." *Sociological Methodology* 72-103.

McCullagh, Peter and J. A. Nelder. 1989. *Generalized Linear Models*. London: Chapman and Hall.

Mislevy, Robert J. and R. Darrel Bock. 1989. "A Hierarchical Item-Response Model for Educational Testing." Pp. 57-74 in *Multilevel Analysis of Educational Data*, edited by R. D. Bock. San Diego, CA: Academic Press.

Muthén, B. O. 1987. *LISCOMP. Analysis of Linear Structural Equations With a Comprehensive Measurement Model*. Mooresville, IN: Scientific Software Inc.

Numerical Algorithms Group. 1985. *The GLIM Release 3.77 Manual*. Downers Grove, IL: Author.

Prosser, Robert, Jon Rasbash, and Harvey Goldstein. 1991. *ML3. Software for Three-Level Analysis. Users' Guide*. London: University of London, Institute of Education.

Roberts, K. and Leigh Burstein. 1980. *New Directions for Methodology for Social and Behavioral Sciences*. Vol. 6. San Francisco: Jossey-Bass.

Robinson, W. S. 1950. "Ecological Correlations and the Behavior of Individuals." *American Sociological Review* 15:351-57.

Searle, S. R., G. Casella, and C. E. McCulloch. 1992. *Variance Components*. New York: Wiley.

Van den Eeden, Pieter and Harry J. M. Huettner. 1982. "Multilevel Research." *Current Sociology* 30:1-181.

Wong, George Y. and William M. Mason. 1985. "The Hierarchical Logistic Regression Model for Multilevel Analysis." *Journal of the American Statistical Association* 80:513-24.

Joop J. Hox is an associate professor of methods and statistics at the Department of Education, University of Amsterdam, the Netherlands. He is a staff member of the Dutch Interuniversity

Graduate School of Psychometrics and Sociometrics and teaches courses on multivariate methods and modeling at the Graduate School of Education of the University of Amsterdam. His main research interests are multilevel modeling, survey research methodology, and problems of operationalization and measurement. He has edited several books and published in both Dutch and international journals.

Ita G. G. Kreft is an associate professor at the School of Education, California State University, Los Angeles. Her research interests are the teaching of statistics and multilevel analysis. She has published in Sociology of Education, Sociological Methods & Research, *and the* Journal of Educational Statistics.